Vrije Universiteit Brussel

Faculteit Wetenschappen en Bio-ingenieurswetenschappen
Vakgroep Fysica

Measuring the anomalous couplings in the Wtb vertex using the Matrix Element Method at the LHC

Annik Olbrechts

Promotor: Prof. Dr. Jorgen D'Hondt

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

February 2016

Doctoral examination commission:

Prof. Dr. J. D'Hondt (Vrije Universiteit Brussel)
Prof. Dr. B. Craps (Vrije Universiteit Brussel)
Prof. Dr. F. Blekman (Vrije Universiteit Brussel)
Prof. Dr. F. Maltoni (Université Catholique de Louvain)
Prof. Dr. N. Van Eijndhoven (Vrije Universiteit Brussel)
Prof. Dr. P. Vanlaer (Université Libre de Bruxelles)
Dr. P. Van Mulders (Vrije Universiteit Brussel)
Prof. Dr. W. De Meuter (Vrije Universiteit Brussel)

Contents

Introduction

Achieving insight in the most fundamental aspects of our universe requires a detailed knowledge of the building blocks of all matter. Hence a thorough understanding of all the fundamental particles and interactions amongst them is of crucial importance. These concepts are studied meticulously in the field of elementary particle physics and have been summarised in a theory denoted the Standard Model. This theory has been experimentally verified with great precision and all obtained result are in excellent agreement with the provided predictions. One of the more outstanding achievements of the Standard Model was realised in the summer of 2012 with the observation of the Brout-Englert-Higgs boson, predicted in 1964. Even though the Standard Model can be considered as a very successful theory, it still has a couple of short-comings and can not be considered as a "theory of everything". Hence one should continue to experimentally confirm the various predictions made by the Standard Model.

In order to achieve the challenging conditions required to fully test the Standard Model, high-energetic particle accelerators such as the Large Hadron Collider at CERN near Geneva are constructed. This is a 27 km long circular particle accelerator designed to deliver proton-proton collisions at a record-breaking centre-of-mass energy of 14 TeV. For the measurement discussed in this thesis, the considered top-quark pair events were produced during the 2012 run, which operated at a centre-of-mass energy of 8 TeV, and were collected by the Compact Muon Solenoid experiment.

In this thesis, such a consistency test of the Standard Model will be performed since it will be investigated whether the interaction vertex of the top-quark pair decay corresponds with the predicted left-handed description and is not influenced by new-physics phenomena. This will be achieved by directly measuring the best estimate of one of the coupling coefficients in this so-called Wtb interaction vertex using a Matrix Element method.

In Chapter 1 an overview of the theoretical framework of the Standard Model will be given together with the specific details related to the top-quark interaction vertex. The experimental setup of the CERN accelerator complex will be discussed in Chapter 2. An overview of the generation and simulation of proton-proton collisions is given in Chapter 3 together with the specific reconstruction algorithms. These algorithms describe how the electronic signals of the CMS detector are translated into actual physical objects such as electrons and muons. The selection of top-quark events and the reconstruction of the desired event topology is discussed in Chapter 4. In Chapter 5 the technique used to measure the right-handed tensor coupling of the Wtb interaction is introduced and the necessary normalisations and calibrations required to correctly apply this technique are discussed. In Chapter 6 the actual measurement

is performed using this Matrix Element method and the relevant uncertainties are discussed. Finally in Chapter 7 the agreement of the obtained result with both the theoretical expectation and the currently available exclusion limits is given.

The analysis discussed in detail in Chapters 5 and 6 corresponds to the main subject of my PhD research and provides the first direct and most accurate measurement of the right-handed tensor coupling. In addition, I also contributed to the measurement of the W-boson helicity fractions using top-quark pair events produced during the 2011 run, which has been published [1].

Chapter 1

Anomalous couplings in the top-quark sector

In elementary particle physics the ultimate goal is to achieve a complete and profound understanding of the most fundamental constituents surrounding us, and preferably doing so using a single theory. The quest for such a "theory of everything" led to the development of the Standard Model of elementary particle physics, which symbolises the first step towards a grand unification of all fundamental interactions. The most important aspects of the Standard Model, particle content and theoretical framework, will be discussed in detail in Section 1.1.

Further emphasis will be put on the specific characteristics of the top quark, the heaviest quark existing within the Standard Model, which plays an important role in assessing the Standard Model at high energies. Hence this sector is extremely interesting to search for influences of new physics phenomena, for example in the form of anomalous couplings as will be explained in Section 1.2.

1.1 Standard Model of elementary particle physics

The extensive inquiry performed during the 20^{th} century to define the elementary particles and the corresponding discoveries continuously altered the understanding of the fundamental interactions. Every new achievement left the physics community divided and often required the development of a brand new framework capable of describing the observations. Hence the Standard Model, officially established in the early 1970s, actually consists of many ingenious contributions from many renown physicists [2–4]. For the last decades the general view on elementary particles is believed to be rather stable, especially since every new discovery validated the Standard Model.

1.1.1 Particle content

Within the Standard Model (SM) the elementary particles are categorised based on their spin. Fermions, containing both leptons and quarks, have half-integer spin while bosons, also called force mediators, have integer spin. The collection of fermions can be stored into three separate generations, characterised by increasing mass, as shown

in Table 1.1. Each fermion f has an antiparticle, which is defined to have identical mass but opposite electrical charge and is generally denoted as $\bar{f}$. Only for the charged leptons, l^-, the notation l^+ is used for their respective antiparticle.
Even though the Standard Model consists of three fermion generations the first one is sufficient to describe all stable matter around us. This because a single atom consists of an electron circulating around a proton-neutron nucleus, which are bound states of up- and down-quarks with respective quark-content *uud* and *udd*. Starting from such an atom every known chemical element can be formed.

Table 1.1: Overview of the fermions in the Standard Model and their corresponding electrical charge.

Generation	**Quarks**		**Leptons**	
1^{st}	up u	down d	electron neutrino ν_e	electron e^-
2^{nd}	charm c	strange s	muon neutrino ν_μ	muon μ^-
3^{rd}	top t	bottom b	tau neutrino ν_τ	tau τ^-
Electrical charge	+2/3	-1/3	0	1

Separating fermions into leptons and quarks is motivated by the different fundamental forces they interact with. The Standard Model comprises three of the four fundamental interactions: the electromagnetic force responsible for holding electrons to nuclei in atoms and binding atoms into molecules, the weak force used for describing radioactive decay of subatomic particles and finally the strong interaction which ensures the stability of ordinary matter by confining quarks into hadrons. The only missing piece of the puzzle is gravity which is unfortunately not yet included. The leptons only interact through the weak and electromagnetic force although the neutral ones, the neutrinos, are obviously not influenced by the latter one. The quarks on the other hand are, besides by the two previously mentioned interactions, also affected by the strong force.
Within the Standard Model force mediation is represented by the exchange of a spin-1 boson, which have all been summarised in Table 1.2. The number of force mediators belonging to a specific interaction is determined from the underlying theoretical framework and depends on the type of charge mediated by the boson. The electromagnetic interaction is described by a single force mediator, while the weak one has three massive bosons and the strong force even contains eight gluons. The large number of gluons follows from the fact that for the strong interaction each quark occurs in three different colours. Since this colour charge is mediated during the interactions, eight possible gluon combinations exist.

A final, but definitely not less important, boson which is incorporated in the Standard Model is the spin-0 Brout-Englert-Higgs boson. This particle is responsible for providing mass to all other particles through the mechanism of electroweak symmetry breaking, as will be explained in Section 1.1.2. Its existence was postulated in 1964 but was only discovered rather recently in 2012 [6, 7]. Both its mass and spin have

Table 1.2: Overview of the spin-1 force-carriers in the Standard Model and their mass [5].

Force	**Boson**		**Mass** (GeV)
Strong force	gluon	g	0
Electromagnetic force	photon	γ	0
Weak force	W-boson	$W^{\pm}$	80.385 ± 0.015
	Z-boson	Z^0	91.1876 ± 0.0021

been verified experimentally with much detail [8, 9], and the obtained mass value is in perfect agreement with the Standard Model predictions:

$$m_{BEH} = 125.09 \pm 0.21\,(stat.) \pm 0.11\,(syst.)\,\text{GeV} \tag{1.1}$$

1.1.2 Interactions through gauge invariance

The Standard Model goes much further than merely being an exhaustive collection of elementary particles, the supporting theoretical framework is that of a relativistic quantum field theory. From this mathematical description, based on gauge invariance under the fundamental forces, follow the interactions between the fermions and bosons in an automatic manner. This will be illustrated for invariance under a general local gauge transformation since each interaction follows the same principle.

As mentioned before, fermions are half-integer spin particles and can thus be represented by a Dirac spinor field:

$$\mathcal{L}_{Dirac} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi \tag{1.2}$$

The imposed local gauge invariance requires the fermion fields, and the overall Lagrangian, to be invariant under the following general transformation:

$$\psi \rightarrow U(x)\psi = \exp\left(-i\vec{\alpha}(x)\cdot\frac{\vec{\tau}}{2}\right)\psi \tag{1.3}$$

where $\vec{\alpha}$ are the rotation parameters in the symmetry group represented by the Lie group generators $\vec{\tau}$[1]. Invariance of the Dirac Lagrangian under this transformation can only be accomplished by replacing the partial derivative ∂_{μ} by a covariant derivative D_{μ}. This however comes at the price of introducing new gauge fields A_{μ}, which interact with the fermion fields with coupling strength g.

$$D_{\mu} = \partial_{\mu} - ig\vec{A}_{\mu}\cdot\frac{\vec{\tau}}{2} \tag{1.4}$$

[1]It should be noted that these rotation parameters depend on x since this implies the considered transformation is a local one. It is precisely the presence of local gauge transformation that will introduce the fermion interactions.

Inserting this covariant derivative results in an additional term in the Dirac Lagrangian, which describes the interaction between the fermion fields ψ mediated by the gauge fields A_μ. Since the covariant derivative should transform under the gauge transformation as the fermion fields, the local changes are incorporated by this vector field.

$$\mathcal{L}_{Dirac} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi + g\bar{\psi}\gamma^\mu\psi\vec{A}_\mu \cdot \frac{\vec{\tau}}{2} \tag{1.5}$$

Fundamental fermion interactions in the Standard Model

In the above explanation the introduced matrix $U(x)$ has been defined as the most general rotation matrix of the symmetry group SU(N). This procedure can however be simplified in order to obtain the three gauge interactions of the Standard Model.

Quantum chromodynamics gauge transformations

As mentioned before, the strong interaction is represented by the quantum number colour such that each quark has three equivalent states. Therefore the fermion fields should be seen as a three-component column vector such that the quantum chromodynamics (QCD) symmetry group is SU(3) containing eight gauge fields G^a_μ. The generators τ in Equation (1.3) are for this symmetry group the Gell-Mann matrices λ^a_i, giving the covariant derivative of the strong interaction the following form:

$$D_\mu = \partial_\mu - ig_S\frac{\lambda^a}{2}G^a_\mu \tag{1.6}$$

where g_S is the coupling constant of the strong interaction.
The three-component or triplet representation is only valid for particles carrying colour charge, otherwise they are merely $SU(3)_C$ singlets.

Electroweak gauge theory

The electroweak interaction combines the electromagnetic and weak theory and should be able to explain the parity violation observed in the latter. The smallest group capable of doing so is $SU(2)_L \times U(1)_Y$ where the subscript L stands for left-handed[2] and Y for the weak hypercharge. The overall covariant derivative which should be used for the electroweak interaction is thus:

$$D_\mu = \partial_\mu - ig\frac{\tau_i}{2}W^i_\mu - ig'\frac{Y}{2}B_\mu \tag{1.7}$$

where g and g' are the respective coupling strengths of the weak and electromagnetic interaction and τ_i the Pauli matrices.

This gauge invariance introduces a total of four gauge fields, three from the $SU(2)_L$ transformations and one from the $U(1)_Y$ one. The physical gauge fields, the electromagnetic photon A_μ and the weak vector bosons, $W^\pm_\mu$ and Z^0_μ,

[2] Left-handed and right-handed fermions can be distinguished using the left-handed and right-handed operator $P_{L,R} = \frac{1}{2}(1 \mp \gamma_5)$ with γ_5 defined as the fifth gamma matrix ($\gamma_5 = \mathrm{i}\gamma_0\gamma_1\gamma_2\gamma_3$).

can be derived from these gauge fields in the following way:

$$\begin{aligned} A_\mu &= W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W \\ W_\mu^\pm &= \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp i W_\mu^2\right) \\ Z_\mu &= W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W \end{aligned} \tag{1.8}$$

The angle θ_W used in these equations is the weak mixing or Weinberg angle, defined as:

$$\tan\theta_W = \frac{g'}{g} \tag{1.9}$$

Only the left-handed fermions can be represented as a doublet in SU(2) while the right-handed ones are again singlets and therefore do not interact with the gauge fields W_μ^i.

The abelian or non-abelian nature of the underlying gauge group signficantly influences the characteristics of the introduced gauge fields. Only in the latter case self-interactions among the gauge fields are allowed, as is the case for the gluons and the three vector bosons. The photon on the other hand is not able to have any self-interactions.

Electroweak symmetry breaking

The mathematical framework of gauge invariance explains in detail the interactions of the fermions and bosons, their mass acquirement however remains a big mystery. Simply introducing a bosonic mass term of the form $m^2 A_\mu A^\mu$ would violate gauge invariance. The same even holds for a fermionic mass term, $m_f \psi\bar{\psi}$, which would violate the SU(2)×U(1) symmetry because of the different transformation rules for right- and left-handed fermions.
Nevertheless, observations of massive fermions and bosons indicate that the Standard Model, in order to remain trustworthy, should be extended in a way to accommodate mass terms for both particle types. A solution is given by the principle of spontaneous symmetry breaking, known as the Brout-Englert-Higgs (BEH) mechanism [10–12], postulated in 1964. It introduces a single scalar doublet which leaves the Lagrangian invariant but breaks the ground state of the vacuum.

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{1.10}$$

The Lagrangian of this BEH field can contain the following gauge-invariant terms:

$$\begin{aligned} \mathcal{L}_{BEH} &= (D^\mu\phi)^\dagger (D_\mu\phi) - V(\phi) \\ &= (D^\mu\phi)^\dagger (D_\mu\phi) - \mu^2(\phi^\dagger\phi) - \lambda(\phi^\dagger\phi)^2 \end{aligned} \tag{1.11}$$

where μ^2 and λ (> 0) are two real values representing a mass parameter and the scalar self-interaction strength, respectively.

In case the mass parameter is positive the potential only has the trivial minimum at $\phi = 0$ and Equation (1.11) simply describes a massive scalar particle with mass μ and quartic coupling strength λ. However if the mass parameter is negative the situation is rather different since a non-unique vacuum state is retrieved for the potential resulting in spontaneous symmetry breaking once a vacuum expectation value is chosen.

$$\left\langle \phi^\dagger \phi \right\rangle = v^2 = \frac{|\mu^2|}{\lambda} \tag{1.12}$$

In order to study the particle spectrum in the theory small perturbations around this minimum should be considered:

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{1.13}$$

From the four original fields of the scalar doublet only one remains: the BEH field H. The three other real fields have been absorbed by the massless vector fields of the weak interaction converting them into massive fields. The BEH boson H^0, originating from the BEH field, acquires itself a mass $m_H = \sqrt{2\lambda}v$ while the photon remains massless. The mass of the vector bosons of the weak interaction is given by:

$$M_W = \frac{1}{2}vg \qquad\qquad M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} \tag{1.14}$$

The principle of electroweak symmetry breaking illustrates elegantly how the bosons acquire mass within the Standard Model, but no mass term for the fermions is yet included. Their mass, however, also follows from the same BEH mechanism but in a slightly less trivial manner.
The existence of the additional BEH field ϕ allows for the introduction of the following gauge-invariant terms in the Lagrangian:

$$\mathcal{L}_{Yukawa} = -Y_{ij}\bar{\psi}_{L,i}\phi\psi_{R,j} + h.c. \tag{1.15}$$

with Y_{ij} the unknown Yukawa matrices. Hence the fermion masses arise from the Yukawa interactions describing the couplings of the fermions with the BEH field.

For the quarks the weak-interaction eigenstates, considered up to now, have been observed to differ slightly from the mass eigenstates. Hence a matrix conversion is required which diagonalises the mass matrix. This is done by the 3×3 Cabibbo-Kobayashi-Maskawa (CKM) matrix [13, 14], which represents the probability of a transition from a quark q into a quark q' by the matrix element $|V_{q'q}|$.

$$\begin{pmatrix} d^{weak} \\ s^{weak} \\ b^{weak} \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.16}$$

1.1.3 Unanswered questions in the Standard Model

The Standard Model is regarded as an extremely successful theory, which has been experimentally verified up to the percent level or better [5]. However, it still bares some important shortcomings which cannot be ignored and should be understood in order to denote the Standard Model as a "theory of everything".

Grand Unified Theory
The successful unification of the weak and electromagnetic interaction into the electroweak one sparked hope of one day representing the three forces of the Standard Model by a single one. Such a unification of the elementary forces is currently not yet explicable by the Standard Model since it requires new physics at a very high energy scale ($\Lambda_{GUT} \sim 10^{16}\,\text{GeV}$). This Grand Unified Theory (GUT) is believed to be a first step towards the incorporation of gravity in the Standard Model.

Hierarchy problem
The measured vector boson masses indicate that the principle of electroweak symmetry breaking should occur at an energy scale of $\mu^2 \sim (100\,\text{GeV})^2$. The large energy gap up to the GUT or Planck scale ($\Lambda_{Planck} \sim 10^{19}\,\text{GeV}$), where the latter is the energy regime where the gravitational attraction becomes comparable to the other fundamental interactions, requires a significant fine-tuning of at least 28 orders of magnitude. A scale difference of this extent, also known as the hierarchy problem, is far from desirable for any theory.

Dark matter and energy
Cosmological observations have pointed out that the matter described by the Standard Model only constitutes about 4.8% of the matter in the universe [15]. The remaining part is occupied by dark matter (25.8%) and dark energy (69.4%), two cosmological concepts that are extremely challenging to detect.

Neutrino masses
The fact that neutrinos are massless weakly interacting particles is embedded in the theory of the Standard Model but recent observations by the Sudbury Neutrino Observation [16] and Super-Kamiokande [17] collaboration have offered clear evidence that neutrinos can convert from one flavour to another. This implies at least one neutrino species has non-zero mass requiring the introduction of a lepton mixing matrix, by analogy with the quark mixing one, known as the Pontecorvo-Maki-Nakagawa-Sakata matrix [18, 19].

The shortcomings outlined above are by no means exhaustive but have been listed in order to indicate that the Standard Model should not be regarded as a theory capable of describing all aspects of experimental physics. However, these couple of unsettled issues do not weigh up to the numerous successes of the Standard Model, which are deemed as firm enough to incorporate new physics phenomena through theoretical extensions.
Supersymmetry is one of the more widely accepted suggestions, which introduces an additional symmetry relating bosons and fermions. In this framework each particle has a superpartner with identical quantum numbers, except for the spin parameter differing

by a half-integer. Supersymmetry close to the electroweak scale would provide a natural solution for the hierarchy problem without requiring the significant fine-tuning. As a bonus the lightest supersymmetric particle is a possible dark matter candidate.
Other extensions are considered as well, for example by including additional dimensions or substituting particles by strings. These type of theories require a significant number of cosmological observations rendering the theory certification more challenging. Hence in order to decide upon the correct Standard Model extension much more experimental data of both cosmological and elementary particle physics processes are required. The latter can be achieved by constructing state-of-the-art particle colliders.

1.2 Importance of the top-quark interaction vertex

One of the regions of interest to look for new physics phenomena in elementary particle physics is the top-quark sector. This elementary particle, the heaviest one in the SM, has been discovered more than 20 years ago, in 1995, but nevertheless remains a challenging research subject due to its important role in beyond the Standard Model theories.
The high mass of the top quark has rendered its observation very arduous since it requires extreme energy conditions to be produced. On the other hand, it is exactly this high mass that makes the top quark such an interesting particle to investigate. This because the top quark is the only elementary particle for which the Yukawa coupling is of the order of 1, such that it is likely to assume the top quark might shed some light on the principle of electroweak symmetry breaking.
A detailed overview of the most recent and most important top-quark properties will be given here since a thorough understanding of the top-quark sector might help to constrain new physics models. Afterwards the specific details of the top-quark decay vertex, the Wtb one, will be discussed. Within this interaction vertex anomalous couplings can appear, which are currently only constrained indirectly. The way these anomalous couplings are connected with the measured properties of Section 1.2.1 will be discussed in detail in Section 1.2.2.

1.2.1 Measured top-quark properties

The energy regime required to produce the heavy top quark was first realised at the Tevatron collider [20], where it was finally discovered in 1995 by the CDF [21] and D∅ [22] experiments. However since a couple of years the Tevatron has been superseded by the Large Hadron Collider (LHC) at CERN [23] as top-quark factory, which can produce top quarks in ample amounts due to the higher centre-of-mass energy. At hadron colliders top quarks can be produced either in pairs or singly, although the former one is the more dominant production method.
The top-quark pair-production cross-section can be determined theoretically in a very precise manner and compared to the measured cross-sections at the LHC. These theoretical and experimental cross-section values have been summarised in Table 1.3, all providing for excellent agreement. Single production of top quarks also occurs but with lower probability and higher background, rendering detection significantly more

Table 1.3: Comparison between the theoretical predictions [24, 25] and experimental measurements [26–29] of the $t\bar{t}$ pair production cross-section $\sigma_{t\bar{t}}$.

		Theory prediction (pb)	Measured $\sigma_{t\bar{t}}$ (pb)
Tevatron:	$p\bar{p}$ at $\sqrt{s} = 1.96$ TeV	$7.164^{+0.391}_{-0.475}$	7.60 ± 0.41
LHC:	pp at $\sqrt{s} = 7$ TeV	$172.0^{+12.1}_{-13.4}$	174.5 ± 6.2
LHC:	pp at $\sqrt{s} = 8$ TeV	$245.8^{+16.6}_{-18.7}$	241.5 ± 8.5
LHC:	pp at $\sqrt{s} = 13$ TeV	832^{+40}_{-46}	825 ± 114

challenging. This production mechanism consists of three distinct processes: s-, t- and tw-channel, which have a different relative contribution at the Tevatron and the LHC since the latter one is dominated by processes with initial gluons.
Single top-quark physics has been an extremely challenging research subject over the last couple of years, especially since this production mechanism was only for the first time observed in 2009 [30, 31] by the CDF and D∅ collaborations. Moreover, evidence of the single top quark s-channel at the LHC was only observed in 2015 by ATLAS [32]. The single top-quark sector is of particular interest in order to accurately determine the CKM matrix element V_{tb} since the production cross-section is directly proportional to the squared of the matrix element. The latest results [33] clearly demonstrate that its value is close to unity, implying that the top quark decays predominantly into a W-boson and a bottom quark.

$$|V_{tb}| = 0.998 \pm 0.038(exp.) \pm 0.016(theo.) \quad \Rightarrow \quad |V_{tb}| > 0.92 \text{ @ } 95\% \; CL \tag{1.17}$$

Even though top-quark production can be governed either by strong or electroweak interactions, its decay is purely characterised by electroweak processes. As follows from the V_{tb} coefficient, this occurs with a branching probability close to 100% and results in a W-boson and a b-quark. The produced W-boson is known to be unstable and thus decays directly: either into a quark and anti-quark (mainly $W \to u\bar{d}$ or $c\bar{s}$) or into a charged lepton and corresponding neutrino ($W \to \ell^+\nu$ with $\ell = e$, μ or τ). Since the quarks can take three different colours there are 6 distinct hadronic decay channels such that the leptonic branching ratio of the W-boson is about 33% while the hadronic one is close to 67%.
Thus in the top-quark pair production context, the production process of interest for this thesis, three decay channels can be distinguished: the hadronic channel where both W-bosons decay into a quark-antiquark pair, the dilepton one where they both decay into a charged lepton and neutrino and finally the semileptonic channel where one of the two W-bosons decays hadronically and the other one leptonically. In this thesis the semi-leptonic (with $\ell = \mu$) top-quark decay-channel will be studied, which has a branching ratio of 14.8% per lepton flavour.

Searches have been performed at the LHC to ascertain that no other top quark decays occur, which would hint towards a discrepancy of the Standard Model. However all observations are in agreement with the expected rate, such as the flavor-changing

neutral-current top-quark decays ($t \to Zq$, Hc and Hu) for which the branching fractions are lower than 0.05%, 0.47% and 0.42% at 95% CL, respectively [34, 35]. Also the baryon-number violating decays $t \to \bar{b}\bar{c}\mu^+$ and $t \to \bar{b}\bar{u}e^+$ have been found to be heavily suppressed with an individual branching fraction smaller than 0.15% at 95% CL [36].

Besides the production cross-section is also the top-quark mass an extremely important parameter which can be used to constrain the electroweak sector of the Standard Model [37]. Prior to the discovery of the BEH boson a precise determination of the top-quark and W-boson mass allowed for the prediction of the BEH boson mass. Currently this electroweak fit procedure serves as an important consistency test of the Standard Model for which an accurate measurement of the top-quark mass is crucial [38].

$$m_t = 172.44 \pm 0.13(stat) \pm 0.47(syst)\,\text{GeV} = 172.44 \pm 0.48\,\text{GeV} \tag{1.18}$$

Due to its high mass the top quark has a very short lifetime, shorter than the QCD hadronisation scale, such that it can be studied as a free quark. As a result, the top-quark spin information is transmitted to its decay products allowing for a study of the top-quark spin correlations and the W-boson helicity fractions.
Within the Standard Model, there is a spin correlation between the spin of the top quark and the anti-top quark. Experimentally this is verified by determining the spin correlation relative to the SM prediction of f_{SM}, for which the current most precise result is in good agreement with the expectations [39]: $f_{SM} = 1.20 \pm 0.14$.
For the W-boson helicity fractions, the right-handed polarisation is supposed to be suppressed in the Standard Model. This measurement serves as an excellent test of the Standard Model and exploits the fact that the top-quark spin information is transfered to its decay products. The obtained results are compatible with the Standard Model expectations since the positive, longitudinal and negative polarisation are found to be -0.009 ± 0.021, 0.659 ± 0.027 and 0.350 ± 0.026, respectively [40].

1.2.2 Phenomenology of the Wtb interaction

Each measurement discussed before serves as an individual consistency check of the Standard Model and, in case no deviation from the Standard Model prediction is found, results in constraints on the related SM extensions. Actual searches for new physics signatures can be done in two distinct manners: either by looking for the production of new particles or by probing novel interactions of the known Standard Model particles. In this thesis the latter approach will be discussed, which allows to represent the overall interaction vertex through effective field theories (EFT) [41]. This way the Standard Model can be interpret as an effective low-energy theory influenced by higher dimensional interaction terms, or so-called anomalous couplings.

$$\mathcal{L}_{EFT} = \mathcal{L}_{SM} + \sum_i \frac{C_i}{\Lambda^2}\mathcal{O}_i \tag{1.19}$$

where c_i are the coupling coefficients of the SM fields to the new fields, Λ the characteristic scale of new physics and $\mathcal{O}_i$ the higher-order effective operators.

For the Wtb coupling this Lagrangian can be simplified significantly using the vertex-function approach because only a small fraction of the introduced operators influence the top-quark sector. At first the dimension-five operators can be neglected since their inclusion would result into lepton number violation. In addition, no operators higher than dimension-six will be considered since the new-physics effects are expected to be very small due to the current excellent agreement of the Standard Model with experimental measurements. Hence, assuming that the new couplings arise from the dimension-six operators allows to parametrise the general effective Wtb interaction vertex in the following way [42, 43]:

$$\begin{aligned}\mathcal{L}_{Wtb} &= -\frac{g}{\sqrt{2}}\bar{b}\gamma^{\mu}\left(V_L P_L + V_R P_R\right) t W^{-}_{\mu} + h.c. \\ &\quad -\frac{g}{\sqrt{2}}\bar{b}\frac{i\sigma^{\mu\nu}q_{\nu}}{m_W}\left(g_L P_L + g_R P_R\right) t W^{-}_{\mu} + h.c. \end{aligned} \tag{1.20}$$

where V_L, V_R, g_L and g_R are the complex coupling coefficients for the vector and tensor interaction, respectively. These coupling coefficients are related to the coefficients of the dimensional-six operators as [43]:

$$V_L = 1 + C^{(3)*}_{\phi q}\frac{v^2}{\Lambda^2}; \qquad g_L = \sqrt{2}C^{*}_{bW}\frac{v^2}{\Lambda^2}; \tag{1.21}$$

$$V_R = \frac{1}{2}C^{*}_{\phi\phi}\frac{v^2}{\Lambda^2}; \qquad g_R = \sqrt{2}C_{tW}\frac{v^2}{\Lambda^2}; \tag{1.22}$$

In the Standard Model limit both tensor couplings and the right-handed vector coupling vanish at tree-level, corresponding with the observed left-handedness of the weak interaction. The operators P_L and P_R are defined as the left- and right-handed operators representing the different treatment of left-handed and right-handed fermions in the weak interaction.

$$\psi_L = P_L\psi = \frac{1-\gamma_5}{2}\psi \quad \& \quad \psi_R = P_R\psi = \frac{1+\gamma_5}{2}\psi \tag{1.23}$$

This so-called parity violation in the weak force has been postulated in 1956 by T.D. Lee and C.N. Yang and soon after experimentally verified [44, 45]. The observed maximal parity violation could only be incorporated in the theoretical framework of the Standard Model by introducing a vector-axial (V-A) structure for the weak interaction that treats the right- and left-handed fermions differently. The left-handed behaviour of the weak interaction can be verified experimentally by measuring the W-boson helicity fractions since the right-handed polarisation is not allowed within the Standard Model. As mentioned before, excellent agreement with the Standard Model predictions has been obtained for this measurement.

Direct measurement of the Wtb coupling coefficients

A direct measurement of the coupling coefficients in the Wtb interaction vertex would also serve as an excellent test of this V-A structure of the weak interaction, but has not yet been accomplished. Indirect information from the measurement of the W-boson

helicity fractions [1, 46, 47] did allow to put limits on the values of the different coefficients as can be seen from Table 1.4 and Figure 1.1. The limits have been obtained from the combined $t\bar{t}$ dileptonic and semileptonic measurement executed by ATLAS while the visualisations have been taken from the single top-quark and semileptonic $t\bar{t}$ measurements of CMS. Purely looking at the limits obtained from the W-boson helicity fractions allows for a second region near $Re(g_R) \gg 0$, which completely contradicts the Standard Model predictions. However this solution is disfavoured by the measured cross-section for single top-quark production.
The different coupling coefficients are considered to be real since this follows from the restrictions when assuming CP conservation, which will also be the case for the measurement discussed in this thesis.

Table 1.4: Summary of the indirect limits obtained for the different anomalous coupling coefficients at 95% confidence level obtained by ATLAS using a 1.04 fb^{-1} data at 7 TeV [46].

$Re(V_R)$	$Re(g_L)$	$Re(g_R)$
[-0.20, 0.23]	[-0.14, 0.11]	[-0.08, 0.04]

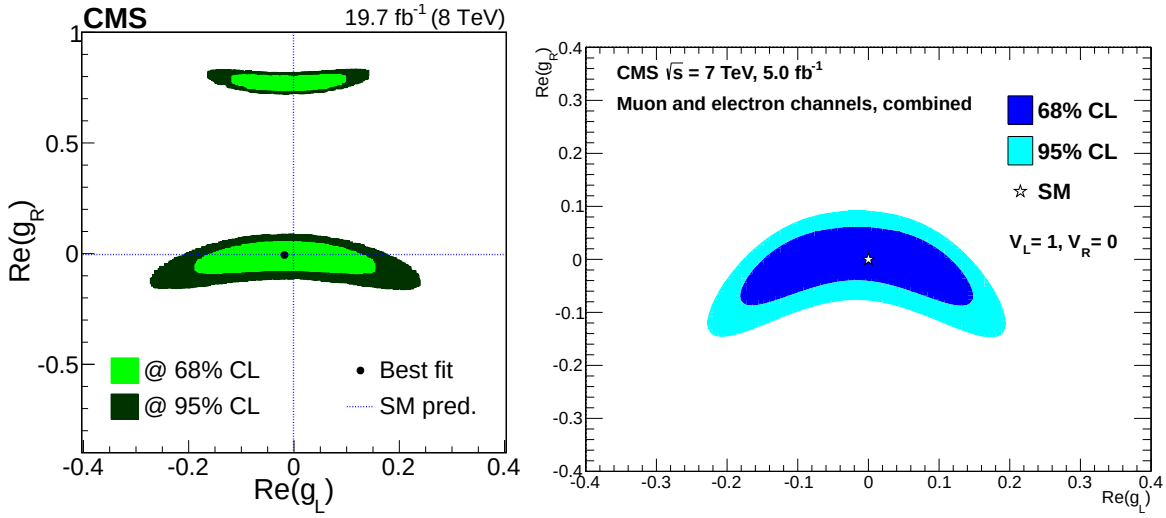

Figure 1.1: Exclusion limits for the real part of the anomalous tensor couplings while restraining the vector couplings to the SM predictions using the combined results of the single top-quark measurement (left) and the semileptonic $t\bar{t}$ pair one (right) [1, 47]. The latter one also has this second region near $Re(g_R) \gg 0$ but is not shown.

Performing such a direct measurement is a rather challenging procedure since it affects the entire decay process of the top quark and thus influences all decay products. To make matters even worse, there is no physical observable which allows to visualise the full impact of the different anomalous couplings. Hence the effect of the different coupling coefficients can only be predicted partially by looking at observables known to be sensitive to changes in the interaction vertex, such as the W-boson helicity fractions, the angular asymmetries and spin asymmetries.

Each of these observables has in common that they can be represented using angular distributions of the top-quark decay products allowing for a partially and simplified visualisation of the influence of the anomalous couplings. In this thesis the angle θ_l^* between the lepton and the reversed top quark, measured in the rest-frame of the W-boson, will be considered in order to outline the optimal analysis strategy to study the anomalous couplings in the Wtb interaction vertex. This observable has been used extensively for the measurement of the W-boson helicity fractions.

The benefit of using this angle θ_l^* and the connection with the W-boson helicity fractions is that the dependency of the top-quark decay on the coupling coefficients of the Wtb interaction can be visualised. This dependency allows to describe the partial widths of the top-quark decay, $\Gamma_{0,R,L}$, as a function of the four coupling coefficients [48]. Since these partial widths are directly related to the helicity fractions through $F_i = \Gamma_i/\Gamma$ the influence of each of the Wtb coupling coefficients can be visualised by looking at the relative change observed in the angular distribution. Hence using the formula given in Equation (1.24), which describes the normalised differential decay rate for unpolarised top quarks, the individual contribution of each coupling coefficient can be depicted.

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta_l^*} = \frac{3}{8}(1+\cos\theta_l^*)^2 F_R + \frac{3}{8}(1-\cos\theta_l^*)^2 F_L + \frac{3}{4}\sin^2\theta_l^* F_0 \qquad (1.24)$$

From the angular distributions shown in Figure 1.2 can be concluded that the angular distribution is not influenced in case only one coupling is set as non-zero. As a result the only way a direct measurement can be performed is by looking at the interplay of two such couplings. Hence it has been opted for in this thesis to restrain the V_L coefficient to its expectation value of the Standard Model, $V_L = 1$, while measuring one of the three remaining anomalous coupling coefficients.
The most relevant coupling coefficient to consider has also been identified by looking at these angular distributions and by carefully calculating the predictions of the theoretical connection with the decay width of the top-quark. These calculations indicate that the only coupling with a non-suppressed linear term is the right-handed tensor coupling g_R, which is also clearly the most sensitive one as can be seen from the bottom right plot in Figure 1.2.

Now that the most sensitive coupling coefficient has been identified, the significantly simplified formulas linking the width of the top-quark decay with the coupling of interest are given in Equation (1.25) to (1.27). The left-handed vector coupling is still present in these formulas but will be set equal to 1 when the measurement of this

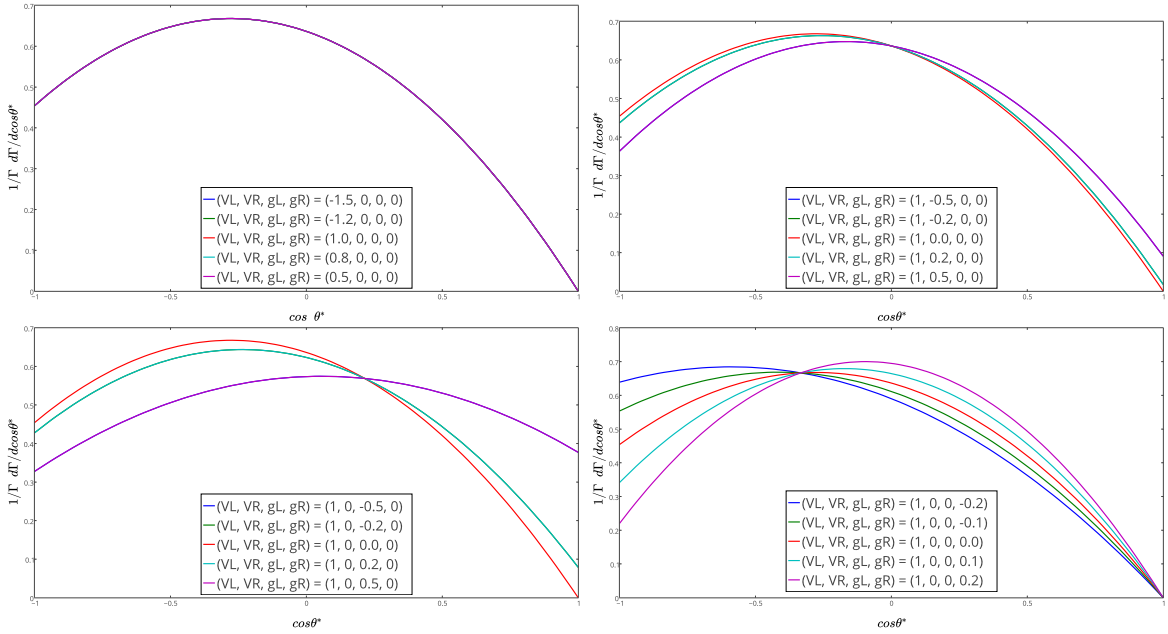

Figure 1.2: Influence of varying a single coupling coefficient on the $\cos\theta^*$ distribution. From upper left to upper right the considered coupling constants are V_L, V_R, g_L and g_R, respectively.

right-handed tensor coupling is performed.

$$\begin{aligned}\Gamma_0 &= \frac{g^2|\vec{q}\,|}{32\pi}\left\{\frac{m_t^2}{m_W^2}V_L^2(1-x_W^2-2x_b^2-x_W^2x_b^2+x_b^4)+g_R^2(1-x_W^2+x_b^2)\right.\\ &\quad\left.-2\frac{m_t}{m_W}V_Lg_R(1-x_W^2-x_b^2)\right\} \qquad (1.25)\\ \Gamma_{R,L} &= \frac{g^2|\vec{q}\,|}{32\pi}\frac{m_t^2}{m_W^2}\left\{V_L^2(1-x_W^2+x_b^2)-2\frac{m_t}{m_W}V_Lg_R(1-x_W^2-x_b^2)\right.\\ &\quad\left.+g_R^2(1-x_W^2-2x_b^2-x_W^2x_b^2+x_b^4)\right\}\\ &\quad\pm\frac{g^2}{64\pi}\frac{m_t^3}{m_W^2}\left\{-x_W^2V_L^2-g_R^2(1-x_b^2)+2x_WV_Lg_R\right\}\\ &\quad\times(1-2x_W^2-2x_b^2+x_W^4-2x_W^2x_b^2+x_b^4) \qquad (1.26)\\ \Gamma &= \frac{g^2|\vec{q}\,|}{32\pi}\frac{m_t^2}{m_W^2}\left\{V_L^2(1+x_W^2-2x_b^2-2x_W^4+x_W^2x_b^2+x_b^4)\right.\\ &\quad+g_R^2\left(2-x_W^2-4x_b^2-x_W^4-x_W^2x_b^2+2x_b^4\right)\\ &\quad\left.-6x_WV_Lg_R(1-x_W^2-x_b^2)\right\} \qquad (1.27)\end{aligned}$$

where the notation x_W stands for m_W/m_t and similarly x_b for m_b/m_t. The variable $|\vec{q}\,|$ is defined as:

$$|\vec{q}\,| = \frac{1}{2m_t}\sqrt{m_t^4+m_W^4+m_b^4-2m_t^2m_W^2-2m_t^2m_b^2-2m_W^2m_b^2} \qquad (1.28)$$

From these formulas is clearly visible that the dominant interference term, $V_L * g_R$, is not scaled down by the factor x_b such that it cannot be neglected, as is the case

for the other two coupling coefficients. Hence the reason why the right-handed tensor coupling is the most sensitive one of the three anomalous coupling coefficients present in the Wtb interaction vertex. As an extra advantage the results acquired from the g_R coefficient do not depend on the bottom-quark mass, which is often neglected compared to the W-boson and top-quark mass. Within the Wtb interaction vertex this so-called massless-b limit is assumed in various theoretical and experimental research studies even though it has a noteworthy effect on the other coupling constants.

The full details on the Wtb interaction vertex are slowly but surely being revealed due to several indirect constraints with outstanding precision. However these measurements are gradually being dominated by systematic uncertainties and thus require either more sensitive observables or otherwise a new analysis technique using a direct approach. The analysis discussed in this thesis will follow the second approach and will mainly focus on the right-handed tensor coupling since it is expected to be the most sensitive coupling. The followed experimental strategy, however, has been developed in a general manner and could be extended in order to measure the two other coupling coefficients as well.

Chapter 2

The CMS experiment at CERN's accelerator complex

The Standard Model of elementary particle physics, for which the main successes and shortcomings have been discussed extensively in Chapter 1, results in very precise predictions. However it is only acknowledged as an effective theory up to an energy scale of about 1 TeV. Physics beyond this energy is studied with specific high-energetic particle colliders such as the Large Hadron Collider (LHC) located at CERN (European Organization for Nuclear Research) near Geneva. The LHC provides proton-proton collisions at a record-breaking energy and is currently the world's most energetic particle collider.
Several different experiments surround the LHC, each with a specific physics goal ranging from general high-luminosity physics to dedicated plasma-studies and even long-lifetime neutrino interactions. In this chapter attention will mainly be devoted to the Compact Muon Solenoid (CMS) experiment, which is the LHC general-purpose experiment used for recording the data processed within this thesis.

2.1 The Large Hadron Collider

The need for a particle collider with the dimensions of the LHC was driven by a quest to understand the nature of the electroweak symmetry breaking, for which the Brout-Englert-Higgs mechanism is presumed to be responsible, and to investigate physics at the TeV scale. Such high energies can only be achieved when state-of-the-art technology is used for both accelerating and colliding the particles, and for recording the provided interactions.

When the design of the LHC machine was approved in 1994 it was decided to reuse the existing 26.7 km Large Electron Positron (LEP) tunnel, previously excavated in the 1980's and positioned between 45m and 150m below the Earth's surface. Avoiding the construction of a new tunnel was a huge cost-saver but presented some stringent limitations on the machine's design. For example the space limitation in the tunnel compelled the use of so-called twin-bore magnets where both proton rings are contained within a single magnet structure.
The LHC is designed to provide proton-proton collisions with a beam energy of 7 TeV

each, resulting in a centre-of-mass energy $\sqrt{s}$ of 14 TeV. This is a seven-fold energy increase compared to the previous most energetic particle collider, the Tevatron collider, which yielded proton-antiproton collisions between 1983 and 2011. In order to reach these extreme energy conditions the LHC exploits the presence of the extensive accelerator complex at CERN to gradually increase the beam energy.
When the proton beams are circulating within the LHC at the desired beam energy they can be steered in order to collide head-on in the dedicated interaction regions. Of the eight interaction regions existing in the LEP tunnel only four have actually been equipped with particle detectors for the LHC data-taking. The ATLAS [49] and CMS [50] experiments are the two largest ones and are intended as general-purpose detectors studying a broad range of physics processes while the ALICE [51] and LHCb [52] experiments search for a specific type of physics interactions. The ALICE experiment serves mainly for heavy-ion physics while LHCb is dedicated to heavy-flavour physics. Within this thesis data collected at the CMS detector during the first era of data-taking has been analysed, which started in March 2010 and continued until December 2012. These collisions did not take place at the design beam energy of 7 TeV but at a reduced beam energy of 3.5 TeV and 4 TeV for the 2010-2011 and 2012 data-taking, respectively, leading to proton-proton centre of mass energies of 7 and 8 TeV.

The LHC's design, driven by the LEP legacy

The decision to accelerate protons instead of electrons, as was the case for LEP, was lead by the extremely high beam energy desired for the parton interactions. Less massive particles loose a significant amount of their energy due to synchrotron radiation when travelling at high speed in a circular orbit. With velocities close to the speed of light, as foreseen at the LHC, this energy loss would represent a limiting factor for the final energy reachable when accelerating electrons. This effect is almost negligible for protons.
The physics goals set by the LHC also compelled to step away from the approach adopted at the Tevatron collider: providing collisions between protons and anti-protons. This because the desired number of collisions, represented by the luminosity $\mathcal{L}$, would not be attained when using anti-protons. The design luminosity was fixed to reach a record-breaking value of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, corresponding to around 1 billion interactions per second. Such a high rate of collisions is only feasible when the partons are confined into dense bunches of about 10^{11} particles. In order to fulfil these challenging conditions the interacting partons need to be produced in adequate amounts, which is not realisable for anti-protons. As a result is was decided to provide proton-proton interactions at the LHC.
The relatively small circumference of the LHC tunnel combined with the extremely high velocity of the circulating protons compelled the use of superconducting magnets. Although superconductivity was already used extensively in previous accelerators, the LHC pushed the existing technology to a new level by operating at a record-breaking temperature of 1.9 K resulting in the production of a 8.33 Tesla strong magnetic field. These superconducting dipole magnets, of which 1232 exist in the LHC, are responsible for bending the proton trajectory.
However, the use of particles with the same electric charge introduces an additional

challenge since two separate magnetic fields are required to bend the protons in opposite directions. For the LHC this lead to the development of twin-bore magnets consisting of two separate beam pipes each surrounded with individual superconducting coils within the same mechanical structure. A schematic cross-section of such an LHC dipole magnet is given in Figure 2.1.

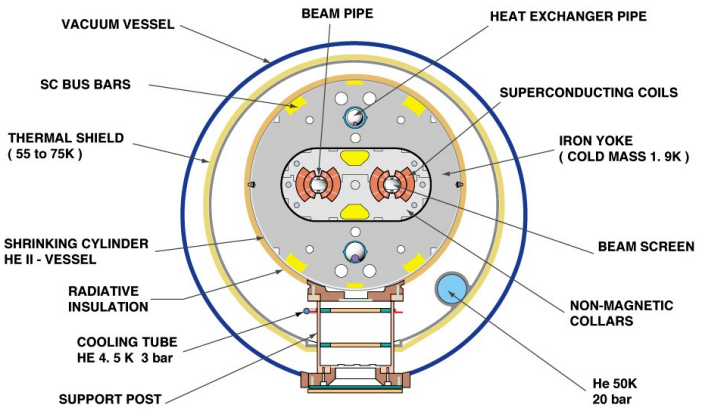

Figure 2.1: Cross-section of an LHC dipole magnet with the twin-bore design [53].

The tunnel of the Large Hadron Collider is not a perfect circle but consists of eight arcs and eight straight sections. The 15 m long dipole magnets are located in the arcs and are responsible for preserving the circular orbit of the protons. The actual acceleration of the protons is provided by the dedicated radio-frequency (RF) cavities, which are positioned in one of these straight sections. A second type of magnets, so-called quadrupole magnets, are located throughout the entire tunnel and ensure that the beams remain aligned and do not drift apart. These quadrupole magnets are able to squeeze the beam either vertically or horizontally and are also installed just before the different detectors where they provide an additional squeezing to increase the chances on a collision.

The LHC injection chain

The LHC does not only reuse the existing LEP tunnel it also benefits strongly from the complete accelerator complex present at CERN in order to reach the record energy of 7 TeV. The protons run through a series of interconnected linear and circular accelerators and are only passed on to the next in line once they attain the maximum speed possible for that specific accelerator. A schematic overview of the sequence used for the proton acceleration at the LHC is given in Figure 2.2.
The entire LHC injection chain starts with a box of hydrogen atoms from which protons are electrically stripped and accelerated to an energy of about 50 MeV when passing through the Linac2. Following to this linear acceleration, the protons are injected into the first circular accelerator, the Proton Synchrotron Booster (PSB), where they remain until reaching an energy of 1.4 GeV. Once the protons are energetic enough they will be shot into the Proton Synchrotron (PS) and become accelerated to 25 GeV, the maximum speed reachable by the PS. Afterwards they are finally inserted in the last

accelerator of the injection chain, which is the Super Proton Synchtrotron (SPS). The SPS is responsible for accelerating the protons up to 450 GeV after which the LHC is capable of accelerating them further until the nominal energy of 7 TeV.

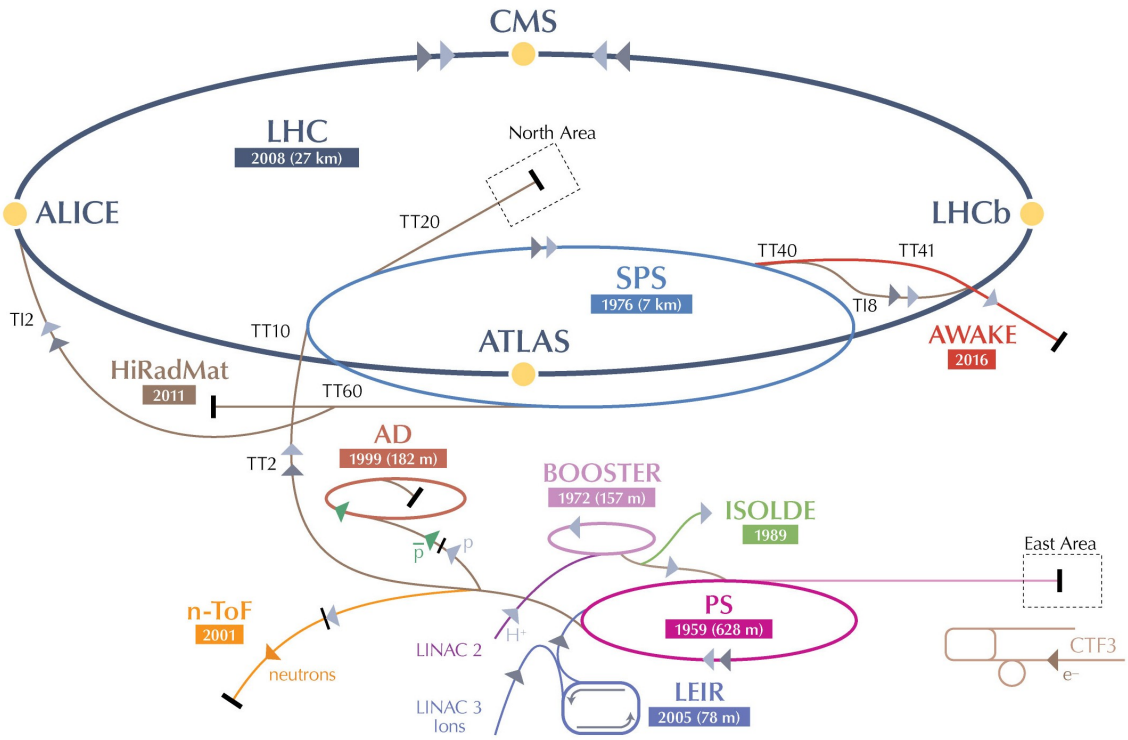

Figure 2.2: Detailed overview of the LHC injection chain [54].

The protons accelerated at the LHC are confined in large proton bunches since this significantly increases the possibility of observing at least one proton-proton interaction during a single bunch crossing. The design of the LHC is optimised for providing *pp* collisions with not only a record-breaking beam energy, but also with the highest luminosity ever recorded. In order to reach the design luminosity the LHC should store multiple proton bunches in a so-called bunch train with the smallest separation possible. This is a common technique adopted in most accelerators with the important difference that the LHC can reach much smaller bunch separations than ever before, up to 25 ns.

The proton bunches are created by the RF cavities responsible for accelerating the protons, a process which starts already during the first stages of the LHC injection chain. These type of cavities produce a resonant electromagnetic wave that oscillates at a given frequency and accelerates the protons up to an ideal energy. Hence inverting the electromagnetic field ensures that the protons become organised into discrete packets since protons travelling too fast will undergo a deceleration while those arriving late will feel an additional push. The adopted accelerator sequence is constructed such that each successive accelerator is capable of containing more bunches allowing a gradual filling of the bunch train.

Particle detectors

The Large Hadron Collider has dedicated areas in the tunnel where beam crossings are provided and particle detectors are placed in order to measure the particle activity present during proton-proton collisions. The four interaction regions utilised by the main LHC experiments CMS, LHCb, ATLAS and ALICE are represented as yellow blobs in Figure 2.2.
As shortly mentioned before the two largest experiments are the two general-purpose experiments CMS and ATLAS. They are both designed to cover a wide range of physics processes and are constructed following an onion-layered structure around the interaction point to avoid particles escaping detection. More detail about the CMS detector and the different layers it contains will be discussed in Section 2.2.
The two smaller experiments, ALICE and LHCb, tend to focus on a dedicated type of physics processes and are optimised as such. The ALICE experiment is a heavy-ion detector which is interested in the Pb-Pb collisions also provided by the LHC machine. It aims to gather information about the quark-gluon plasma, a phase of matter where the quarks and gluons are no longer confined in hadrons. The LHCb experiment focuses on heavy flavour physics and will try to explain why the universe seems to constitute almost entirely of matter and not of anti-matter. The design of this detector is rather distinctive since it does not follow the standard shape of an enclosed detector positioned symmetrically around the interaction region, but instead consists of a half-open structure designed to accurately measure forward particles.
The CERN site contains many other smaller experiments which try to grasp the physics behind some of the remaining unsolved mysteries and hope to result in useful future applications. Specific for the LHC tunnel are the three experiments which are located in the vicinity of one of the main particle detectors sharing the same experimental cavern. The first one is the TOTEM [55] experiment, which is placed close to the CMS detector along the beampipe and is designed to investigate the proton structure in the very forward region. The LHCf [56], installed near ATLAS, will also focus on physics in the very forward region, but aims to better understand hadron interaction models used in the simulation of high-energy cosmic rays. The last experiment, MOEDAL [57] is deployed around the LHCb interaction region and will try to detect the magnetic monopole, a hypothetical particle with magnetic charge.

Run-I (2010-2012) data taking

The LHC is designed to provide proton-proton collisions with a centre-of-mass energy $\sqrt{s} = 14$ TeV and the first proton beams made circulating in September 2008 were indeed configured for these conditions. However a mechanical incident on 19 September 2008 delayed the observation of the first LHC collisions to March 2010 and forced the LHC to run at a reduced energy of 3.5 TeV per beam. The same beam energy was kept during the 2011 run and was only slightly increased during the final year of the Run-I data taking period in order to provide collisions at $\sqrt{s} = 8$ TeV. In between the proton-proton runs dedicated ion-ion or proton-ion collisions were scheduled for the specific heavy-ion studies performed at the different experiments.
The number of collisions produced during the proton-proton collisions can be rep-

resented by the integrated or instantaneous luminosity, expressed in terms of pb^{-1}. The difference between the integrated and instantaneous luminosity lies in the considered time span. The instantaneous luminosity is the number of collisions provided by the LHC each second, while the integrated luminosity is the accumulated number of interactions during a longer period of time. Figure 2.3 shows the integrated luminosity delivered to the CMS experiment for *pp* collisions for the three years long Run-I data-taking period. Summing the daily contributions results in a total value for the integrated luminosity during the three years of Run-I of respectively, 45.0 pb^{-1}, 6.1 fb^{-1} and finally 23.3 fb^{-1}.

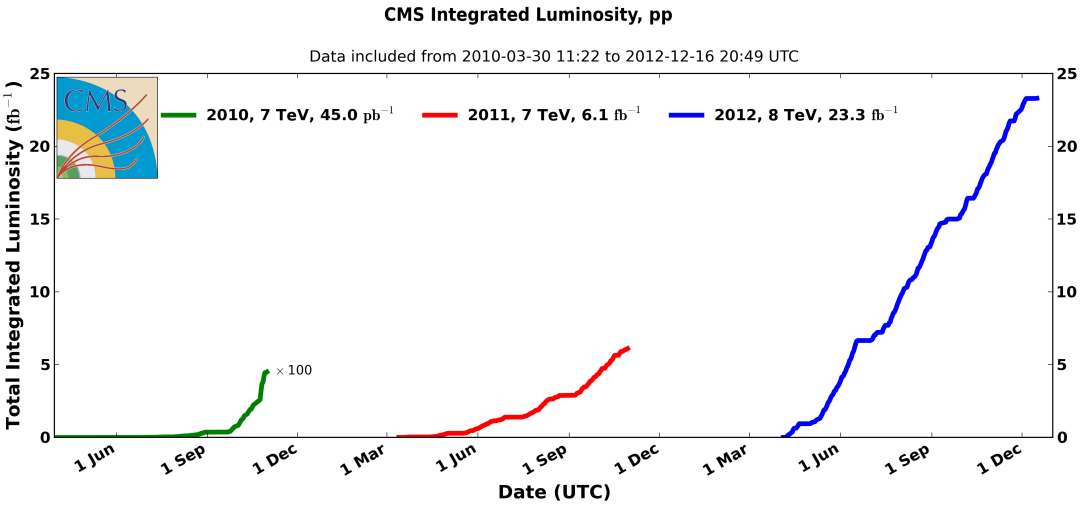

Figure 2.3: Overview of the integrated luminosity delivered to the CMS detector during the 2010-2012 data-taking period [58].

2.2 The Compact Muon Solenoid detector

One of the two main-purpose particle detectors of the Large Hadron Collider is the Compact Muon Solenoid (CMS) experiment designed to perform a wide variety of physics measurements. Therefore the specifications of the CMS experiment follow from the LHC physics program goals requiring good identification and momentum resolution throughout the entire detector. In order to efficiently measure all the different particles emerging from the interaction point, the CMS apparatus [59] consists of four separate subdetectors which are all optimised to identify specific types of particles: a tracking detector, an electromagnetic and hadronic calorimeter, and a muon system. The first three layers of the CMS detector are confined within the high-field superconducting solenoid magnet of 3.8 Tesla, as depicted in Figure 2.4. From a geometric perspective each of the subdetectors consists of a cylindrical barrel part centred around the interaction point for which both ends are closed hermetically by an endcap structure.

The layout of the CMS detector is mainly driven by the superconducting solenoid since the subdetectors positioned inside the magnet bore need to be as compact as

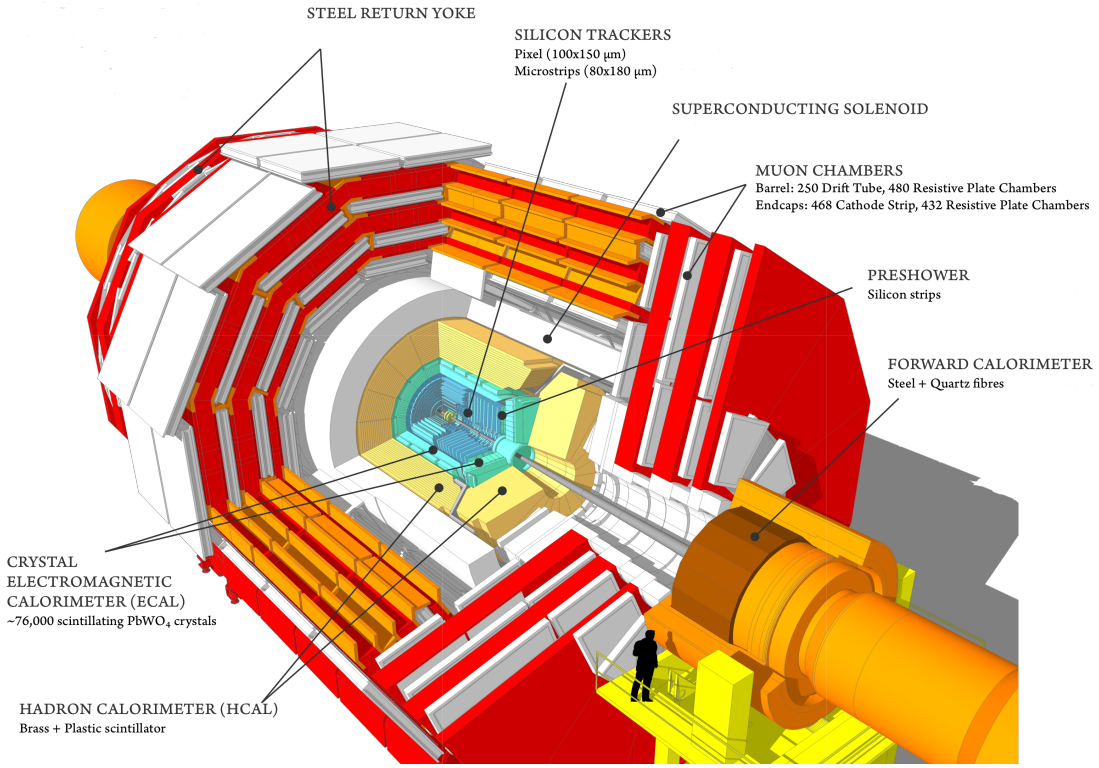

Figure 2.4: CMS layout with all subdetectors clearly visible [60].

possible without any loss of granularity. In addition the size of the muon system is restricted to four stations because of the magnet's return field.
The tracking detector is placed closely around the beam pipe and consists of a separate silicon-based pixel part and a strip detector as will be explained in Section 2.2.1, in order to guarantee track reconstruction in the high density environment close to the interaction point. Around the tracking detector the electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL) are positioned, which are responsible for accurately measuring the energy of the particles emerging from the interaction region. More detail about the characteristics of the calorimeter systems will be given in Section 2.2.2. Finally, surrounding the magnet coil, the muon chambers interleaved with the steel return yokes can be found. This part of the CMS detector, discussed in detail in Section 2.2.3, provides an accurate muon identification, crucial for distinguishing promising muon final state signatures from the extensive background. The overall dimensions of the full CMS detector are a total length of 21.6 m and a diameter of 14.6 m resulting in a total weight of 12 500 tons.

The CMS experiment has adopted a proper coordinate system for which the origin is centred at the nominal collision point within the detector. The y-axis is pointing upwards and the x-axis radially inwards toward the centre of the LHC. Hence, according to the right-hand rule, the z-axis follows along the anticlockwise-beam direction. This coordinate system can easily be converted into a spherical coordinate system where the azimuth angle ϕ is measured in the x-y plane and the polar angle θ from the z-

axis. From this coordinate system the pseudo-rapidity η can be derived, a variable used extensively in accelerator physics since it has the advantage to be invariant with respect to Lorentz boosts along the beam axis. Therefore this variable, defined as,

$$\eta = -\ln \tan \frac{\theta}{2} \tag{2.1}$$

is used to describe the angle of a particle with respect to the z-axis. The pseudo-rapidity is closely related to the rapidity, denoted with the symbol y and defined in Equation (2.2). Since this variable requires both the energy and the total momentum of a particle to be known, the rapidity is more challenging to determine. However in the case of high-energy collisions, both quantities are almost identical.

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \tag{2.2}$$

The Large Hadron Collider is able to provide a bunch-crossing rate of about 40 MHz, however, the current state-of-the-art computer systems are not capable of handling such a large rate of data. Hence the CMS experiment has been equipped with a dedicated multi-layered online event-selection system and uses a specialised computing system to store, transfer and manipulate the recorded data as will be explained in Section 2.2.4. The capability of the CMS detector to efficiently perform this complex data handling is discussed in Section 2.2.5.

2.2.1 The silicon tracking apparatus

The CMS tracking detector is located at the most inner point of the magnet bore close to the interaction point and is hence exposed to the harsh radiation environment produced by the proton-proton collisions. In order to survive these challenging conditions and still be capable of providing fast and accurate read-out of the particle's hits, it was decided to fully equip this tracking detector with active silicon, making it the largest silicon tracker ever constructed.
For the CMS tracking apparatus two different detection techniques have been adopted. The most inner part of the detector consists of pixel cells of size $100 \times 150\,\mu\text{m}^2$, capable of achieving similar track resolution in both $r - \phi$ and z direction, while the outer part contains silicon micro-strip sensors with diverse track resolution. This choice is motivated by the varying particle flux conditions throughout the tracking detector, which start out rather extreme at low radii but decrease when moving further away from the interaction region. Therefore the technology used in the first layers of the CMS tracker should be able to identify individual particle hits in a very dense track regime, which is feasible for silicon pixels. However the use of silicon strips is sufficient in the outer regions of the tracking detector where the track density is significantly lower. From Figure 2.5, which shows the geometry of the CMS tracker, the two structures are clearly visible together with the general subdetector layout containing central barrel layers closed with endcap disks allowing a larger η-coverage.

The pixel part of the tracker subdetector consists of three cylindrical barrel layers (BPix) placed at radii 44, 73 and 102 mm and two endcap disks (FPix) positioned at

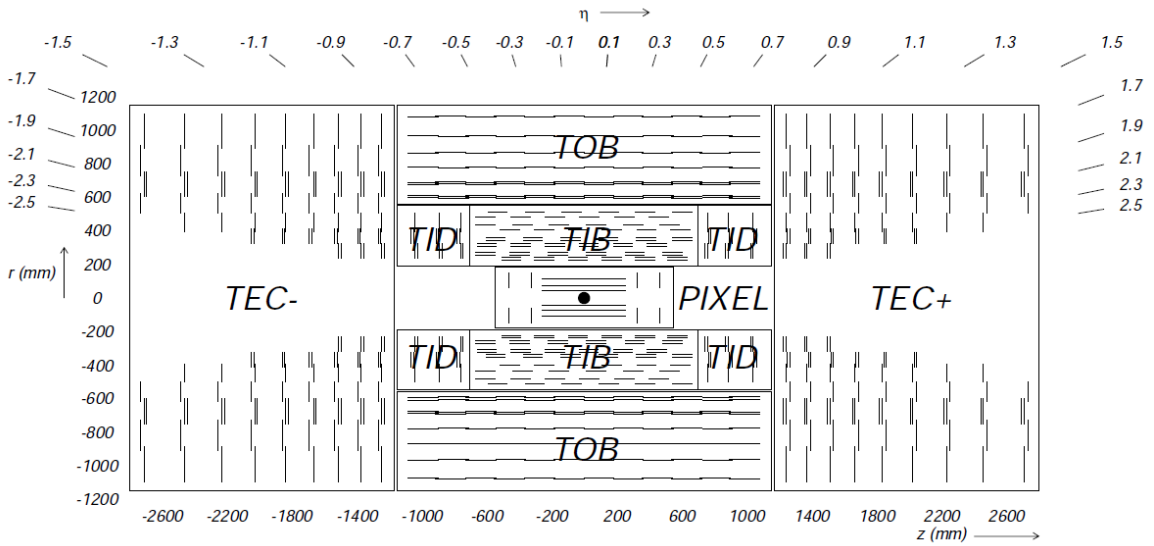

Figure 2.5: Schematic cross-section of the CMS tracking detector [50].

z = ± 345 and ± 465 mm on each side. The barrel layers itself are 570 mm long and thus extend up to $z = \pm 285$ mm resulting in, combined with the chosen positioning of the two endcap disks, at least three tracking points for almost the full pseudo-rapidity range. The silicon strip subdetector is further divided into an inner and outer part, both of which follow the barrel-endcap structure. The Tracker Inner Barrel (TIB) contains 4 concentric barrel layers located at radii 255.0, 339.0, 418.5 and 498.0 mm which extend from -700 mm to +700 mm along the beampipe. The Tracker Inner Disks (TID) are the endcap configuration for the inner strip detector and are each made up of 3 disks placed between ± 800 and ± 900 mm in z. Finally the Tracker Outer Barrel (TOB) and Tracker End Cap (TEC) further extend the overall dimensions of the CMS tracker detector to a diameter of 2.4 m and a length of 5.4 m by adding 6 detection layers and 9 disks.

Reconstruction of charged particle tracks

The track reconstruction algorithm is a computational complex and iterative process. It is designed to start from the innermost hit of the pixel detector and proceed outwards layer by layer. Hence this algorithm is active in the most dense environment of the tracker and therefore requires an efficient search for hits and a fast trajectory propagation. The track reconstruction within CMS is defined as the Combinatorial Track Finder [61] (CTF) and can be decomposed into four different steps: seed generation, pattern recognition, ambiguity resolution and track fitting.

Seed generation

This step of the track reconstruction provides initial trajectory candidates from pairs of pixel hits. The starting trajectory parameters and associated uncertainty can be determined from only five parameters since the magnetic field remains quasi-uniform in a large part of the tracker volume. Hence either

three hits or two hits combined with a beam spot constraint are sufficient to determine the seed trajectory of potential tracks.

Pattern recognition or track finding
This part of the CTF algorithm determines which hits are compatible with the seed trajectory using a combinatorial Kalman Filter [62] method. It starts by scanning for layers likely to be intersected by the track candidate. Then the trajectory parameters are extrapolated to this layer and the compatible hits are identified. Finally the track parameters are updated by adding one of the compatible hits and the procedure is repeated until the outermost layer is reached.

Ambiguity resolution
Since for each of the compatible hits the trajectory candidates are grown in parallel, a single seed can result in multiple tracks or two identical tracks can originate from a different seed. This possible double-counting is avoided by excluding specific tracks based on the number of hits shared among them.

Track fitting
The remaining mutually exclusive trajectory candidates are recalculated during the last step of the iterative tracking algorithm. It uses a Kalman Filter method on the full list of hits starting from the innermost one. Afterwards a smoothing stage is applied in the form of an outside-in Kalman Filter based on the result of the first one. This approach yields optimal estimates of the parameters.

Executing this CTF sequence multiple times and removing the hits associated with reconstructed tracks after each iteration significantly decreases the combinatorics. By first identifying the more straightforward track candidates the streamlined collection of unmatched hits allows the recovery of tracks that slightly deviate from the simplified pattern. Therefore the track reconstruction algorithm is implemented such that first the prompt tracks are identified and only afterwards the ones originating from outside the luminous region of the proton-proton collisions.

The performance of the track reconstruction is outstanding, and muons are reconstructed better than any other charged particle [61]. For isolated muons with $1 < p_{\mathrm{T}} < 100\,\mathrm{GeV}$ the tracking efficiency is higher than 99% for the entire η-range of the tracker and does not depend on p_{T}. The transverse momentum resolution for a muon with $p_{\mathrm{T}} = 100\,\mathrm{GeV}$ and $|\eta| < 1.6$ is of the order of 2-3% but quickly deteriorates for higher pseudo-rapidity values. The efficiency for reconstructing the trajectories of charged hadrons in $t\bar{t}$ events varies between 85 and 95% depending on the p_{T} and η value.

Primary vertex reconstruction

Once the full collection of reconstructed tracks is recovered the location and corresponding uncertainty of the associated proton-proton interaction vertices should be determined. Since only prompt tracks originating from near the interaction region are relevant for the primary vertex, these type of tracks need to be selected first. Afterwards the tracks that appear to originate from the same interaction vertex are clustered

based on their z-coordinates. Finally the actual vertex fitting can be applied on the candidates which contain at least two tracks using an adaptive vertex filter [63]. This fitting procedure computes the vertex parameters and assigns a weight to each track in the vertex, reflecting the probability that it actually belongs to the considered vertex.

The complete primary vertex reconstruction results in an accurate measurement of the position of the primary vertices. For events with a reconstructed jet with transverse energy $E_{\mathrm{T}} > 20\,\mathrm{GeV}$ the obtained resolutions are about $10\,\mu$m in x and y, and $12\,\mu$m in z [61].

2.2.2 The calorimetry subdetectors

Any charged particle emerging from the interaction point will be detected by the tracker detector and its trajectory will be reconstructed. However in order to fully identify the observed particle this information should be combined with the measured energy deposits, a task performed by the electromagnetic and hadronic calorimeters. The ECAL is designed to measure the parton showers produced by electrons and photons while the HCAL will absorb the hadron showers.

The electromagnetic calorimeter

The ECAL is a hermetic homogeneous detector made of lead tungstate ($PbWO_4$) crystals, which ensures the detector to be fast, fine in granularity and radiation resistant. The ECAL subdetector contains in total 75 848 of lead tungstate scintillating crystals, of which 61 200 are placed in the barrel part (EB) while the remaining are equally distributed among the two endcap structures (EC). The pseudo-rapidity coverage of the EB is $|\eta| < 1.479$, which is further extended by the two endcaps to $|\eta| < 3.0$, as can be seen in Figure 2.6. The crystals used in the two subdetector parts vary slightly, for the barrel crystals with a front-face cross-section of $2.2 \times 2.2\,\mathrm{cm}^2$ with a total length of 23 cm were constructed while for the endcap these values take $2.86 \times 2.86\,\mathrm{cm}^2$ and 22 cm, respectively. The ECAL contains besides the barrel and endcap an additional third substructure: the Preshower (ES). This is a sampling calorimeter consisting of lead radiators combined with silicon strip sensors specifically designed to identify and reject signals originating from neutral pions and minimum ionising particles. The ES is only 20 cm thick and is restricted to the $|\eta|$-coverage between 1.653 and 2.6.

The energy for photons and electrons can be measured very precise with the electromagnetic calorimeter. Their energy resolution has been measurered with electrons originating from Z-boson decays, for which resolutions of 2% are obtained in the central region and 2-5% elsewhere [64].

The hadronic calorimeter

The HCAL subdetector surrounds the electromagnetic one and is the last part of the CMS experiment confined within the solenoid. It is a brass/scintillator calorimeter, motivated by the fact that brass is an efficient hadron absorber on a short scale and

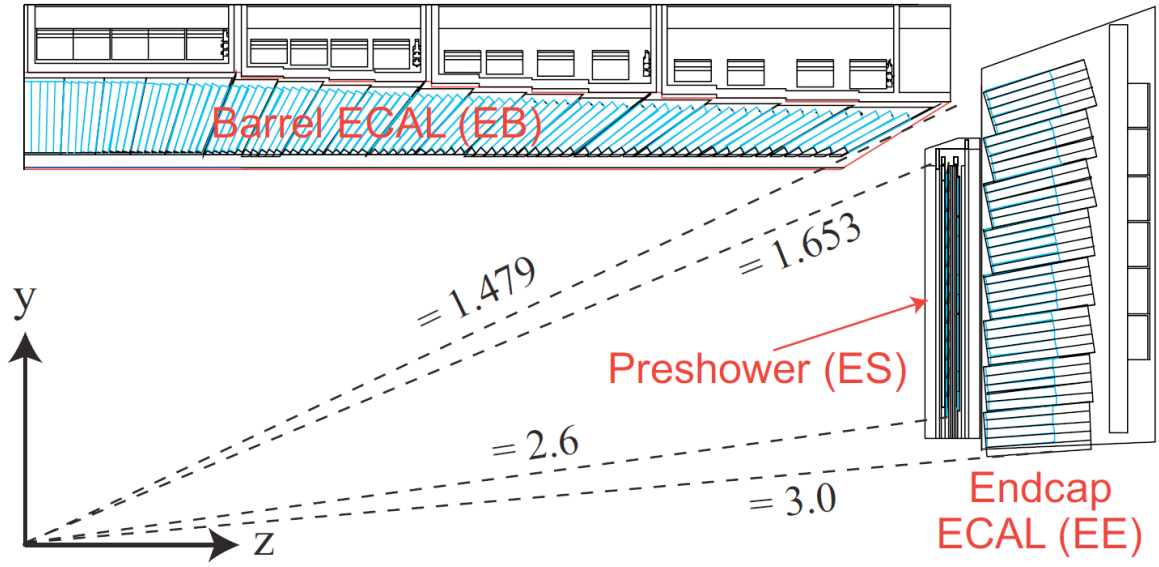

Figure 2.6: Overview of one quarter of the electromagnetic calorimeter with the different substructures and their respective pseudo-rapidity coverage [59].

that the use of scintillating tiles is interesting when only limited space remains for the active medium. The geometrical design of the hadronic calorimeter is depicted in Figure 2.7 and also consists of a barrel and endcap part.

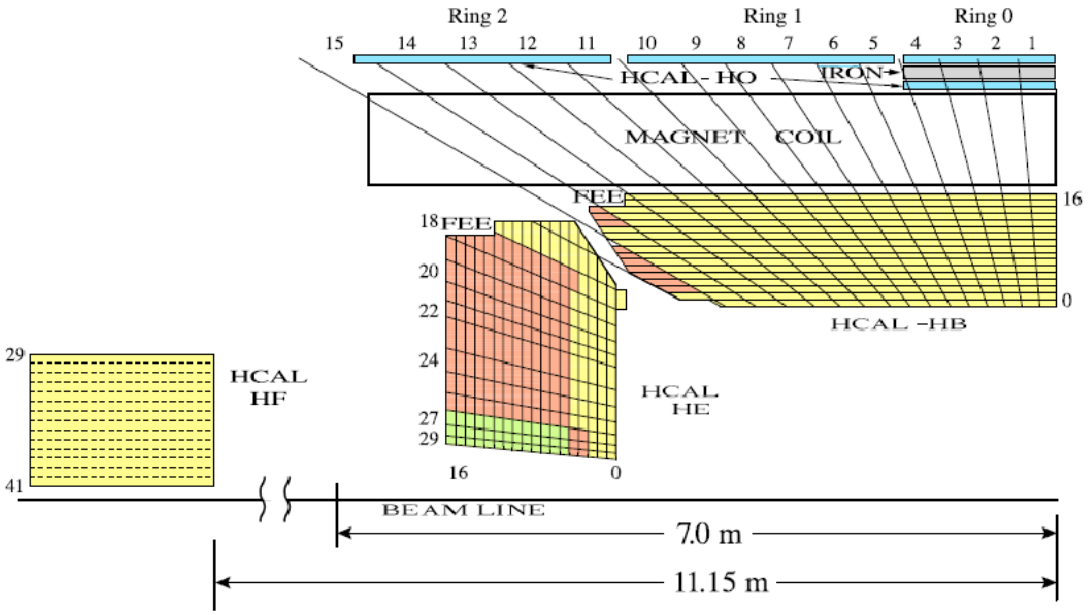

Figure 2.7: Layout of the hadronic calorimeter [65].

The barrel (HB) consists of 17 scintillator plates interleaved with brass plates. The individual scintillator tiles have a size of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$, cover an overall pseudo-rapidity range of $|\eta| < 1.4$ and are contained within radii 1777 mm and

2876.5 mm. The endcaps (HE) are made up of 19 scintillator plates with the same thickness and brass plates of 7.8 cm. The pseudo-rapidity coverage is extended to $|\eta| < 3.0$ and even an overlap in coverage is provided between $1.3 < |\eta| < 1.4$. For the HE scintillator tiles of similar size as the HB have been used up to $|\eta| < 1.74$, afterwards their size is increased to maximally $\Delta\eta \times \Delta\phi = 0.350 \times 0.174$.
Only the HB and HE are not sufficient for measuring all hadronic activity produced within a proton-proton interaction. For the barrel part the restricted size available due to the confinement in the solenoid results in possible leakages of hadron showers in the muon chambers. As a result, the central part of the hadronic calorimeter has been extended by placing an outer tail catcher (HO) just outside the magnet coil, which has been limited to the $|\eta|$-range of 1.26. Also the endcap calorimeters require a complementary structure capable of measuring energy deposits in the very forward region. These forward calorimeters (HF) are positioned at 11.2 m from the interaction point, cover the range $3.0 < |\eta| < 5.0$ and are thus exposed to enormous particle fluxes. Since the HF needs to survive at least a decade in these harsh conditions, the choice has been made to construct this calorimeter as a steel/quartz fibre calorimeter.

The performance of the hadronic calorimeter has been determined using charged pions. The obtained energy resolution is about 24% for 20 GeV pions and improves to approximately 13% for a pion of 100 GeV.

2.2.3 The muon system

The last subdetector of the CMS experiment, the only one which is positioned completely outside the magnet coil, is the muon system. The sole goal of the muon chambers is to accurately identify and measure the muons created during proton-proton interactions. Enlarging the CMS detector with the muon system almost doubles the size of the overall experiment making it the largest subdetector of the four. The large surface which needs to be covered by this single substructure combined with the varying radiation conditions resulted in the use of three different detector technologies, which are all listed in Figure 2.8.
The central region of the muon system consists of four concentric layers which cover $|\eta| < 1.2$ and are equipped with drift tubes (DT) while the endcap part ($|\eta| < 2.4$) has four stations containing cathode strip chambers (CSC). The DTs are designed for the low rate which is expected in the barrel region and thus have a slower response time than the CSCs in the endcap. The spatial resolutions obtained for the two detector systems are similar, 80-120 μm for the DTs and 40-150 μm for the CSCs [66]. For the pseudo-rapidity region $|\eta| < 1.6$ additional resistive plate chambers (RPC) are added in order to ensure a fast response with good time resolution. Since the RPCs are specifically designed for an accurate time measurement the corresponding spatial resolution is of less importance and is found to be 0.8-1.2 cm [66].

2.2.4 Online event filtering

The Large Hadron Collider is designed to provide about 40 million bunch crossings per second during which multiple simultaneous proton-proton collisions can occur. Storing

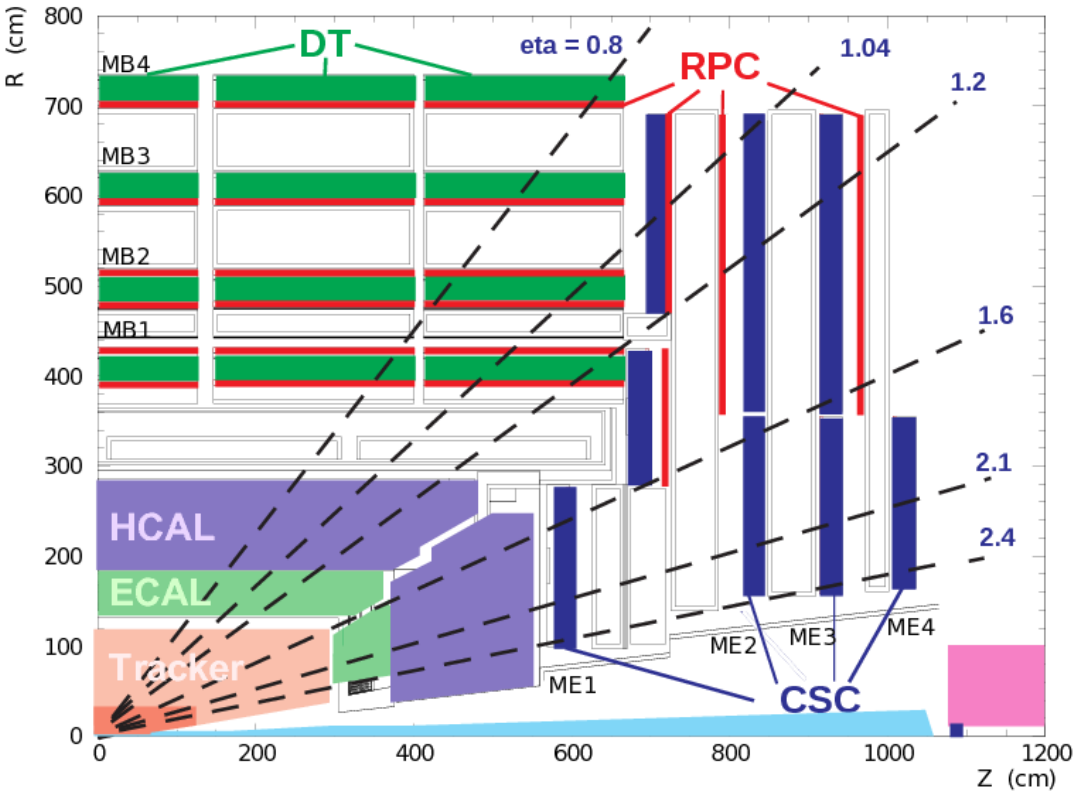

Figure 2.8: Cross section of one quarter of the CMS muon system [59].

such a large amount of data is not feasible, hence the event rate should be drastically reduced by filtering out the seemingly interesting interactions.
This event filtering process is performed by a dedicated multi-layered trigger system responsible for reducing the bunch crossing rate of about 40 MHz to 400 Hz, which is the maximal rate that can be stored. For the CMS trigger system this rate reduction is obtained using two complementary triggers, where the first one (Level-1 or L1) is designed to execute fast decisions while the second one (High-Level trigger or HLT) is capable of performing complex calculations in case of interesting events.
The L1 trigger is mounted partially on the detector itself and has only access to information from the calorimeter and muon subsystems. Since this trigger layer should decide within 3.2 μs whether an event looks promising enough to be analysed further, the processing of each bunch crossing is pipelined to avoid any dead-time regions. The accepted events are studied in detail by the HLT, which has access to the complete read-out data and has less stringent time constraints. In contrast to the L1 trigger which merely functions as a keep/reject switch, the HLT also serves as a labelling system that tags the selected events with the specific trigger requirements that were satisfied. This information can then later be used during the offline selection to split the events into dedicated datasets.
Even though the online triggering systems provides a radical rate reduction, CMS still requires a dedicated offline computing system to store, transfer and manipulate the unprecedented amounts of recorded data. Since the CMS collaboration prefers the

recorded data to remain accessible throughout the entire lifetime of the experiment, a complex distributed system of large scale with a layered structure is needed.
This computing system is set up as a collaboration between LHC experiments, computing centres and middleware providers and is referred by as the WorldWide LHC Computing Grid (WLCG). This allows to host the majority of the computing resources outside the CERN area since only the Tier-0 centre is located at CERN. The next step in this hierarchical tiered structure is provided by some of the large national computing facilities, which host the Tier-1 centres. Finally various Tier-2 centres can be found at partner universities, and one of these is located at the Interuniversity Institute for High Energies (IIHE) in the VUB/ULB computing centre.
The Tier-0 is mainly devoted to recording the detector information from the experimental site (RAW) and reconstruct the first datasets (RECO). It is the only one of the computing centres which is not accessible for analysis use. The following layer of Tier-1's provides storage of a second complete copy of the RAW data and allows more complex reconstruction algorithms for RECO samples. Finally the Tier-2 centres are designed to be used for final-stage analysis and for specialised activities which can be performed offline.

2.2.5 CMS performance during Run-I data-taking

The number of proton-proton interactions delivered by the LHC can differ from the number of interactions actually recorded by the experiments. Such a data loss can for example be caused by a technical malfunction in one of the detector's subsystems or by an unregulated overload of the trigger rate. However the overall performance of the CMS detector during the Run-I data-taking period, ranging from 2010 to 2012, was really good and any significant data-losses were avoided. The CMS detector reached an efficiency of 92.24%, 90.98% and 93.52% for 2010, 2011 and 2012, respectively. For the 2012 run at $\sqrt{s} = 8\,\text{TeV}$ the comparison between the integrated luminosity delivered by the LHC and recorded by the CMS experiment can be found in Figure 2.9.
Afterwards the recorded data is validated by the offline Data Quality Monitoring (DQM) to ensure that it is suited for physics analysis. This largely automated tool determines whether all subdetectors were working properly during data-taking and monitors the reconstruction of the various physics objects. The analysis discussed in this thesis will be performed using a total of $19.6\,\text{fb}^{-1}$ of integrated luminosity recorded by the CMS experiment.

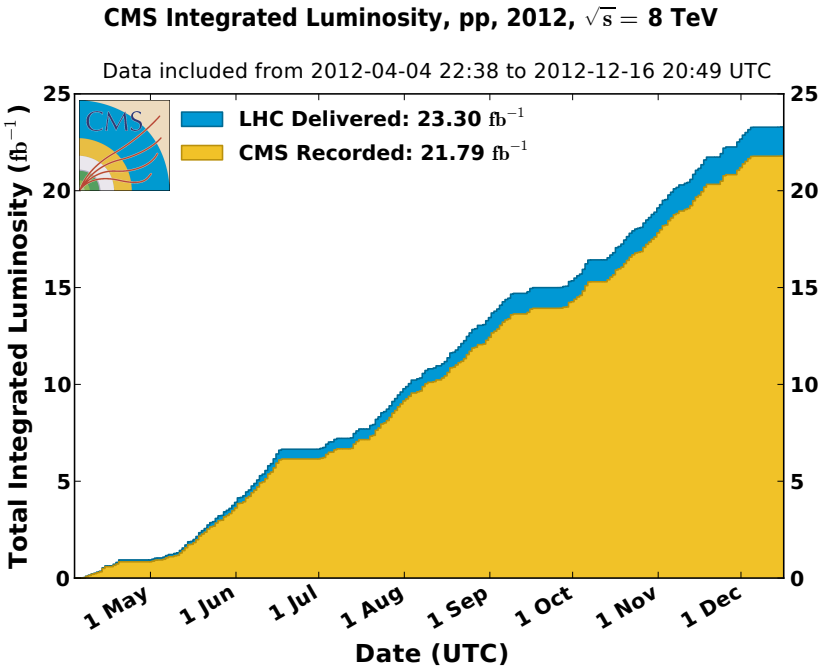

Figure 2.9: Integrated luminosity delivered by the LHC and recorded by the CMS experiment during the 2012 data-taking at $\sqrt{s} = 8\,\text{TeV}$ [58].

Chapter 3

Event simulation and reconstruction

An accurate understanding of simulated collision events and their reconstruction in particle detectors is crucial for detailed studies of the collected data. The use of event generators at hadron colliders allows to investigate the performance of data analysis strategies and techniques, for example the expected power to discriminate signal from background processes.
This is accomplished by a detailed simulation of the different processes taking place during proton-proton collisions, explained in Section 3.1. An accurate description of the phenomena is to be combined with a realistic representation of the detector response as briefly discussed in Section 3.2. The remaining part of this chapter, Section 3.3, contains the reconstruction of physical objects from the true or simulated electronic readout of the CMS detector.

3.1 QCD at hadron colliders

The composite nature of protons together with the high-momentum transfers reachable at the LHC significantly complicates the event structure. The different processes taking place during a single proton-proton collision can be factorised [67–71]. An overview of the factorised subprocesses is given in Figure 3.1 and briefly discussed below.

Parton Distribution Functions
In proton-proton collisions both incoming protons can be viewed as a collection of partons whose momentum fraction x within the hadron is parametrised by the so-called parton distribution functions.

Hard scattering
Hard scattering is the perturbative process of two colliding partons, one originating from each proton, that creates the high-energetic final-state particles. It can be represented by a factorised product of the initial- and final-state contributions as described in Section 3.1.1.

Parton shower
This phase of the event generation process describes approximately the higher-order corrections induced by emission of additional gluon and/or quarks, as will be explained in Section 3.1.2. Depending whether this radiation originates

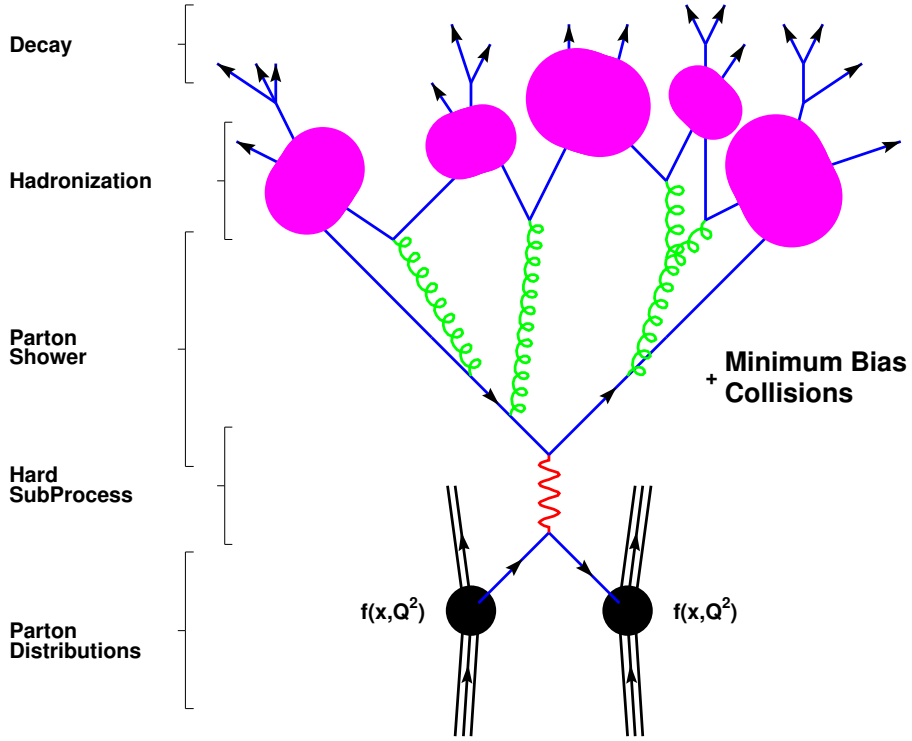

Figure 3.1: Schematical overview of the consecutive steps of the event generation process [72].

from the incoming or outgoing partons it is denominated, respectively, Initial State Radiation (ISR) or Final State Radiation (FSR).

Hadronisation

The collection of receding post-shower partons is combined into experimentally observable colour-neutral hadrons as required by colour confinement. This hadronisation process is described by QCD-inspired phenomenological models as discussed in Section 3.1.3.

The main challenge at hadron colliders is to reconstruct the missing information about the partons of the hard interaction from the observed quantities in an event. The event generation process at the LHC is even more arduous due to the diversity of QCD phenomena in the accessible range of momentum transfer Q^2. The interaction produces during the hard interaction a few high-energetic outgoing leptons, gauge bosons or partons of which the latter afterwards transform non-perturbatively into final-state hadrons. This large variation in energy range, and corresponding QCD coupling strength, implies that only the event generation's high-momentum transfer contribution can be derived perturbatively from the QCD Lagrangian while the other aspects have to be expressed using phenomenological non-perturbative models.

3.1.1 Hard Scattering

Most events studied at the LHC involve high-momentum transfers in order to create massive particles. The inclusive production cross section of an observable final state X from hadrons h_1 and h_2 can be factorised into:

$$\sigma_{h_1 h_2 \to X} = \sum_{a,b \in \{q,g\}} \int dx_a \int dx_b f_a^{h_1}(x_a, \mu_F^2) f_b^{h_2}(x_b, \mu_F^2) \int d\Phi_{ab\to X} \frac{d\hat{\sigma}(\Phi_{ab\to X}, \mu_F^2)}{d\Phi_{ab\to X}} \quad (3.1)$$

From Equation (3.1) can be concluded that the hadronic cross section, valid for all orders in perturbation theory, is actually a convolution of a perturbative short-distance component $\hat{\sigma}_{ab\to X}$, calculable from Matrix Elements, and an approximate long-distance one, represented by the parton distribution functions (PDF). The PDF $f_a^h(x_a, \mu_F)$ is the probability of encountering parton a with momentum fraction x_a in parent hadron h when this is probed at energy scale μ_F. This factorization scale μ_F symbolises the transition from the short-distance process to the long-distance one. The partonic scattering cross section $\hat{\sigma}(\Phi_{ab\to X}, \mu_F^2)$ depends on the final state phase space $\Phi_{ab\to X}$.

Equation (3.1) serves as the starting point for event simulation in general-purpose Monte Carlo event generators which, due to the perturbative nature of the parton-level differential cross section, can be expanded in orders of the QCD coupling α_S. Originally these calculations were performed at leading order (LO), corresponding to $\mathcal{O}(\alpha_S^2)$, however this only describes the simplest processes taking place in hadron colliders and does not correspond to reality where additional radiation on top of X occurs. Moreover the current theoretical precision needed to study QCD at colliders requires at least next-to-leading order (NLO) calculations. Hence much effort has been devoted in order to overcome the infrared singularities in QCD allowing to extend the matrix element generators to perform these NLO calculations in an automated way and thereby significantly improving accuracy and predictive power.
In order to generate an event sample which allows for a theoretical description of real collision events several event generators can be used. For example `HERWIG` [73] and `PYTHIA` [74], which are two general-purpose event generators but are both limited to leading-order calculations. Hence a more precise description of the data is obtained for the multi-leg LO `MadGraph` [75] generator and the NLO generators `POWHEG` [76] and `MC@NLO` [77], which are therefore used more frequently.

MadGraph

MadGraph is a tree-level matrix-element generator for decays and $2 \to n$ scatterings. Real higher-order corrections, through the inclusion of additional final-state partons, can be incorporated, but no virtual higher-order corrections are considered. Currently the matrix-elements for $t\bar{t}$ production with up to 3 additional partons and W-boson production with up to 5 additional partons can be calculated. The actual event-generator is then provided by the `MadEvent` event generator using these matrix-elements.

POWHEG and MC@NLO

`POWHEG` and `MC@NLO` are two event generators which are capable of calculating

NLO corrections for a limited number of processes and, even more important, correctly matching them with the additional particles created during the parton shower step, which will be discussed in detail in 3.1.2.

3.1.2 Parton shower

The hard interaction is followed by the iterative process of higher-order emissions defined by the Parton Shower (PS) formalism. The partons formed during the hard scattering are prone to gluon radiation emission, $q \to qg$, and gluon branching, $g \to gg$. The first type of QCD parton branching corresponds in QED to Brehmstrahlung while the second QCD process has no analogy in QED and is caused by QCD's non-abelian nature. Both processes are incorporated in the PS formalism which sequentially lowers the transverse momentum of the contributing partons until the QCD confinement limit is reached, resulting in a parton cascade.

The parton shower formalism's objective is to convert the inclusive cross section for the production of parton a into the exclusive cross section taking into account a number of additional less-energetic particles [78]. Hence the complex $2 \to n$ process will be decomposed into a hard interaction with momentum transfer Q^2 and a succession of gluon radiations each with momentum transfer Q_i^2; a justifiable approach in the approximation $Q_i^2 \ll Q^2$ which is defined as the collinear limit. The Alterelli-Parisi splitting functions [79], denoted $P_{ba}(z)$, describe this collinear splitting of parton b into parton a and are defined as:

$$P_{qq}(z) = \frac{4}{3}\frac{1+z^2}{1-z} \qquad P_{qg}(z) = \frac{4}{3}\frac{1+(1-z)^2}{z} \tag{3.2}$$

$$P_{gq}(z) = \frac{n_f}{2}(z^2+(1-z)^2) \quad P_{gg}(z) = 3\frac{z^4+1+(1-z)^4}{z(1-z)} \tag{3.3}$$

where n_f represents the number of quark flavours.
These splitting functions are divergent in the case of $z \to 1$, corresponding to soft gluon emission as z is the momentum fraction carried away by the parton a. Since reality is known to be finite, these soft divergences, together with the collinear divergencies, have to be excluded by introducing a cutoff scale on the transverse momentum k_t below which all remaining perturbative effects are absorbed by the parton distribution functions. The freedom of choosing this factorisation scale μ_F^2, generally around 1 GeV^2, necessitates the introduction of the DGLAP (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi) evolution equations [80, 81], which represent the fact that any parton a may have been produced by the branching of parton b at slightly higher scale $\mu_F^2 + d\mu_F^2$:

$$\mu_F^2 \frac{df_a^h(x,\mu_F^2)}{d\mu_F^2} = \sum_{b\in\{q,g\}} \int_x^{z_{max}} \frac{dz}{z}\frac{\alpha_S}{2\pi} P_{ba}(z) f_b^h(x/z,\mu_F^2) \tag{3.4}$$

Even though the introduction of the factorisation scale resolved the divergencies, the branching probability in Equation (3.4) can still exceed unity. This because also the virtual divergencies which lead to cancellations have been removed by this cutoff. Hence

total conservation of probability should be restored by adding an additional term to the DGLAP equations:

$$\mu_F^2 \frac{df_a^h(x,\mu_F^2)}{d\mu_F^2} = \sum_{b\in\{q,g\}} \int_x^{z_{max}} \frac{dz}{z} \frac{\alpha_S}{2\pi} P_{ba}(z) f_b^h(x/z,\mu_F^2) - f_a(x,\mu_F^2) \sum_{b\in\{q,g\}} \int_{z_{min}}^{z_{max}} dz \frac{\alpha_S}{2\pi} \frac{1}{2} P_{ab}(z) \quad (3.5)$$

The above equation can be further simplified by identifying the Sudakov form factor, which represents the probability for a parton not to undergo a branching between the energy scales $t^{'}$ and t.

$$\Delta_a(t,t^{'}) = \exp\left\{ - \sum_{b\in\{q,g\}} \int_t^{t^{'}} \frac{dk_T^2}{k_T^2} \int_{z_{min}}^{z_{max}} dz \frac{\alpha_S}{2\pi} \frac{1}{2} P_{ab}(z) \right\} \quad (3.6)$$

The parton shower algorithm outlined above is applicable for both initial and final state radiation since their branching probabilities are similar. However the actual implementation in the Monte Carlo event generators is performed in an entirely different manner. Initial state radiation is simulated by employing a backward evolution: the Monte Carlo event generator starts from the desired hard interaction and surrounds the initial partons with additional radiation only afterwards. This because each parton branching significantly reduces the energy of the initial partons and therefore the possibility to produce the hard process of interest, such as top-quark pair production. Final state radiation on the other hand is taken into account in a much more straightforward way: the parton branching starts at the hard interaction scale Q^2 and is sequentially lowered until the factorisation scale μ_F^2 is reached.

Combine hard scattering with parton showering

From the overview given above can be understood that the Matrix-Element and Parton-Shower algorithm have some crucial differences in how to simulate $X+n$-jet topologies. The former one reliably describes the simulation of well separated hard partons but lacks information about the collinear and soft partons while for the latter one this is exactly the contrary. Hence for accurately describing the entire event simulation chain up to the final-state hadrons both approaches should be combined. However this is not a straightforward process since possible double-counting can occur. A hard parton of a $X+2$-jet event can originate either from a $X+2$-jet fixed-order matrix-element calculation or otherwise from a hard emission during the showering of the $X+1$-jet event.

Different approaches exist for correctly dealing with this double-counting issue and the ones used in this thesis are outlined below. A distinction should be made whether the matrix-element calculations have been performed at LO or NLO accuracy since the latter significantly complicates the combination procedure.
In case LO matrix-elements have to be combined with the parton shower the MLM

approach [82] is applied which imposes a cut on the jet transverse momentum to ensure any hard jet in the event to originate from the hard interaction. So an event is rejected in case more than the requested number of jets have transverse momentum above the merging scale. The hard jets produced in this way are certain to be described by tree-level matrix elements since the merging scale is chosen to be larger or equal than the matrix-element cutoff scale. As in the parton shower algorithm, this approach can be depicted by introducing Sudakov factors that represent the probability for not undergoing a hard scattering below the merging scale during the showering process. The MLM approach is used in `HERWIG` and `PYTHIA` and results in a parton shower structure with LO accuracy which is applied in a broad range of LHC analyses.
Even though the LO results using the MLM approach are successful in describing shapes of experimental distributions considerable gain can be reached by extending to NLO accuracy. The main challenge of matching NLO calculations with parton showers is to overcome the additional double-counting introduced by the approximate NLO corrections included in the parton shower generators. One of the first acceptable NLO matching methods which correctly tackles this double-counting issue was the so-called $MC@NLO$ algorithm, which only applies the parton-shower algorithm on PS-corrected NLO matrix-elements. The correction term is obtained by first computing the NLO matrix-element corrections to n-body decay, then calculating how the first shower of a n-body decay would populate the $n+1$-body phase space and finally subtracting this approximate shower calculation from the exact NLO matrix-element. The downside of the $MC@NLO$ approach is two-fold: the subtraction of the two contributions can lead to negative weights and the subtraction terms are generator-dependent such that for now only `HERWIG` can be used for performing the parton shower.
Hence a new NLO matching method was developed to overcome both the presence of negative weights and the generator-dependency of the $MC@NLO$ approach. The `POWHEG` approach starts from the hardest emission using full NLO accuracy and applies normal showering afterwards. This implies that only one emission beyond LO should be generated in order to obtain NLO accuracy. In this thesis the `POWHEG` approach is combined with the `PYTHIA` parton shower algorithm.

3.1.3 Hadronisation

The missing link in the event generation process is how the quarks and gluons produced during both the hard interaction and the showering turn into experimentally observable colour-neutral hadrons. This step is defined as the hadronisation or fragmentation process and is represented by phenomenological models since it cannot be calculated from first principles due to the corresponding low energy scales. Two distinct models for describing this non-perturbative process are used today: the Lund string model and the cluster model. The former one is implemented in `PYTHIA` while the latter one is used by `HERWIG`.
The Lund string model [83] is based on linear confinement, which states that the potential V between a quark-antiquark increases with separation distance r due to the presence of a strong QCD colour field:

$$V = \kappa r \quad \kappa \sim 1 \frac{\text{GeV}}{fm} \tag{3.7}$$

Hence the kinetic energy of such a parton pair will transform into potential energy and accumulate while receding. Once sufficient energy is stored in the colour string stretched between the quark q and anti-quark $\bar{q}$, the string will split into a new $q\bar{q}$ pair with a colour string surrounding each parton pair. A fraction of the potential energy will be absorbed by the parton creation and, as a consequence, lowering the remaining energy during each following string splitting until no subsequent splittings can occur. The probability for the creation of a quark with mass m and transverse momentum p_{T} during such a splitting is given by:

$$\exp\left(-\frac{\pi m^2}{\kappa}\right)\exp\left(-\frac{\pi p_{\mathrm{T}}^2}{\kappa}\right) \tag{3.8}$$

However, the above formula only represents the formation of light u-, d- and s-quarks since the presence of the mass term implies that the production of heavier quarks is suppressed during this step of the event generation process.
The transition of these free quarks into bound states is described by the Lund fragmentation function which gives the probability of a colour string to produce a hadron h with mass m_h, transverse momentum p_T and longitudinal momentum fraction z during the string-breaking process. The fragmentation function exhibits a "left-right" symmetry since the splitting sequence should be identical whether is started from the quark or anti-quark.

$$f(z) \propto \frac{1}{z}(1-z)^a \exp\left(-\frac{b(m_h^2 + p_{\mathrm{T,h}}^2)}{z}\right) \tag{3.9}$$

with a and b free parameters of the model. In order to overcome the suppression of heavier hadrons an additional $1/z^{bm_Q^2}$ [67] factor has to be taken into account.
The second hadronisation model, the cluster fragmentation model [78], is based on the preconfinement property of QCD and splits the gluons non-perturbatively into $q\bar{q}$ pairs after the parton shower. From this clusters, or colour singlet combinations of partons, can be created which transform into hadrons either directly or through splitting processes depending on their mass.

3.1.4 Additional event activity

The previous sections have provided a detailed overview of the event generation process from start to finish, but were limited to the ideal situation where only one parton present in each proton contributes to the production of final-state hadrons. A more realistic representation would be to take into account the additional activity which occurs in coincidence with the primary parton collision.

The term Underlying Event (UE) has been adopted as collective noun to depict the types of additional interactions during a single hadron-hadron collision that could possibly alter the final-state. At the LHC where protons are used as incoming particles, two distinct soft phenomena contribute to the UE: the beam remnants and the multiple parton interactions (MPI) [84]. The beam remnant is defined as the remainder of the proton after the hard-interacting partons are extracted. The non-zero colour charge of the beam remnant implies the creation of additional hadrons during the hadronisation

process is possible. Multiple parton interactions represent the distinct scattering processes that could take place between other incoming partons. The presence of MPI can be understood from the composite nature of protons implying that each parton is as likely to undergo scattering interactions within one single hadron-hadron interaction. The jets produced from the MPI are in general less energetic than the principal hard interaction and the chances of producing an extra hard interaction during this process are very rare. As a result the underlying event is mainly saturated with low-energetic partons which tend to travel along the beamline such that it can be studied by selecting a specific topological structure.
The complexity of the underlying event lies in the fact that it involves both non-perturbative and perturbative QCD making it, at least for the moment, practically impossible to understand the physics. Hence Monte Carlo models have to be tuned; i.e. constraining free parameters using existing data; in order to accurately describe the collider physics.

Another important contribution to additional event activity which significantly complicates the final-state topology of hadron-hadron collisions is pileup (PU) or additional hadron-hadron collisions. This type of interactions occur because hadron-hadron collisions are not performed between two single hadrons but between bunches of hadrons. Therefore it is again likely for multiple hadrons to result in a hard interaction surrounded with UE during one bunch crossing. A distinction can be made whether the PU originates from the same bunch crossing as the hard interaction or from a previous bunch crossing, which are defined as in-time PU and out-of-time PU, respectively. The latter one can manifest itself when the bunches are spaced such that the next one arrives before the previous one evacuated completely from the detector.

3.2 Simulating detector response

The Monte Carlo event generators adopt an approach independent of the considered accelerator complex, besides starting with the correct incoming partons, and can thus be applied in a very broad physics range. In order to study specifically proton-proton collisions collected at the CMS detector, the simulated events are pushed through a dedicated detector simulation chain based on the `GEANT4` software toolkit [85]. This software package contains a full geometrical description of the CMS detector and a detailed mapping of the magnetic field, implemented in a flexible way allowing the activation and deactivation of specific detector subsystems.
This detailed detector simulation treats the simulated events as actual data originating from the interaction point and propagates them through the entire detector while taking into account energy loss caused by interactions with the detector material. The simulated energy deposits in the different subsystems are converted into electronic signals based on the actual detector behaviour resulting in a completely identical treatment as for real data. During this step pileup is included by overlaying the primary interaction with generated proton-proton events simulated in an identical manner. The mixing procedure applied depends on the specific subdetector considered, especially for incorporating out-of-time pileup since different numbers of preceding and succeeding

bunch crossings should be regarded. Employing the full simulation chain results in very good agreement with actual data and is therefore widely used in many physics analyses, including this thesis. However this full simulation, known as FullSim, is very time-consuming, in general several minutes are necessary for processing a single event, resulting in the development of a fast simulation chain, denominated as FastSim. The simplified geometry adopted in the FastSim approach reduces the CPU-time significantly. An overview of the Monte Carlo samples used in this thesis are listed in Table 3.1.

3.3 Physics object reconstruction

The detector readout results in a collection of electronic signals which still need to be translated into actual particle signatures, a process performed by the reconstruction algorithms. Within CMS emphasis was placed on developing a physics-object reconstruction algorithm capable of combining information from multiple subdetectors. Compared to older algorithms that only used specific parts of the detector for reconstructing particles, such a combined approach is capable of exploiting the different optimisations of each of the separate subdetectors. The algorithm requires a set of building blocks distributed across the entire detector which are afterwards linked in order to construct stable final-state particles.
Since this particle-flow event reconstruction uses the muon and electron candidates obtained from the standard reconstruction algorithms, these will be discussed first in Sections 3.3.1 and 3.3.2. Afterwards the specific aspects of this reconstruction algorithm will be explained in Section 3.3.3. The reconstruction of the jets, the identification of b-quark jets and the reconstruction of the missing transverse energy will be given in Sections 3.3.4, 3.3.5 and 3.3.6.

3.3.1 Muon reconstruction

The muon reconstruction algorithm is designed to fully exploit the excellent reconstruction efficiency in both the tracker and the muon system. Hence tracks reconstructed in the inner tracker and the muon system separately are combined into actual muon candidates. In order to distinguish these two types of muon-seeds they are called *tracker track* and *standalone-muon track*, respectively.
The identification of standalone-muon tracks is performed in two consecutive steps [59, 86]. First local reconstruction starts by constructing track segments from the detected hits in the DT and/or CSC chambers. Afterwards the track segments found in the innermost chambers are used as seeds for the standard reconstruction algorithm based on the Kalman Filter technique [62]. First an inside-out Kalman Filter is applied which propagates the muon track to the next layer, compares with the measured energy deposits and updates the track parameters accordingly. Once the outermost layer of the muon system is reached, a second Kalman Filter is used to calculate the track parameters at the innermost muon station. Finally, in order to improve the momentum resolution, an additional beamspot constraint is applied to the track parameters before

Table 3.1: Overview of the simulated samples at $\sqrt{s} = 8\,\text{TeV}$. Showering and hadronisation is performed with PYTHIA for all these samples.

Sample	Generator	σ (pb)	# events	$\mathcal{L}$ (fb^{-1})
$t\bar{t}$ + jets, $m_t = 172.5$ GeV	MadGraph			
ℓ + jets		107.7	24.4 M	226.5
dilepton		25.8	11.8 M	455.3
all-hadronic		112.3	30.5 M	271.9
$W \to \ell\nu_\ell$ + jets	MadGraph			
W + 4 jets		264.0	13.2 M	49.9
W + 3 jets		640.4	14.8 M	23.2
W + 2 jets		2159.2	32.9 M	15.2
W + 1 jet		6662.8	22.4 M	3.4
$Z/\gamma^* \to \ell^+\ell^-$ + jets	MadGraph			
Z/γ^* + 4 jets		27.4	6.0 M	218.7
Z/γ^* + 3 jets		60.7	10.6 M	175.0
Z/γ^* + 2 jets		215.0	2.3 M	10.7
Z/γ^* + 1 jet		666.3	23.7 M	35.6
single top	powheg			
t-channel t		56.4	3.7 M	65.6
t-channel $\bar{t}$		30.7	1.9 M	61.8
tW-channel t		11.1	489 k	44.1
tW-channel $\bar{t}$		11.1	493 k	44.5
s-channel t		3.8	245 k	64.6
s-channel $\bar{t}$		1.8	125 k	71.0

Table 3.2: Overview of the dedicated $t\bar{t}$ samples used for systematic studies. Showering and hadronisation is performed with PYTHIA for all these samples.

Sample	Generator	σ (pb)	# events	$\mathcal{L}$ (fb^{-1})
$t\bar{t}$ + jets variations	MadGraph			
Q^2 up, less ISR/FSR		245.8	5.0 M	20.3
Q^2 down, more ISR/FSR		245.8	5.4 M	21.9
ME-PS matching up		245.8	5.4 M	21.8
ME-PS matching down		245.8	5.5 M	19.9

identifying the standalone-muon track.

Proper muon candidates combining information from both the tracker detector and muon system can be obtained using two different methods. In case the muon identification starts from the standalone-muon tracks so-called *global muons* are reconstructed while the collection of tracker tracks gives rise to *tracker muons*. The global muon candidates are reconstructed by identifying a matching tracker track, for each standalone-muon track, by propagating both track parameters onto a common surface. Then for each pair the hits are combined into a global-muon track using an outside-in Kalman Filter. The identification of muon candidates as global muons is especially powerful when a high quality muon track was found in the muon detector. However in some cases it can occur that the standalone-muon reconstruction fails because of a lack of hits. This is most likely to happen in the presence of low transverse momentum muons which are unable to deposit sufficient energy in the muon spectrometer. Hence for these muons the tracker-muon reconstruction is very useful since it extrapolates all tracker tracks with transverse momentum $p_{\mathrm{T}} > 0.5$ GeV and total momentum $p > 2.5$ GeV to the muon system. If at least one muon segment corresponds with the extrapolated track, the tracker track fulfilled the tracker muon requirements and is identified as such.
Since both approaches have specific benefits they are combined in order to have a robust and highly efficient muon reconstruction (95-98%) throughout all energy ranges.

3.3.2 Electron reconstruction

The material budget of the tracker requires a dedicated electron-track reconstruction to correctly incorporate the energy loss caused by Brehmsstrahlung. In addition the electrons are severely influenced by the strong magnetic field which will smear the ini-

tial momentum of the electron, predominantly along the ϕ direction. Hence a more complex electron-reconstruction algorithm [59, 87] is required instead of the general Kalman Filter track-reconstruction approach. This resulted in the development of the electron-reconstruction algorithm based on a Gaussian Sum Filter (GSF) fit [88], which has as main benefit that it is capable of modelling changes in curvature radius throughout the different track layers. Unfortunately the GSF fit is rather CPU intensive and is therefore only be applied on a subset of track seeds defined as electron seeds.

The identification of the subset of electrons seeds relevant for the electron-track reconstruction can be performed by two different seeding algorithms: an ECAL-based or a tracker-based algorithm. The ECAL-based approach starts from the energy deposits recovered in the electromagnetic calorimeter and extrapolates back to the interaction vertex. In order to take into account the Brehmsstrahlung effects, the cluster is enlarged into a so-called supercluster and the extrapolation to the tracker is performed from the energy-weighted average position of this supercluster. The tracker seeds corresponding with hits of the extrapolated supercluster are then defined as electron seeds. The tracker-based approach on the contrary starts from charged-particle tracks reconstructed with the general Kalman Filter reconstruction algorithm. The tracker seeds are in this case obtained using an multivariate analysis method in order to only select the ones compatible with the electron-particle hypothesis.
After the tracker seeds have been identified, the specific electron-track fitting procedure can be applied. As mentioned above, this is done by a GSF fit which describes the energy loss in each tracker layer by a mixture of Gaussian distributions. Such a representation is advisable in the presence of Brehmsstrahlung since the normal Kalman Filter fit only assumes a single Gaussian energy loss distribution for a particle traversing the detector. The track fitting provides electron tracks up to the electromagnetic calorimeter such that the corresponding track parameters can be obtained at the ECAL surface allowing the estimation of the energy loss due to Brehmsstrahlung.

The GSF tracks recovered with this dedicated electron reconstruction algorithm are afterwards translated into actual electron candidates in two different ways: either based on a track-cluster association criterion or by the Particle Flow event reconstruction algorithm. The former one, which depends on the seeding method used, will be discussed here while the Particle Flow approach will be discussed in Section 3.3.3. In case the ECAL-based seeding algorithm is used for identifying the electron seeds, the electron track is associated with the supercluster used for the seed reconstruct based on a geometrical matching. For the tracker-based seeding algorithm the association is done with a Particle Flow cluster based on a MVA combining information on track observables and electron Particle Flow cluster observables.

3.3.3 The Particle-Flow event reconstruction algorithm

In order to reconstruct the direction, energy and type of all stable particles as accurately as possible the particle-flow (PF) event reconstruction algorithm combines the information of the different CMS subdetectors. The obtained collection of individual particles is then used to reconstruct jets and determine missing transverse energy. The

main benefit of the PF approach is the significant gain in efficiency by combining less precise subdetectors with more granular ones [89–91].
The PF algorithm uses a stepwize approach, starting by identifying fundamental elements such as charged-particle tracks, calorimeter clusters and muon tracks. Then the algorithm links these distinct building bricks of the different subdetectors topologically to construct specific building blocks. As a final step the building blocks are converted into stable particles.

Reconstructing and combining the fundamental elements

The building bricks used by the PF event reconstruction algorithm have to be measured with very high efficiency and a low fake rate because most of the stable particles have rather low momentum, even in very energetic collisions. The dedicated identification algorithms developed for the different subdetector bricks are optimised such to effectively deal with this demanding environment.
The iterative tracking algorithm used to reconstruct charged-particle tracks fulfils both requirements. The CMS tracking detector can be considered the cornerstone of the PF event reconstruction since it measures the momentum of charged hadrons with a higher resolution than the calorimeters and even provides a precise determination of the charged-particle direction at the production vertex before any influence from the magnetic field. The iterative tracking algorithm starts from very tight charged-particle seeds and progressively loosens the track seeding criteria. At each iteration hits assigned to the tracks found during the previous iteration are removed.
The calorimeter clusters are reconstructed in a high efficient and low fake rate manner using a clustering algorithm specifically developed for the PF event reconstruction. In this algorithm the seeds are defined as calorimeter cells with energy above a certain threshold. These cluster seeds are then transformed into so-called topological clusters by accumulating calorimeter cells adjacent to the cells present in the cluster. In order to suppress electronics noise the calorimeter cells are required to exceed a given energy threshold. Finally each topological cluster results in several particle-flow clusters, as much as cluster seeds present in the topological cluster.

Since each particle is expected to give rise to multiple building bricks a non-ambiguous linking algorithm that excludes any possible double-counting is applied. This algorithm connects elements presumed to correspond to the same particle and quantifies the quality of the linkage by the distance between the considered elements. For example a charged-particle track is linked with a PF calorimeter cluster if its extrapolated position lies within the cluster boundaries. This specific linking is also performed between charged-particle tracks and ECAL clusters in order to take into account the energy deposited by Bremsstrahlung photons emitted by electrons. Because the above explained clustering algorithm is performed separately in each of the calorimeter sub-detectors linking between different calorimeter clusters is also considered. In this case a linkage is established when the cluster position of the more granular calorimeter is within the cluster envelope of the less granular one. Finally the linking algorithm matches charged-particle tracks and muon tracks based on a global χ^2 track fit in order to create global muons.

Identifying stable particles

After establishing the fundamental elements and the linkages amongst them, the collection of stable particles is reconstructed by the particle-flow algorithm. This occurs gradually: first the PF muons and PF electrons are identified and from the remaining elements the charged hadrons, photons and neutral hadrons are distinguished.
The global and tracker muons can only be promoted to PF muons once the contamination from misidentified charged hadrons is removed. Both contributions can be distinguished based on different criteria, such that three specific selection procedures are applied. At first the so-called isolated selection is applied, which considers only global muons and has the loosest selection of all three since almost no additional neutral particles are expected to lie within their vicinity. The remaining muon candidates are passed to the PF-loose and PF-tight selection, which are developed to identify muons within jets. The PF-tight selection aims to reject hadronic punch-through[1] by combining information from the muon system and the calorimeters while the PF-loose selection tries to recover muon candidates that have a track momentum significantly larger than the corresponding calorimeter deposit, a combination incompatible with the charged hadron hypothesis.
The reconstruction of PF electrons starts from the GSF track for which the outermost track layer position is extrapolated to the ECAL and associated with the closest PF cluster, this to incorporate possible changes in the curvature due to Brehmsstrahlung. Afterwards the energy of the corresponding photon clusters are assigned to the total electron energy. Finally the electron candidates are distinguished from charged hadrons using a multivariate analysis based on variables related to energy and geometrical matching between the track and the cluster, two purely calorimeter-based variables and several genuine tracking quantities.

After the identification of the PF muon and PF electron candidates the remaining charged-particles tracks and PF calorimeter clusters are translated into charged hadrons, photons or neutral hadrons. Whenever a linkage can be performed between a particle track and a calorimeter cluster with compatible energy measurements, it is defined as a charged hadron candidate. In case the calorimeter measurement is larger the excess is assigned to a photon or a neutral hadron depending whether the cluster is found within the ECAL or HCAL, respectively. The collection of remaining calorimeter clusters unable to be linked with a charged particle track are also identified as photons or neutral hadrons.

3.3.4 Jet reconstruction

The reconstruction of jets is less straightforward compared to the other physics objects explained before because jets should be seen as a collection of hadronic activity combined into a single cone. However the event topology of interest, $t\bar{t} \rightarrow b\bar{b}q\bar{q}l\nu_l$, results into a final state containing four jets so reconstructing this object in a correct and accurate way is very important. Clustering algorithms are used to collect the showering

[1]Defined as hadron shower remnants penetrating through the calorimeters and reaching the muon system.

and hadronisation activity of these initial quarks and relate their transverse momenta with this of the jet. However, these reconstructed jets will be influenced by detector imperfections which might alter their transverse momentum. Hence a dedicated set of calibrations will need to be applied.

Jet clustering algorithm

Many different jet clustering algorithms exist but in this thesis only the cluster-based ones [92] will be used and hence explained. This type of jet clustering algorithms starts from a collection of stable partons or calorimeter cells and combines them into a cone with radius R. This clustering procedure uses a distance-based approach and looks for each object i whether another object j can be found within the predefined cone with radius R taking into account the transverse momentum $k_\perp$ of both objects.
The distance measures used in this jet clustering algorithm are given in Equations (3.10) and (3.11) where the first one defines the distance between the two objects while the second one represents the distance between the object i and the beam (B). Here $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, the (y, ϕ) distance between both objects and p can be interpreted as a parameter which controls the relative power between the energy and the geometrical scales.

$$d_{ij} = \min(k_{\perp i}^{2p}, k_{\perp j}^{2p}) \frac{\Delta_{ij}^2}{R^2} \tag{3.10}$$

$$d_{iB} = k_{\perp i}^{2p} \tag{3.11}$$

The value given to the parameter p defined in the two distance definitions, which governs the relative power of $k_\perp$ versus Δ_{ij}^2, results in different cluster-based jet algorithms; two of which are used in this thesis. When this parameter takes the value 1 the $k_\perp$ algorithm can be retrieved and in the case of $p = -1$ the jet algorithm is defined as the anti-$k_\perp$ algorithm, which is used for the jet clustering in this thesis. Within this jet clustering algorithm soft particles prefer to cluster with hard particles implying robust jet boundaries with respect to soft radiation. Both jet clustering algorithms are infrared and collinear safe, meaning that the created jet collection is not sensitive to soft emission and collinear splitting, respectively.

The jet clustering algorithm creates jets by looking for the smallest of these distances. Whenever the distance d_{ij} is smallest, the objects i and j are merged into a single object and stored as such in the list. However, in case the distance d_{iB} is smallest, the object i is removed from the list of input objects and categorised as a final jet. Afterwards the distances are recalculated and this procedure continues until no input objects remain.
The merging of two different objects into a single one is done using one of the existing recombination scheme, in the case of this thesis the E recombination scheme is used. This scheme calculates the four-momentum of the new object by simply adding the four-momentum of its constituents.

Jet energy calibration

The jet-reconstruction algorithm is developed such that the original generator-level particle can be related to the reconstructed PF jet. However the contribution of additional proton-proton interactions and influences of dead or badly functioning regions in the different subdetectors will slightly decrease the transverse momentum of the reconstructed jet. Within CMS a factorised approach has been adopted to resolve in a sequential way most of the undesired effects by dedicated jet energy scale (JES) calibrations [93, 94].
The main bulk of these JES calibrations are determined using purely simulated samples since this allows to also extract information for corners of phase space badly accessible in data and hence to obtain correction factors which are valid in a widespread range. The downside is that whenever discrepancies between data and simulation occur, additional correction factors have to be applied to simulation to ensure good agreement between both.
The jet energy calibrations which will be applied in this thesis are discussed in detail below. Within CMS additional correction terms exist concerning optional calibrations with minor contributions. They correspond to corrections for jet-flavour dependency of the jet response, for the underlying event and for the parton-level energy (L5-L7) and will not be considered here.

Figure 3.2: Schematical overview of the factorised approach adopted in CMS for incorporating jet energy calibrations for data and simulation (MC).

L1 pileup offset correction

The first contribution to the factorised correction chain aims to remove the additional energy deposits originating from pileup interactions in order to only maintain the high-p_{T} scattering. The corresponding correction term is determined purely from simulation and is based on the average difference in transverse momentum between matched jets with and without additional pileup interactions.

The offset energy needed to be subtracted from the jet energy is calculated with a *hybrid jet area method*, which uses the effective area of the jets multiplied by the average energy density in the event and represents the softness of the

jet activity. The full correction formula used at CMS is:

$$\mathcal{C}_{L1}\left(p_{\mathrm{T}}^{raw}, \eta, A_j, \rho\right) = 1 - A_j \frac{\{\rho_0(\eta) + \rho.\beta(\eta).\left[1 + \gamma(\eta).\log(p_{\mathrm{T}}^{raw})\right]\}}{p_{\mathrm{T}}^{raw}} \tag{3.12}$$

with η the jet pseudo-rapidity, A_j the jet area and ρ the per-event p_{T} offset density. The correction factors ρ_0, β and γ are determined in bins of η by fitting the offset function using uniquely matched reconstructed jets from the with-PU sample and without-PU sample. The correction formula is applied to both simulated and data events such that an additional scale factor should be taken into account in order to incorporate the small discrepancy between data and simulation. This scale factor is only applied for data events and is calculated from the PU-offset corrected transverse momenta of the reconstructed jets.

L2L3 Monte Carlo calibration

Now that the p_{T}-response of the jets is independent of the pileup, represented by the number of primary vertices, a following correction should be applied to ensure that the energy of the reconstructed jets corresponds on average with the generated jets at particle level and to obtain an η^{PF}-independent response. This second calibration is again completely simulation-based and is determined by inverting the p_{T}^{gen} and η^{PF} binned response.

$$\mathcal{C}_{L2L3}(p_{\mathrm{T,L1}}^{PF}) = \frac{1}{\left\langle \frac{p_{\mathrm{T,L1}}^{PF}}{p_{\mathrm{T}}^{gen}} \right\rangle (p_{\mathrm{T}}^{gen}, \eta^{PF})} \tag{3.13}$$

The correction factors are derived from a QCD multijet sample for which a jet reconstruction identical to the one used in data is applied.

L2Residual data-based relative (η-dependent) correction

This first data-based calibration, abbreviated to L2Res, aims at removing the residual difference in η^{PF}-dependence between data and simulation. For this dijet events with one jet contained in the barrel ($|\eta| < 1.3$) are used. This approach allows to correct the p_{T} response of all jets relative to the response of central jets based on the expected p_{T} balance between both jets in the event.

L3Residual data-based absolute (p_{T}-dependent) correction

After correcting the relative jet energy scale, the absolute jet energy scale should still be determined. This should only be done for central jets since the L2Res calibration ensures that these values can be used outside the barrel safely. The correction factors are determined using Z+jet (with $Z \to \mu^+\mu^-$ or e^+e^-) and γ+jet events since this allows to exploit the precise measurement of the Z and γ as reference objects.

The L3Res correction tackles the two main remaining differences between data and simulation: slightly lower response for data than for simulation and p_{T}-dependency for the ratio of data to simulation response. These two contributions are independent and can thus be factorised as well. The constant scale factor is determined from the very precise $Z \to \mu^+\mu^-$+jet events while the p_{T}-dependent correction factors are obtained by combining the response of all different decay channels in a global fit.

Each subcorrection in this factorised approach has specific JES uncertainties which are combined into a systematic uncertainty used in physics analyses. The overall uncertainty is obtained by quadratically adding the uncertainties of each level and is dominated by the p_{T}-dependent difference in pileup offset between data and simulation, the uncertainty of the jet p_{T} resolution and the lepton/photon scale uncertainties for the L1, L2Res and L3Res subcorrection, respectively.

Jet energy resolutions

Besides the jet energy scale also the transverse momentum or energy resolution (JER) can be influenced by discrepancies between data and simulation. The measurement of the JER is performed using similar methods as applied for determining the JES, but instead of looking at the mean of the response distribution the width is considered. The jet p_{T} resolution is determined using QCD dijet and γ+jet events and results in an η-dependent data/MC scale factor exceeding unity [93]. Hence the p_{T} resolution is about 10% worse for data than for simulation in the barrel, a value which quickly raises to roughly 25% -on average- in the endcap. In order to account for this difference, the resolution for simulation events is worsened by smearing the energy of the corrected PF jets.

3.3.5 Identification of b-quark jets

Identifying the jets originating from b-quark decays consists of constructing observables in order to exploit the differences between b-quark jets and light jets. The algorithms developed for this purpose, many exist in literature, are capable of distinguishing the event topology of interest from the large bulk of background events which only contain light-parton jets. The different b-jet-identification or b-tagging algorithms rely on the reconstructed physics objects although some minor optimisation requirements are implied for the track selection to improve efficiency [95].
One of the main b-quark jet characteristics is the relatively long lifetime of the b-hadron resulting in the presence of a displaced vertex with respect to the interaction point. Since only the tracking detectors offer the spatial resolution needed to detect the displacement between the primary and secondary vertices, they are reconstructed purely from the track collection. In order to be able to cope with multiple proton-proton interactions the tracks are required to be within a cone of $\Delta R = 0.3$ around the jet axis, defined by the direction of the jet momentum. The actual reconstruction of secondary vertices is an iterative process using an adaptive vertex fit. This fit algorithm estimates the position of the vertex candidate and removes all its associated tracks from the track collection. This fit procedure is repeated until no new vertex candidates can be found. During the first iteration the interaction point is used as a constraint in order to identify the prompt[2] tracks.

The various b-tagging algorithms existing today exploit the properties of the long-lived B meson inside the b-quark jet. These properties result in b-quark jets containing

[2]Prompt tracks are tracks originating near the pp interaction point.

displaced tracks, reconstructed secondary vertices, soft muons or electrons, or a combination thereof. Within this thesis, the Combined Secondary Vertex algorithm is used, which combines the secondary vertex information with the track-based lifetime properties. Because both characteristics are combined the algorithm is also capable of discriminating between b-quark and light-parton jets when no secondary vertex was reconstructed.
Typically, jets are identified or 'tagged' as originating from b quarks by applying a criterion on the b-tagging discriminator output. Three operating points are defined; Loose, Medium and Tight. These operating points correspond to a threshold on the b-tagging discriminator resulting in a misidentification probability for non-b quark jets of about 10% (*Loose*), 1% (*Medium*) and 0.1% (*Tight*) derived by the CMS collaboration at the beginning of data taking on sample of simulated multijet events. In this analysis only the Tight operating point will be considered.

3.3.6 Missing transverse energy

The vast majority of physical objects produced in particle collisions can be reconstructed from the collection of energy deposits. However neutrinos are the exception to the rule since they are weakly interacting and carry neutral charge. Therefore they will traverse the entire detector and escape detection rendering an accurate reconstruction rather challenging.
So in order to ensure the reconstruction of neutrinos or other hypothetical neutral weakly interacting particles, a signal extremely important for many physics analyses, a specific work-around is applied which is based on indirect observations rather than direct measurements. The solution lies to a great extent in the geometrical characteristics of the particle detectors, by requiring them to be hermetically closed such that all other particles are properly detected and cannot leave the detector unseen. Thus the missing transverse momentum, which corresponds to all neutrinos and other weakly interacting neutral particles present in the event, can be defined from the total transverse momentum of all observed final-state particles [96]. This procedure allows to exploit the high reconstruction efficiency of the particle-flow algorithm for the neutrino reconstruction.

$$\vec{\not{E}}_{T,raw} = -\sum \vec{p_{\mathrm{T}}} \tag{3.14}$$

where the sum runs over all reconstructed PF particles.
The imperfect jet energy response in simulation should also be propagated to the $\vec{\not{E}}_T$ by applying the L2L3 jet energy scale corrections, which have been discussed in Section 3.3.4.

$$\vec{\not{E}}_T = \vec{\not{E}}_{T,raw} - \sum \left(\vec{p}^{\,PFjet}_{\mathrm{T},L1L2L3} - \vec{p}^{\,PFjet}_{\mathrm{T},L1} \right) \tag{3.15}$$

where the sum now runs over all PF jets with $p^{PFjet}_{\mathrm{T},L1L2L3} > 10\,\mathrm{GeV}$.

Chapter 4

Event selection and top-quark pair topology reconstruction

From the large number of proton-proton collisions produced at the LHC only a small fraction corresponds to interaction of interest. As a result the biggest challenge of any physics analysis is to obtain, besides a successful object identification and reconstruction, an efficient separation of the event topology of interest from the large bulk of background events. This can be achieved by developing an effective event-selection procedure that excludes events based on specific kinematic and topological requirements, as will be demonstrated in this chapter. Such a procedure starts off with some basic identification and cleaning conditions and proceeds by restricting the kinematics in order to perfectly correspond with the analysis-specific requirements. Hence the number of selected events is reduced gradually until the topology of the remaining events is most likely to correspond with the desired one.

This chapter will focus in Section 4.1 on the general kinematic requirements needed to be fulfilled in order to select semi-leptonical decaying top-quark pair topologies. Section 4.2 will then discuss the various additional selection conditions introduced in order to optimise the event-selection purity and the topology-reconstruction efficiency for this specific analysis. To finalise, Section 4.3 will give an overview of the applied event selection and will demonstrate the obtained agreement between data and simulation.

4.1 Baseline event selection

The goal of the event selection is to keep only the event topologies compatible with the considered decay process and reject background contributions and detector noise mimicking the signal signature. Hence a dedicated selection and cleaning procedure is applied by combining the online trigger system with an offline event selection in order to reduce the stored event rate by specifying the type of final state particles interested in.

4.1.1 Triggering and cleaning of events

As was already briefly mentioned in Chapter 2, CMS possesses a complex trigger system that decides whether the considered event is deemed interesting enough to be stored and processed further. This trigger system uses an exhaustive list of distinct trigger paths, all designed to single out a specific type of final state signature and thus drastically reduce the event rate.

In this analysis the event topology of interest is that of semi-leptonical ($l = \mu$) decaying top-quark pairs, which can be distinguished rather efficiently from background by demanding each event to contain a muon. Such a muon signature is rather distinct and will reduce a large portion of the background, dominated by low-energetic jet processes. Therefore the trigger path applied in this thesis only keeps events with at least one muon for which the kinematic requirements fulfil $p_{\mathrm{T}} > 24\,\mathrm{GeV}$ and $|\eta| < 2.1$.
The different objects retained by the applied trigger path need to pass a couple of cleaning requirements in order to reduce the contribution of electronic noise mimicking the signatures looked for. The first one affects both data and simulation, and ensures that the considered interaction corresponds to a proton-proton collision by demanding that the main primary vertex be recovered within a cylinder of radius 2 cm and length 24 cm around the nominal interaction point. The following cleaning procedures only have to be applied on data, since they verify that data is only recorded when the detector was completely turned on and all subdetectors are working properly.

Since for the trigger efficiencies an excellent agreement has been obtained between data and simulation [97], no correction factor will be applied to simulation.

4.1.2 Lepton selection criteria

The applied trigger path is not specifically developed for identifying top-quark pairs decaying semi-leptonical, such that additional selection criteria are required for the leptons in order to further exclude incorrect event signatures. As a result the kinematic requirements are tightened and require: $p_{\mathrm{T}} > 26\,\mathrm{GeV}$ and $|\eta| < 2.1$.
Still, additional lepton selection criteria are necessary in order to ensure that the stored muon is a well-defined one. These so-called muon identification criteria start from Particle-Flow muons, which have been discussed in Section 3.3.3, and are designed to suppress hadronic punch-through, cosmic muons and muons from decays in flight of other particles. They require the candidate muon to be reconstructed as a global one, and the global-muon track fit, with normalised $\chi^2 < 10$, to contain at least one muon chamber hit. Moreover the muon track should have a minimum of two muon stations with matched segments, contain at least one pixel hit and have more than five tracker layers which have been hit. The latter requirement will guarantee, besides suppressing muons from decays in flight, a good p_{T} measurement for the muon. Finally muon candidates not originating from the primary vertex are rejected by limiting both the longitudinal and transverse impact parameter: $|d_0| < 0.2\,\mathrm{cm}$ and $|\Delta z| < 0.5\,\mathrm{cm}$.
Another important identification criterion is the isolation variable which allows to distinguish prompt muons with high purity from the ones embedded in jets by taking into

account the hadronic activity around the muon candidate. It is defined as the scalar sum of the transverse energy of all the reconstructed particles contained within a cone of radius $\Delta R = 0.4$, excluding the contribution of the muon itself. In order to ensure a correct treatment of the large number of additional proton-proton interactions the isolation variable has to be corrected. For this reason a $\Delta\beta$-corrected isolation variable has been developed, which includes for the charged hadrons (CH) only the partons associated with the primary vertex while for the neutral ones (NH and γ) the estimated PU contribution is subtracted. This contribution can be calculated by halving the PU contribution for charged particles since jets contain on average twice more charged PF particles than neutral ones [98]. The formula to determine this $\Delta\beta$-corrected isolation variable is given in Equation (4.1) and in this analysis $I_{\text{rel}}^{\Delta\beta}$ is required to be smaller than 0.12 in order to guarantee the reconstructed muon is well isolated.

$$I_{\text{rel}}^{\Delta\beta} = \frac{1}{p_{\text{T}}^{\mu}} \left(\sum_{\text{CH}} p_{\text{T}}^{\text{CH}} + \max(0, \sum_{\text{NH}} p_{\text{T}}^{NH} + \sum_{\gamma} p_{\text{T}}^{\gamma} - 0.5 \sum_{PU} p_{\text{T}}^{PU}) \right) \tag{4.1}$$

The events considered in this analysis are required to contain exactly one such well-identified muon, with $p_{\text{T}} > 26\,\text{GeV}$ and $|\eta| < 2.1$. Any event containing an additional lepton, either a PF muon reconstructed as global or tracker muon with $p_{\text{T}} > 10\,\text{GeV}$, $|\eta| < 2.5$ and $I_{\text{rel}}^{\Delta\beta} < 0.2$ or otherwise an electron with $p_{\text{T}} > 20\,\text{GeV}$, $|\eta| < 2.5$, $I_{\text{rel}}^{EA} < 0.15$ and $mvaId > 0$, are rejected.
The latter two variables, I_{rel}^{EA} and $mvaId$, are two of the electron selection criteria that are applied to all electrons present in the event in order to limit the influence of background sources for the electron identification [99]. In contrast to the muon identification, which is based on a cut-based method, the electron identification uses a multivariate approach where different variables are combined into the single $mvaId$ variable. Secondly, the electron-isolation I_{rel}^{EA} is determined in a similar manner as the muon-isolation, but using a different cone size ($\Delta R = 0.3$) and a different pileup subtraction approach. For the electrons this is done by using an effective area A_{eff} and an average energy density ρ, which has proven to be rather efficient in reducing the pileup dependency.

$$I_{\text{rel}}^{EA} = \frac{1}{p_{\text{T}}^{e}} \left(\sum_{\text{CH}} p_{\text{T}}^{\text{CH}} + \max(0, \sum_{\text{NH}} p_{\text{T}}^{NH} + \sum_{\gamma} p_{\text{T}}^{\gamma} - \rho \cdot A_{eff}) \right) \tag{4.2}$$

The lepton selection is then finalised by introducing a correction-factor taking into account the trigger, identification and isolation efficiencies of the selected muons. These values have been determined centrally by the CMS collaboration [97], and, although these efficiencies are almost identical for data and simulation, the correspondig correction factors will be applied in order to ensure an optimal agreement between both.

4.1.3 Jet selection criteria

With the lepton selection clearly established, the next step consists of applying a similar type of identification and cleaning criteria to the selected jets. The goal here is to drastically reduce the fake, badly reconstructed and noise jets while keeping close to 99

% of the real jets. These jet-identification criteria should be applied on the PF jets after the L1L2L3 correction, the jet-energy smearing (both discussed in Section 3.3.4) and the charged hadron subtraction responsible of removing all contributions from charged pileup have been taken into account. These calibrations are needed in order to correct for the small discrepancies observed between data and simulation.

In order to select top-quark pairs decaying semi-leptonical, characterised by a well-isolated lepton and four high-energetic jets, each event has to contain at least four jets fulfilling the requirements: $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.4$. Since each of these jets have to be well separated from the muon identified in the event, all jets for which the ΔR based distance with the muon is lower than 0.3 will be rejected.
The actual jet-identification criteria look at the distribution of the energy fractions and the composition of the different jet constituents. They reject the noise jets by constraining the energy fraction carried by the charged electromagnetic PF particles ($f_{CEM} < 0.99$), the energy fraction carried by the charged PF hadrons ($f_{CH} > 0$), the energy fraction carried by the neutral electromagnetic PF particles ($f_{NEM} < 0.99$) and the energy fraction carried by the neutral PF hadrons ($f_{NH} < 0.99$). Moreover, each PF jet is required to contain at least two constituents ($n_{\mathrm{PFparticles}} > 1$) and at least one charged particle ($n_{\mathrm{charged}} > 0$).

4.2 Analysis-specific event-selection criteria

The event selection criteria discussed in the previous section are kept as general as possible in order to be applicable for various analyses examining similar event topologies. These general selection and cleaning criteria should however be optimised in order to incorporate the necessary analysis-specific requirements. In this analysis, this optimisation is achieved by applying three separate event-selection criteria, which aim to exclude specific types of events. The first one, discussed in Section 4.2.1, exploits the characteristic signature of top-quark pair events for which two of the jets originate from the decay of a b-quark. Afterwards in Section 4.2.2, the criterion used to identify the two remaining jets in the event and decide on the most optimal jet combination will be discussed. Finally Section 4.2.3 focusses more on the kinematic properties of the reconstructed jets and the reduction of the background contributions that can be achieved by restricting the invariant masses of the W-boson and top-quark candidates.

The analysis discussed in this thesis requires a very stringent event selection due to the choice of using a Matrix Element method. This method will be explained in detail in Chapter 5, but for the event-selection optimisation it is sufficient to keep in mind that such a technique examines each event using the full kinematic information, thus requiring a significant processing time. As a result, it has been opted for to restrict the selected number of events as much as possible to avoid spending computational resources on incorrect event topologies.

4.2.1 Background reduction using b-jet identification

Exploiting the presence of two jets originating from the decay of a b-quark in $t\bar{t}$ events is an effective manner of distinguishing the event topology from the background, since this type of decay has the peculiar feature that it gives rise to a displaced vertex. This because the relatively long lifetime of the B-mesons in b-quark decays implies that the decay does not occur at the interaction vertex, as has been explained in Section 3.3.4. In this analysis b-jet identification plays a crucial role in reducing the background contribution since only events with two jets identified as b-jets will be considered. The main background processes for semi-leptonical decaying top-quark events might have events with one jet fulfilling this condition, having two is less likely. Hence the considered background samples; W-boson production in association with jets (W+jets), Z-boson production in association with jets (Z+jets) and single-top production in the t-, tW- and s-channel; will almost be completely negligible after applying this b-tagging requirement.

The b-jet identification algorithms developed by the CMS collaboration are recommended to only be deployed at specific working points, defined as *Loose*, *Medium* and *Tight*. Here the Combined Secondary Vertex (CSV) b-tagging algorithm has been considered, for which these working points correspond to a discriminant value of 0.244, 0.679 and 0.898; a tagging probability of around 85%, 69% and 52%; and a misidentification one of 19%, 5% and 1%; respectively. The impressive efficiency for this b-jet identification procedure can be understood by looking at the distribution of the CSV discriminant for different jet flavours given in Figure 4.1, allowing for a clear distinction between the b-flavoured and light-flavoured jets.

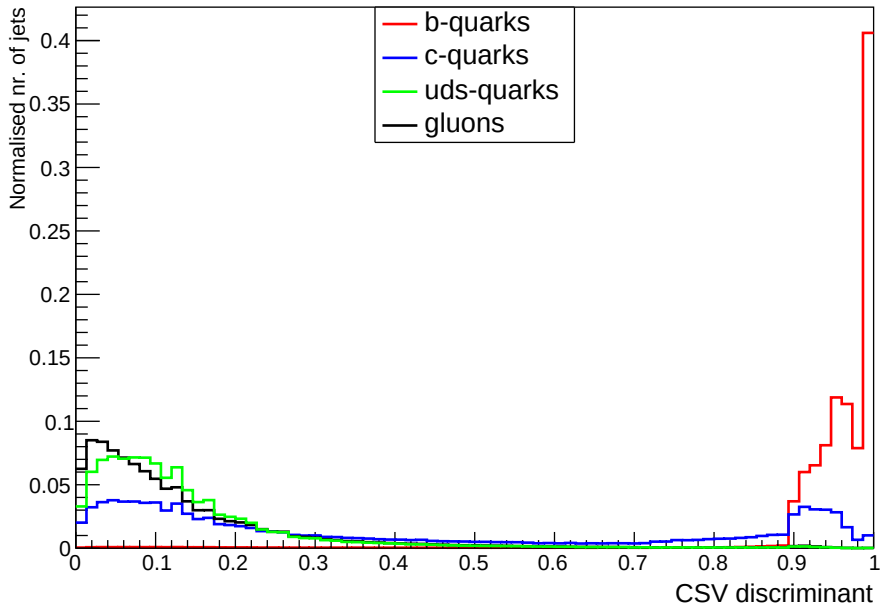

Figure 4.1: Combined Secondary Vertex b-tag discriminant for the different jet-flavours.

Since in this analysis priority is given to selecting event topologies that closely agree with the expected topology the double Tight CSV b-tagging requirement is expected

to be the optimal choice. This is indeed the case, the fraction of selecting the good jet combination increases from 0.409 up to 0.497 and finally reaches a value of 0.547 when tightening the CSV discriminant value from Loose to Tight. These fractions are determined from simulated semi-leptonic $t\bar{t}$ events for which information on the four generator-level partons corresponding to the four jets is available.
It has also been investigated whether an improvement could be observed when the CSV discriminant of the light-jet candidates is restricted, but since the effect was almost negligible it will not be considered further.

After this double b-tagging requirement the background samples have been significantly reduced and the only remaining contribution comes from the single-top quark events, and then in particular the tW-channel. The reason why this specific background process becomes relevant can be explained from a complex interplay between top-quark pair and single-top decays in the tW channel. This because the tW-channel single-top processes are well described at leading order but at next-to-leading order a set of Feynman diagrams is shared among both. For this reason, a diagram-removal approach is applied in order to reduce the effect of this overlap by not adding these types of diagrams in the tW-signal definition [100].

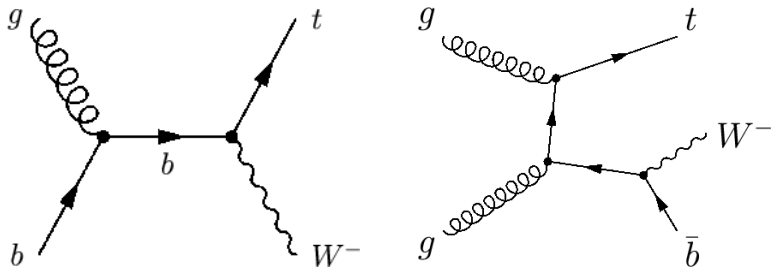

Figure 4.2: Leading-order (left) and next-to-leading (right) order Feynman diagram for tW production. The latter one corresponds to a leading-order diagram for top-quark pair production.

The application of this double Tight b-tagging algorithm is also the motivation why, besides the background samples discussed before, no other background contributions have been considered for this analysis as they will become completely negligible.

4.2.2 Determining the optimal jet combination

With the b-tag working point and the number of b-tags decided on, the topology reconstruction proceeds by assigning the selected jets to the final state particles expected in semi-muonic $t\bar{t}$ events. This reconstruction procedure first identifies the two most energetic jets with sufficiently high CSV-discriminant value and labels all remaining jets as originating from light flavoured partons.
For the two b-jet candidates should then be decided whether they originate from the top quark for which the W-boson decays into two jets, denoted as the hadronically decaying top quark, or from the top quark with the produced W-boson decaying into a muon and corresponding neutrino, the so-called leptonically decaying top quark. This

identification is important since it allows to reduce the number of permutations needed to be considered by the Matrix Element method. From the collection of light jets only the two expected to originate from the hadronical decaying W-boson have to be identified, but not matched with a specific jet since the up- and down-type jet are impossible to be distinguished.

In this analysis, for which a high topology-reconstruction efficiency is desired, it has been studied whether an improvement can be achieved when the two light-jet candidates are chosen from the three leading p_{T} jets based on their likelihood to originate from a W-boson decay instead of using the more general approach of merely continuing with the two most energetic light jets. This light-jet selection is performed simultaneously with the b-jet assignment and is based on a χ^2-method using both the invariant mass of the lepton and the b-jet originating from the leptonical decaying top quark, denoted as m_{lb}, and the invariant mass of the full hadronic decaying top-quark system, m_{qqb} or m_t.
The expected values of these invariant masses have been determined by applying a Gaussian fit on the distribution obtained for all events passing the above-mentioned event selection requirements, resulting in $\hat{m}_{lb} = 108 \pm 32\,\mathrm{GeV}$ and $\hat{m}_{qqb} = 175 \pm 17\,\mathrm{GeV}$. The benefit of using the m_{lb} variable instead of determining the invariant mass of the full leptonic decaying top-quark system is that the reconstruction of the neutrino can be avoided. A small downside is the non-perfect Gaussian behaviour of this invariant mass distribution such that the fit had to be carefully applied onto a limited range of the distribution. The obtained invariant mass distributions together with the Gaussian fit function are both given in Figure 4.3.

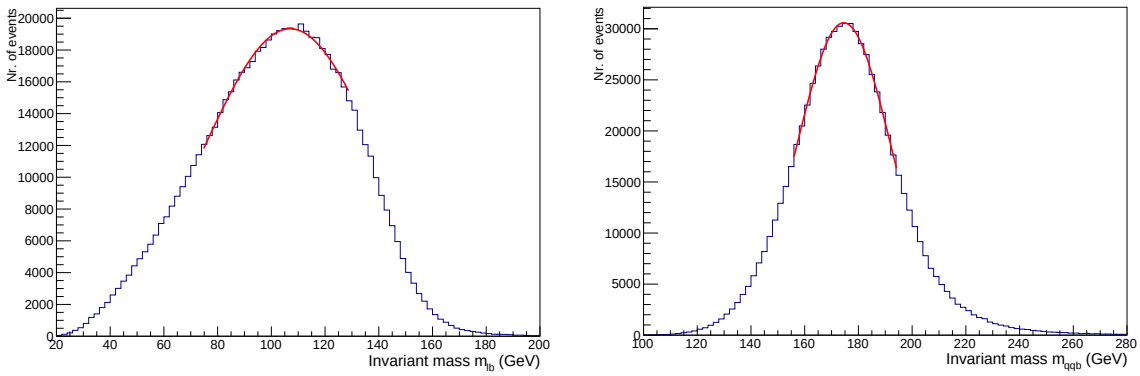

Figure 4.3: Distribution of the invariant masses, m_{lb} on the left and m_{qqb} on the right, together with the applied Gaussian fit function. The considered fit range corresponds to the 1σ interval.

In this analysis the most plausible jet assignment for the b-jets and the light jets is determined with the χ^2-based procedure for which the formula is given in Equation (4.3). For each event the χ^2_i values of the 6 possible jet combinations are calculated and the combination with lowest χ^2_i value is selected. In case the considered event does not have a third jet present in the event, only the permutation between the hadronic

and leptonic b-jet has to be taken into account.

$$\chi_i^2 = \frac{(\hat{m}_{lb} - m_{lb,i})^2}{\sigma^2(\hat{m}_{lb})} + \frac{(\hat{m}_{qqb} - m_{qqb,i})^2}{\sigma^2(\hat{m}_{qqb})} \tag{4.3}$$

Allowing the third light jet to be part of the chosen jet combination significantly improved the topology-reconstruction efficiency and will therefore be applied in this analysis. With these extra jet combinations to choose the most plausible one from, the possibility to select the correct two light jets increases from 63.13% to 76.21%. Due to this more correct determination of the light jet candidates the overall fraction of selecting good events, mentioned earlier to be 0.547, has improved to 0.666. This positive influence can also be seen when comparing the two distributions given in Figure 4.4. Here the left figure contains the χ^2 distribution for the chosen jet combination with and without the inclusion of this third jet, while the right figure shows the difference in shape of this χ^2 distribution for the non-chosen jet combinations.

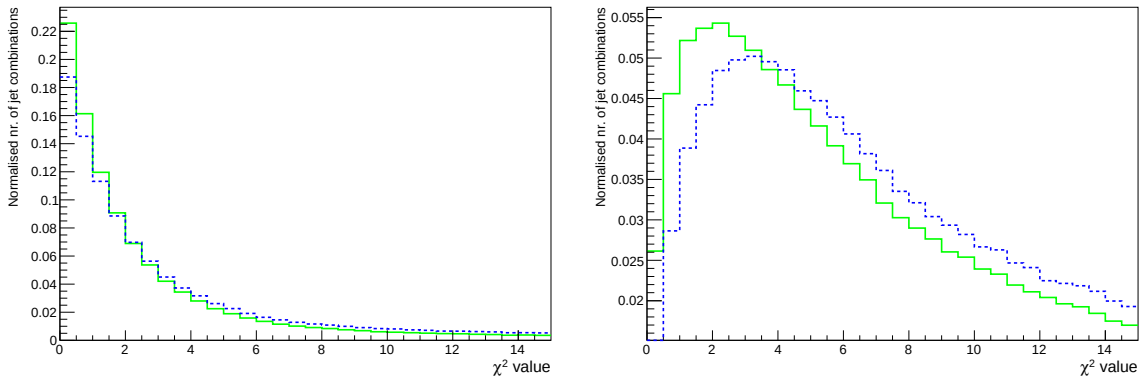

Figure 4.4: Comparison of the χ^2 value of the chosen (left) and non-chosen (right) jet combinations. The solid green line represents the case where the third light-quark jet is allowed to be one of the candidate jets and the dashed blue corresponds to the general approach of selecting the two leading light-quark jets.

4.2.3 Improving the topology reconstruction

Even though the application of a b-jet identification algorithm will significantly reduce the different background contributions, some improvement can still be obtained by excluding events using specific kinematic criteria, as will be demonstrated here. This will also improve the topology reconstruction, which is important for this analysis since the Matrix Element method treats all events as it were semi-leptonic $t\bar{t}$ events. Hence this technique behaves optimally in case more events with the correct event topology are selected.

Therefore two additional event selection criteria have been considered, both relying on the invariant mass distributions. At first the hadronical decaying top-quark system is restricted by requiring the reconstructed invariant masses, m_t and m_W, to lie within a predefined number of standard deviations from the expectation. Secondly a limitation of the χ^2 variable, using information from both top quarks as shown in Equation

(4.3), has been studied. As these two procedures use an overlapping variable to reject undesired events; *i.e.* the invariant mass M_{qqb}; a clear correlation between the two can be observed.

Restricting the mass values of both the hadronical decaying top quark and the W-boson is a rather effective manner of enhancing the probability that the reconstructed event topology actually corresponds to this of a top-quark decay. The expected W-boson mass has been determined by fitting the obtained invariant mass distribution after all event selection criteria have been applied with a Gaussian function, as was the case for the top-quark mass in the χ^2-based method ($\hat{m}_t = 175 \pm 17\,\text{GeV}$). For the W-boson mass this resulted in a value of $\hat{m}_W = 84 \pm 10\,\text{GeV}$, retrieved from the distribution given in Figure 4.5.

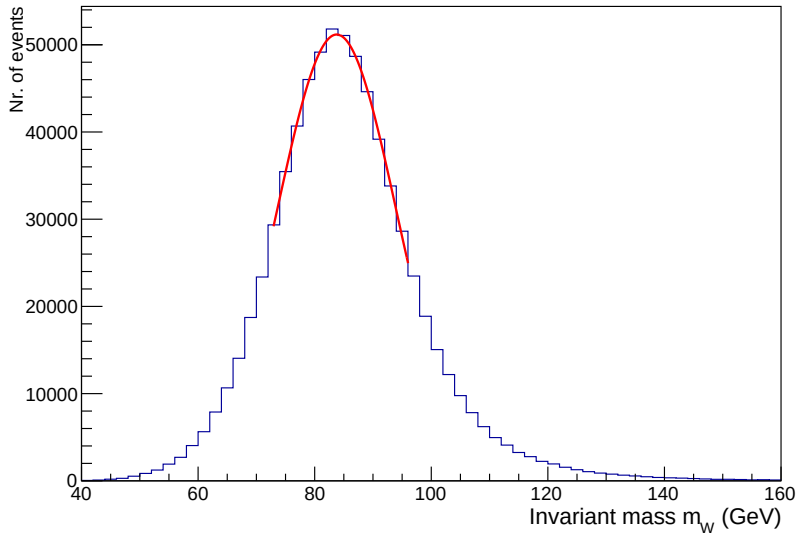

Figure 4.5: Distribution of the invariant mass of the hadronic W-boson, m_W, together with the applied Gaussian fit function. Also here, the considered fit range corresponds to the 1σ interval.

The invariant mass of the W-boson originating from the hadronically decaying top-quark provides a clear distinction between events for which the four jets have been correctly matched with the generator-level parton and events for which at least one jet-assignment, excluding the light-jet permutation, has been done wrongly. Hence from Figure 4.6 can be concluded that a clear improvement in topology reconstruction efficiency can be achieved when restricting this invariant mass. For the top-quark the effect is less significant since its mass is already restricted by the procedure to select the most optimal jet-combination, using the χ^2 formula given in Equation (4.3). Nevertheless, both mass constraint will be applied for completeness since this allows to exclude a small fraction of events for which the W-boson mass is correctly reconstructed but the inclusion of the b-quark jet ruins the overall invariant mass.
The optimal range for these invariant masses, chosen to sufficiently improve the topology-reconstruction efficiency without reducing the event count too drastically, has been

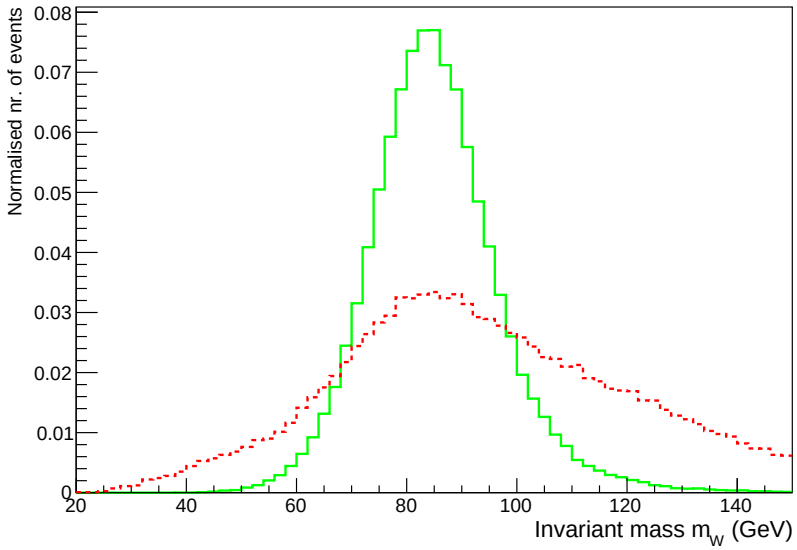

Figure 4.6: Distribution of the invariant mass of the two light-quark jets originating from the W-boson decay for events with correctly reconstructed (solid green) and wrongly reconstructed event topology (dashed red).

identified as the 2σ interval. For this mass range the number of selected semi-leptontic $t\bar{t}$ events fulfilling the double b-tag requirement is halved while the reconstruction fraction has increased to 0.791.

The second considered requirement, limiting the χ^2 variable of the chosen jet combination, aims to exclude the events unlikely to be properly reconstructed. For the events residing in the tail of this χ^2 distribution the possibility to correctly assign all four jets to the considered event topology is as low as 20 %. In order to ensure that the χ^2 restriction results in a cleaner event sample without rejecting the majority of the events, the cut-value to be fulfilled for the chosen jet combination is chosen as $\chi^2 < 10$. Without the mass restrictions being applied, this requirement slightly enhances the reconstruction to 0.699 while reducing the event count for the considered simulated semi-leptonic $t\bar{t}$ sample with 23 %.

The overall effect of these two additional event selection criteria will have less of an impact on the final result than the b-jet identification algorithm, but is nonetheless an efficient way to reduce the contribution of poorly reconstructed events. When demanding both the invariant mass and the χ^2 restriction to be fulfilled, the possibility to correctly reconstruct the entire top-quark pair event topology reaches a value of 79%. The small improvement compared to the individual results, mainly dominated by the mass requirement, can be understood from the large correlation between both methods.

4.3 Influence of full event selection

The full event selection consists of the triggering and cleaning conditions discussed in Section 4.1 combined with the analysis-specific supplementary selection criteria introduced in Section 4.2. The additional event selection criteria applied here have all been developed with the intention to reduce the contribution from the different background processes while remaining with a selected sample dominated by well reconstructed events. This goal is clearly achieved under the influence of the considered selection requirements since for 79% of the semi-leptonic $t\bar{t}$ events the correct jet assignment is chosen.

This full set of event-selection criteria is applied both for data and simulation and, in order to ensure the data is well described by the simulation, a detailed comparison between the two is performed. In order for this comparison to be correct, the different calibrations accounting for the different response of data and simulated events in specific conditions have to be applied.
As has been explained in Section 3.2, the simulation samples are generated using a distribution for the number of pileup interactions meant to correspond with the expected conditions for the considered data-taking period. However, these conditions might vary slightly and thus need to be corrected for once the exact pileup information in data is known. Therefore the number of pileup interactions in simulation is reweighted to the distribution observed in data using a luminosity-based pileup estimate. This method determines the total number of interactions in each bunch crossing from the measured instantaneous luminosity for the corresponding bunch crossing and the total inelastic proton-proton collision cross-section.
The second calibration that needs to be applied corrects for the slightly different efficiency of the b-jet identification algorithm in data and simulation [101]. This aspect has been studied in great detail and resulted in the development of b-tag efficiency scale-factors. These are analytic functions depending on the transverse momentum of the jets, to account for the discrepancies by correcting the measured jet-dependent efficiencies in simulation.
The determination of the b-tag event-weight starts by computing the efficiency ε_f for a jet of flavour f to be identified as a b-jet using simulated semi-leptonic $t\bar{t}$ events. From these efficiencies an overall event probability P_{sim} can be calculated by multiplying the efficiencies of all jets residing in the event[1], as shown in Equation (4.4). The corresponding event probability as expected in data, P_{data} in Equation (4.5), is determined in the same way but every jet-efficiency is corrected using the aforementioned scale-factor SF.

$$P_{sim} = \prod_{i=tagged} \varepsilon_i \prod_{j=non-tagged} (1-\varepsilon_j) \tag{4.4}$$

$$P_{data} = \prod_{i=tagged} SF_i\,\varepsilon_i \prod_{j=non-tagged} (1-SF_j\,\varepsilon_j) \tag{4.5}$$

[1]It is important to note that in Equation (4.4) the labels *tagged* and *non-tagged* refer to whether the considered jet fulfilled the applied b-tag requirement and not whether it has been labelled as one of the two b-jet or light jet candidates in the event selection.

From these two probabilities the overall event-weight can be determined in a straightforward manner by comparing the probability for data with the one in simulation:

$$\text{event-weight } w = \frac{P_{data}}{P_{sim}} \tag{4.6}$$

The two considered calibrations both influence the individual event kinematics and therefore modify the shape of the kinematic distributions. The overall event count obtained in simulation, however, is not affected by the pileup reweighting since this correction factor averages out to unity. The importance of this correction can be visualised in Figure 4.7, which contains the primary vertex multiplicity with and without this specific reweighting applied. Comparing the two distributions clearly illustrates a better shape description, and thus a significantly improved agreement between data and simulation, when taking into account the calibration factor.
The b-tag event-weight, on the other hand, has an average value of 0.87 originating from the stringent requirement of demanding two jets to fulfil the Tight working-point requirement of the CSV algorithm. Hence the event count in simulation has largely been overestimated and is therefore scaled down in order to correctly take into account the difference in b-jet identification efficiency.

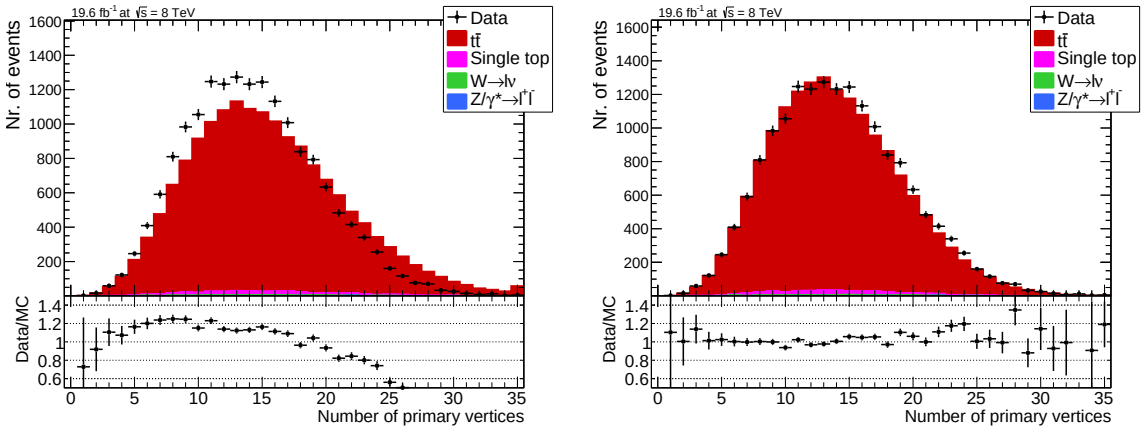

Figure 4.7: Distribution of the number of primary vertices before and after the pileup reweighting is applied.

Once all necessary calibrations have been included the actual comparison between data and simulation can be performed at a luminosity of $19.6 \pm 0.5\,\text{fb}^{-1}$, which is the amount of data collected by the considered isolated-muon triggers during the 2012 data-taking period. The final number of events obtained after demanding both the data and the simulated samples pass through the complete event-selection chain are given in Table 4.1. As expected, the background contribution is very small due to the stringent event-selection criteria.
This table confirms the good agreement between data and simulation after applying the full event-selection chain and introducing the necessary calibration factors shown in the right distribution in Figure 4.7. This agreement between data and simulation can also be observed in other kinematic variables, and some of the more relevant ones

Table 4.1: Event count obtained after applying the full event-selection chain for data together with the expected values and corresponding statistical uncertainties in simulation. The quoted uncertainties are purely statistical and demonstrate the limited size of the different samples.

Sample	Number of events
$t\bar{t}$	16173 ± 41
W+jets	66 ± 7
Z+jets	17 ± 4
Single top	359 ± 24
Total	16 615 ± 48
Observed	16943

for the discussed analysis will be given below.

The first set of distributions that are displayed in Figure 4.8 show the variables used during the final step of the event-selection chain, discussed in Section 4.2.3. Hence the distribution of the χ^2 variable used to select the most optimal jet combination and both the distribution of the invariant mass of the hadronically decaying top quark and hadronically decaying W-boson are given here. In order to clearly show the effect these additional event-selection criteria have on the different simulated samples and data, the distributions will be shown before the cuts have been applied. In the final analysis the χ^2 is required to be lower than 10, the top-quark mass has to be reconstructed between 141 and 209 GeV and the reconstructed mass of the W-boson should be between 63 and 104 GeV, both corresponding to their respective 2σ-range. Within the range that will be kept for further analysis, an very good agreement between simulation and data can be observed.
Afterwards, in Figures 4.9 up to 4.11, the distribution of the transverse momentum p_{T} and the pseudo-rapidity η of the selected lepton and four reconstructed jets can be found. These two kinematic variables have been chosen since they serve as the input for the Matrix Element method, together with the azimuthal angle ϕ, as will be discussed in detail in Chapter 5. Also here no significant discrepancy between data and simulation can be observed.

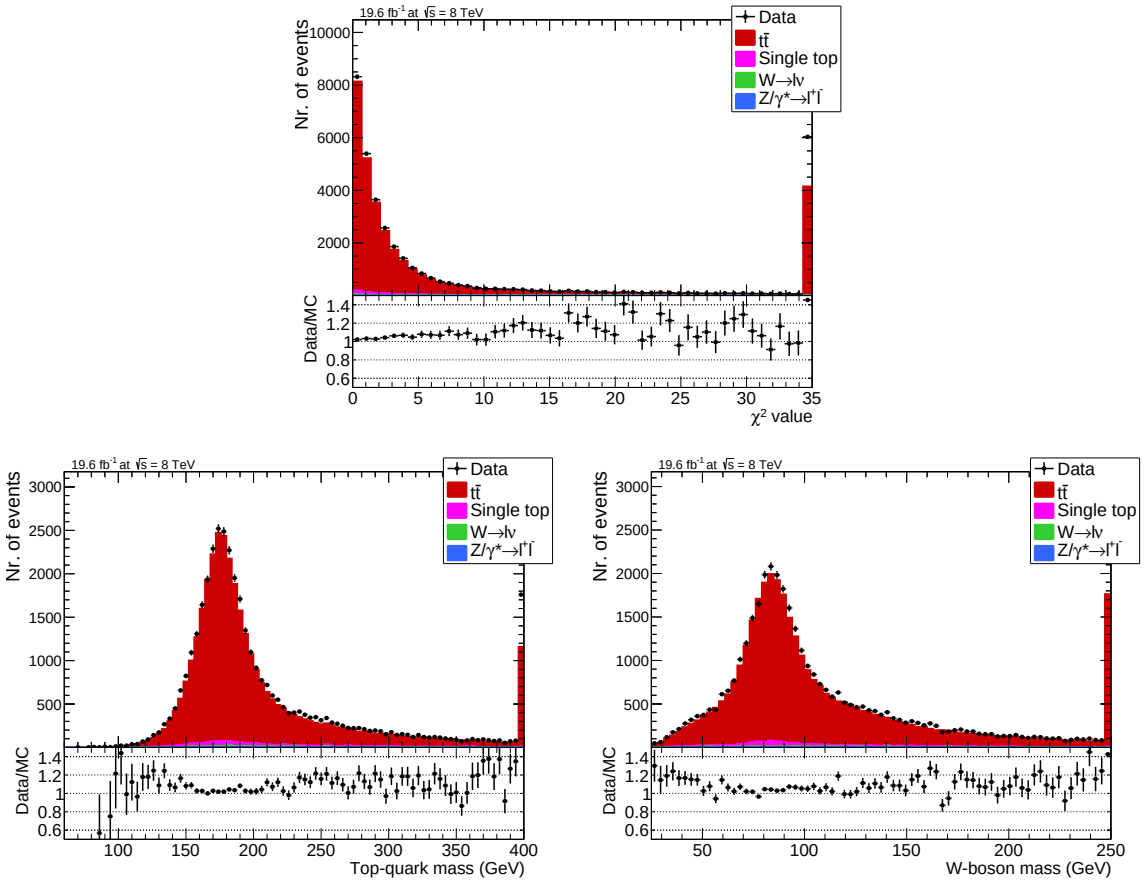

Figure 4.8: Distribution of the three variables used during the final step of the event-selection chain to ensure the correct event topology is reconstructed. The χ^2 of the chosen jet combination (upper plot) is required to be lower than 10 and the mass of the top-quark (bottom left) and W-boson (bottom right) have to be reconstructed within their respective 2σ-range.

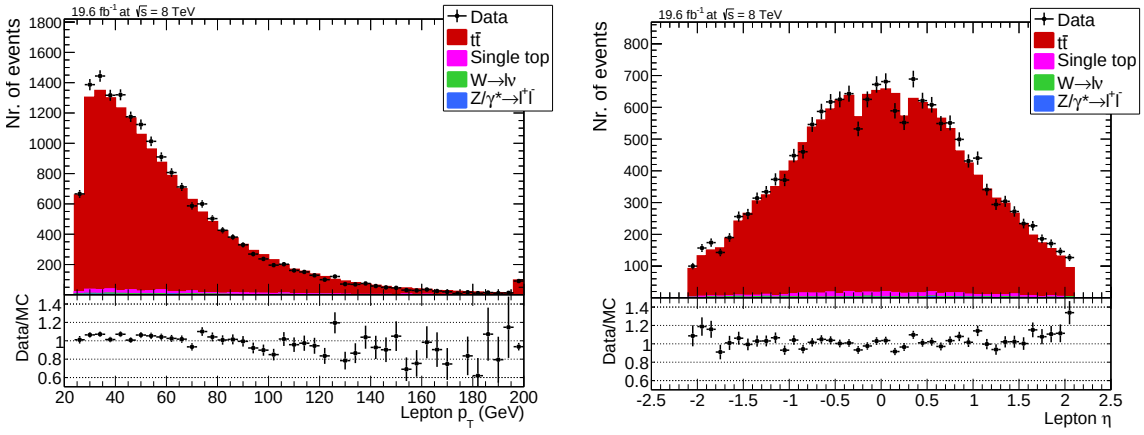

Figure 4.9: Comparison between data and simulation for the transverse momentum (left) and pseudo-rapidity (right) of the selected muon.

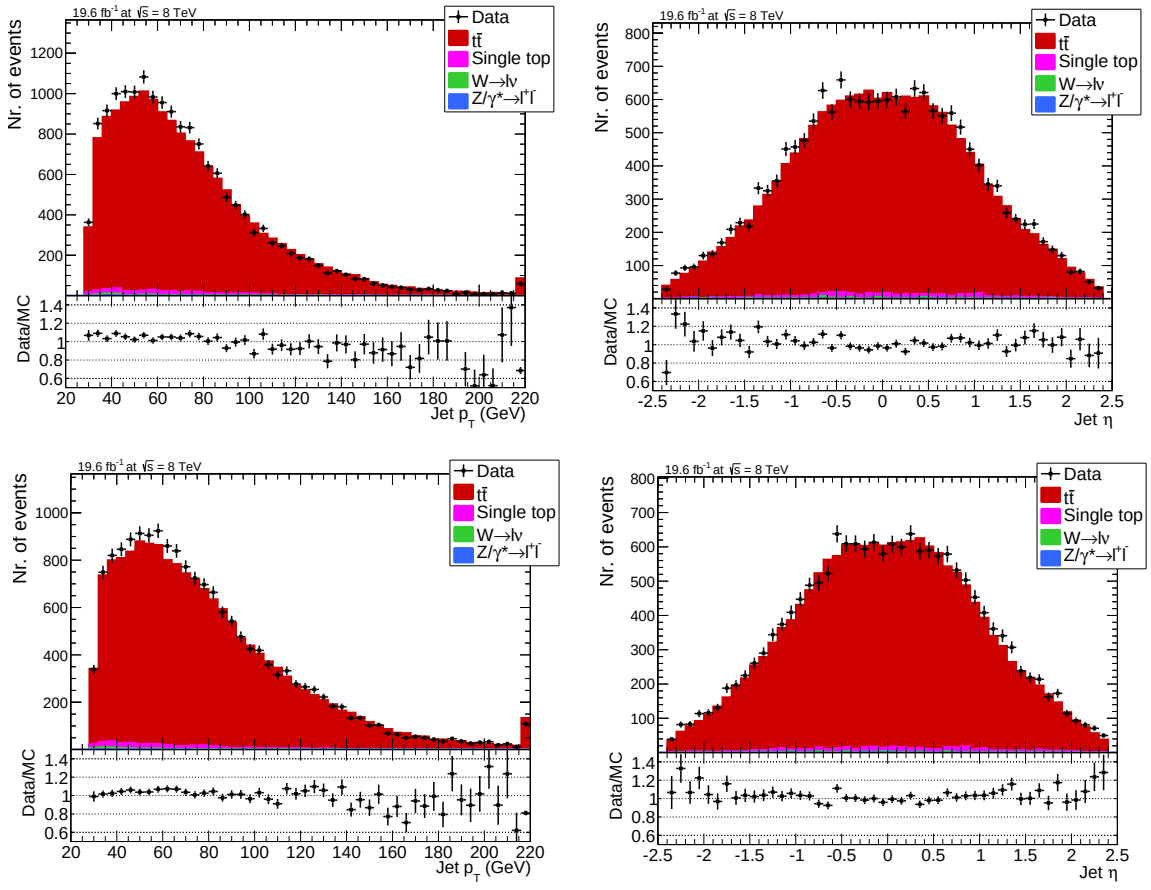

Figure 4.10: Comparison between data and simulation for the transverse momentum (left) and pseudo-rapidity (right) for the two reconstructed b-jets. The properties of the b-jet originating from the hadronic decaying top-quark are described in the upper distributions and the leptonic b-jet in the lower ones.

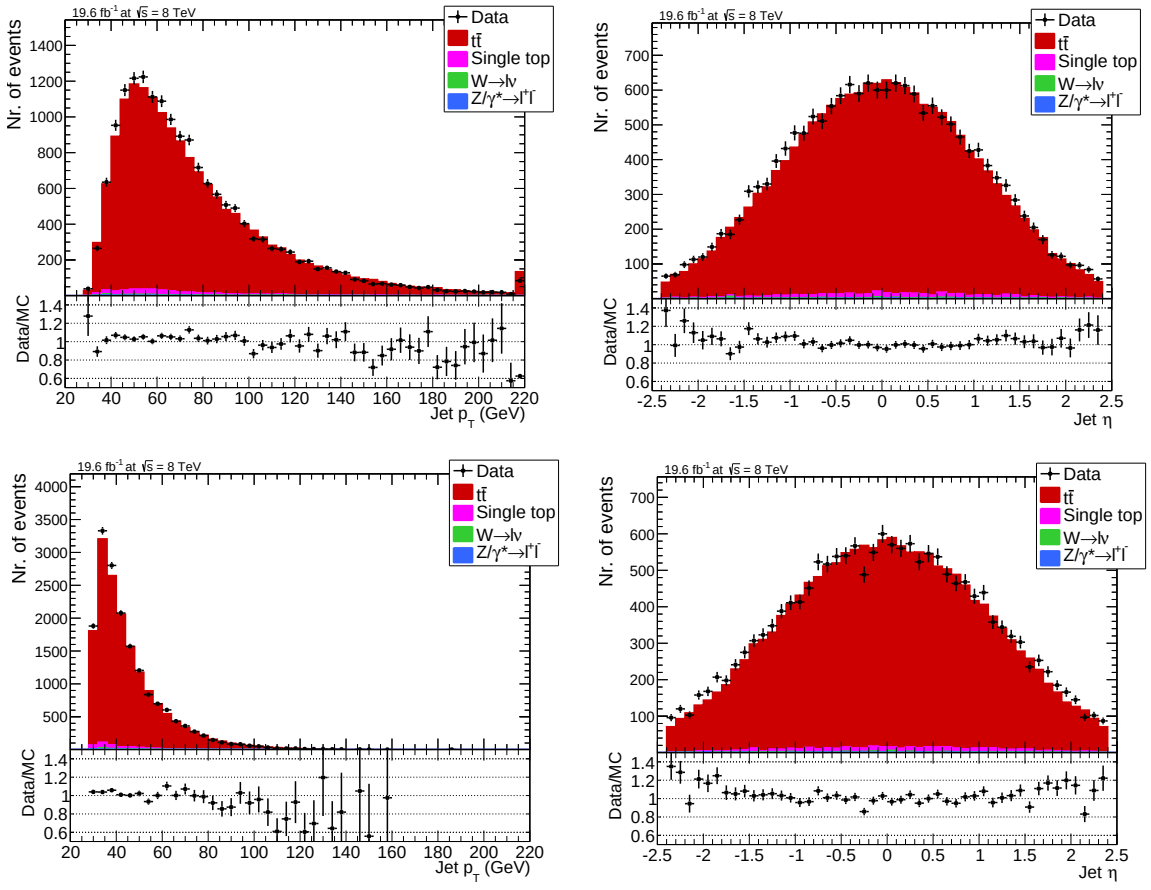

Figure 4.11: Comparison between data and simulation for the transverse momentum (left) and pseudo-rapidity (right) for the two reconstructed light-quark jets. These two jets cannot be differentiated and are therefore the only remaining permutation left in the considered jet combination.

Chapter 5

The Matrix Element method

The measurement of the right-handed tensor coupling of the Wtb interaction discussed in this thesis is performed using the Matrix Element method. This is an advanced analysis technique which allows to extract theoretical information from experimentally observed collision events without requiring any prior knowledge of the possible new-physics scenarios. The Matrix Element method assigns a probability to each theoretical hypothesis on an event-by-event basis, by calculating the matrix-element of the considered process. The obtained event probabilities are then combined into a likelihood and the most probable hypothesis is determined using a likelihood-maximisation method.

A detailed overview of the technicalities and applicability of the Matrix Element method will be given in this chapter. At first the procedure used to obtain the event probabilities from the corresponding matrix-element can be found in Section 5.1. Since in this analysis this matrix-element should incorporate the anomalous couplings of the Wtb interaction, a specific theoretical model has been developed as will be discussed in Section 5.2. To ensure a realistic description of the detection principles and the stochastic effects influencing the reconstruction process is obtained, dedicated resolution functions are introduced. These transfer functions, which smear the kinematic information of the parton-level events, have been determined using the selected signal events as will be discussed in Section 5.3. Since the Matrix Element method will consider various theoretical hypotheses for which the corresponding cross-section can significantly differ, the obtained event probabilities have to be normalised. A detailed explanation for this normalisation procedure can be found in Section 5.4.
The last part of this chapter, Section 5.5, will be devoted to the practical application of the method and will discuss how the Matrix Element estimator can be obtained from the overall likelihood. In order to demonstrate this concept, a measurement of the top-quark mass will be given as an example.

5.1 The Matrix Element method using MadWeight

The Matrix Element method has been developed several years ago to make maximal use of the kinematic information available in experimental events. Since this method is capable of analysing processes with a complicated final state, typically containing several jets and missing energy, it has been used extensively in the top-quark physics

sector at the Tevatron [102–104]. Given the challenging conditions the LHC is faced with, the use of the Matrix Element method has been revived recently and it has found applications in several physics areas [105–108]; including the search for the Standard-Model Brout-Englert-Higgs boson.
The fact that the Matrix Element method is capable of dealing in an efficient manner with event signatures involving missing energy together with the option it provides to determine the best estimate of any theoretical parameter from the available experimental events[1], has made it the appropriate analysis technique for the measurement of the g_R coefficient. Hence this right-handed tensor coefficient will be studied by using the Matrix Element method to calculate an individual event probability for several values of this g_R parameter. Applying a maximisation method on the overall likelihood then allows to determine the g_R value corresponding the best with the available experimental information.

The details of this statistical procedure can be found in Section 5.1.1. The actual calculation of the event probabilities will be performed by the automated MadWeight [109] integrator as will be discussed in Section 5.1.2. The main benefit of this procedure is that it is implemented in the MadGraph framework allowing for an easy access to the Feynman diagrams of various processes.

5.1.1 Likelihood definition and evaluation

The Matrix Element method analyses each event separately and determines an individual event probability, which corresponds to the agreement observed between the kinematic information provided by the experimental event and the theoretical information represented by the considered squared matrix-element. This event probability describes the convolution between the theoretical phase-space and the measured experimental phase-space. For hadron colliders such as the LHC, the interaction is initiated by partons such that the energy fraction of the proton carried by the partons are not well determined. Hence the parton distribution functions $f_i(q_i)$ need to be taken into account. The master formula reads:

$$P(x|\alpha) = \frac{1}{\sigma_\alpha * Acc_\alpha} \int d\Phi(y)\, dq_1\, dq_2\; f_1(q_1)\, f_2(q_2)\, |M_\alpha(y)|^2\, W(x,y) \qquad (5.1)$$

with $d\Phi(y)$ the phase-space measure, $M_\alpha(y)$ the tree-level matrix-element, $W(x,y)$ the transfer function, x the measured kinematic variables, y the variables of the partonic-level final state of the considered process and α the set of parameters to be measured. In the analysis discussed in this thesis, α will correspond to the right-handed tensor coupling g_R.

The resolution functions $W(x,y)$ are a crucial aspect of the Matrix Element method since these give the probability to reconstruct the variables of the partonic-level final

[1]The Matrix Element method also has a second widely used application, where it analyses two competing hypotheses and determines which one corresponds the most with the available experimental information.

state y as the measured kinematic variables x. Hence these will describe the evolution of the parton-level configuration into a reconstructed event using a realistic detector response, and will be determined in Section 5.3.
A second important aspect is the normalisation of the probability by the integrated cross-section σ_α and the detector acceptance Acc_α, which ensures the obtained event probability corresponds to a probability density[2]: $\int dx P(x|\alpha) = 1$. This normalisation is not performed automatically by the MadWeight integration and should therefore be included afterwards as will be discussed in Section 5.4.

The individual event probabilities are then combined into an overall likelihood $\mathcal{L}_{MEM}$ from which the best estimate of the considered theoretical parameter is obtained through a maximisation-method.

$$\mathcal{L}_{MEM}(x|\alpha) = \prod P(x|\alpha) \tag{5.2}$$

However, in practice it is more convenient to convert the likelihood values into χ^2 values using $\chi^2 = -2\ln\mathcal{L}$ since this allows to sum the full collection of event probabilities.

$$\chi^2(x|\alpha) = -2\ln\mathcal{L}_{MEM}(x|\alpha) = -2\sum \ln P(x|\alpha) \tag{5.3}$$

As a result this best estimate, or so-called Matrix Element estimator, will be determined using a minimisation method on the $\Delta\chi^2_{\text{MEM}} = \chi^2_{MEM} - \chi^2_{MEM,min}$, where this $\chi^2_{MEM,min}$ variable corresponds to the lowest overall χ^2_{MEM} value.

5.1.2 MadWeight

The Matrix Element method makes maximal use of the kinematic information available in experimental events and should therefore ideally result in one of the more powerful tools to extract theoretical information from a sample of experimental events. Unfortunately calculating the individual event probabilities is a rather complicated procedure, especially since the convolution between the hard-scattering process and the kinematic smearing results in a challenging multi-dimensional integration procedure, which seriously hampers the applicability of the Matrix Element method.

Nevertheless due to the challenging conditions at the LHC and the various interesting event signatures characterised by missing energy, the Matrix Element method has an extensive number of possible applications. Hence a dedicated algorithm has been developed which evaluates the event probabilities in a fully automated manner and ensures an optimised phase-space mapping for a more efficient integration. This highly-flexible phase-space integrator, which uses the adaptive Monte Carlo integrator VEGAS [110], has been denoted MadWeight [109].
This MadWeight integrator, fully implemented in the MadGraph framework, significantly facilitates the use of the Matrix Element method. Prior to the existence of this tool, a separate integration procedure had to be developed for each considered final-state signature and detailed knowledge on the technical details of both matrix-element

[2] Assumming both a flat prior for α and a normalised transfer function.

generation and phase-space integration was required in order to apply the Matrix Element method.
However, even with this optimised integration procedure the Matrix Element method remains a very time-consuming analysis technique. Therefore it should be avoided to calculate the probabilities of events for which the expected final state particles are not, or only partially, recovered. Hence the choice to introduce an stringent event selection as discussed in Chapter 4. In addition the computing time has also been reduced by distinguishing the two b-quark jets present in top-quark pair events. As a result, for this analysis, where the semi-muonic decay of the $t\bar{t}$ events is studied, only the permutation between the two light jets has to be performed during the integration procedure.

Due to the complexity of the integration procedure, for some events this integration can fail implying that the MadWeight integrator is not capable of providing an event probability. Even if this occurs for only one of the considered theoretical hypotheses, the entire event should nevertheless be excluded in order to avoid introducing a bias on the overall likelihood. This might become relevant for the measurement of the right-handed tensor coupling, which will be discussed in the Chapter 6, since for this study the full available collection of simulation events will be considered. Fortunately this effect only occurs very rarely, less than 0.7% of the events of one specific sample of background events has been affected such that this feature is not expected to influence the overall result.

5.2 Implementing the anomalous Wtb Lagrangian

Since the implementation of the MadWeight integration procedure in the MadGraph framework, the option to analyse personally developed models describing new-physics phenomena has been significantly facilitated. Hence, any model can be created with FeynRules [111], which is a Mathematica-based package to calculate Feynman rules, and be translated to MadGraph using the dedicated interface. The developed model should only contain some basic information, such as the particle content, the parameters and the Lagrangian, which allows the FeynRules package to derive the Feynman rules.
In this analysis a new model has been created specifically for the study of the anomalous couplings in the Wtb interaction vertex using this FeynRules package. This model has been defined as an addition to the Standard Model, hence the entire particle content and parameters of the Standard Model have been kept. The description of the anomalous couplings is then included by introducing four new complex parameters, the four coupling coefficients, and the full Lagrangian described in Equation (1.20). Since the parameter of interest for this analysis, the g_R coefficient of the Wtb interaction vertex, is associated with the decay of the top-quark and is thus not foreseen to change the final state particles, the particle content has not been altered.
Within the model some simplifications have been introduced since the complexity of the developed model is directly related to the processing time required by the MadWeight integration procedure. Hence, the light-flavoured quarks (u-, d-, c- and s-quark) are assumed to be massless and, just as is the case for the Standard Model decays, the

CKM-suppressed W-boson decays have been suppressed.

Besides calculating the Feynman rules for the developed theoretical model, the FeynRules package also ensures the introduced Lagrangian fulfils the basic set of requirements, such as hermicity, gauge invariance, *etc.*. The downside of developing such a brand new model is that there is no straightforward way to ensure the model perfectly describes the introduced parameter since no prior knowledge is available. For the top-quark decay vertex some influences of the coupling coefficients can be visualised by looking at the distortions of the angular distribution of the top-quark decay products, as has been mentioned in Section 1.2.2. Hence, a thorough comparison has been performed with the distributions obtained when simulating events with different values of the coupling coefficients using the developed model. The obtained outcome, shown in Figure 5.1, corresponds well with the theoretical expectations, given previously in Figure 1.2. In addition, the model has also been used to calculate the most optimal value of some of the well-known parameters of the Standard Model, for which no unexpected deviations have been observed.

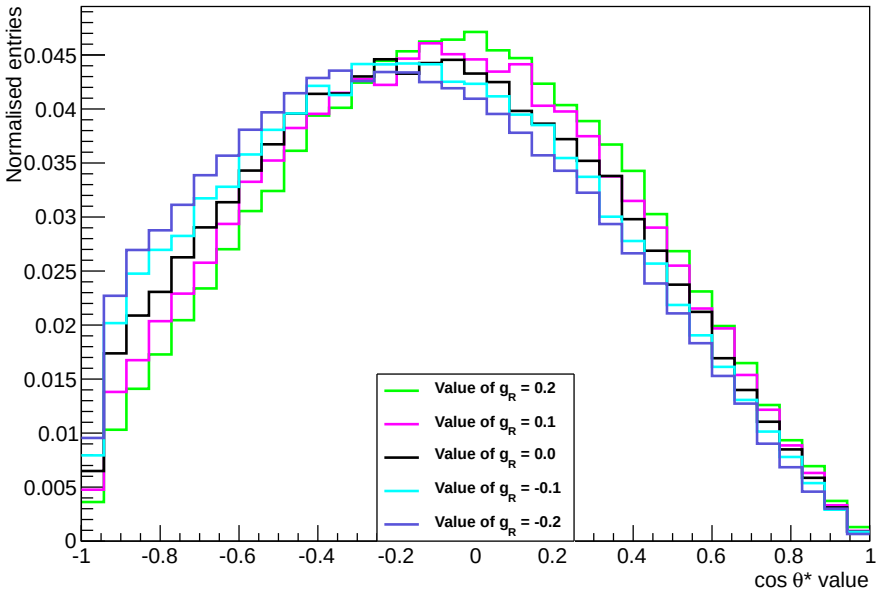

Figure 5.1: Distribution of the $\cos\theta^*$ variable obtained from the model specifically developed to perform the measurement of the right-handed tensor coupling in the Wtb interaction vertex. All other coupling coefficients are equal to their Standard Model prediction ($V_L = 1$ and $V_R = g_L = 0$).

5.3 Resolution functions

As was already briefly mentioned in Section 5.1, the Matrix Element method requires carefully calculated resolution function in order to link the three momenta of these final-state partons with the corresponding momenta of the reconstructed physics objects. These resolution functions are an important aspect of the Matrix Element method and

take into account the combined effect of parton showering, hadronisation and detector response.
However, in order to be able to determine these resolution functions, some strong simplifications have to be imposed. Since a resolution function should exist for each particle type, describing both its direction (ϕ and θ) and energy, it is assumed that these functions are uncorrelated. This allows to reformulate the resolution function $W(x, y)$ in Equation (5.1) using a factorised approach:

$$W(x,y) = \prod_{i}^{N} W(x_i, y_i) = \prod_{i}^{N} W_i^E(x^i, y^i)\, W_i^\eta(x^i, y^i)\, W_i^\phi(x^i, y^i) \tag{5.4}$$

where the index i runs over the different types of physics objects in the considered event topology.

This factorised description can be simplified even further since the transfer functions for the object directions can be represented with a Dirac-δ function. This because both angles are determined very precisely and therefore correspond well with the measured objects, as can be seen from Figure 5.2. The differences shown in these distributions are determined using the parton-level and reconstructed physics object to have an angular distance ΔR smaller than 0.3, a similar condition as what is applied in Chapter 4. As a result, the only remaining phase-space variable for which a transfer function should be determined is the energy, for which the Dirac-δ assumption would not be valid due to the finite precision on its measurement. The transfer function of the energy variable will therefore be represented using a Gaussian-like function.

In this analysis, which focuses on the semi-muonic decay of top-quark pairs, a dedicated transfer function should be developed for the jets and the muon in the event. However due to the possible different behaviour of the light- and heavy-flavoured jets, it has been opted to determine two separate transfer functions for the jets. Hence the two jets identified as originating from the decay of the W-boson and the two jets assigned to the b-quark decay will be treated independently. The transfer function for the muons on the other hand will also be represented with a Dirac-δ function since its energy is determined very precisely. The different behaviour of the muon and the jets, the light-flavoured ones in this case, can be seen in Figure 5.3.
Besides simplifying the transfer-function calculation, this approach of using Dirac-δ functions for various phase-space variables and particle types also significantly speeds up the Matrix Element method. Forcing the reconstructed objects to perfectly correspond to those of the parton-level objects implies that the method does not have to integrate over the corresponding phase-space variables.

The transfer function of both the light- and heavy-flavoured jets will be described using a double-Gaussian distribution, for which the formula is given in Equation (5.5). This double-Gaussian representation is an adequate choice to describe the energy difference ($\Delta E = E_{parton}$ - E_{jet}) of jets, which is characterised by a sharp peak and an asymmetric tail. From Figure 5.3 can furthermore be concluded that the width of the overall ΔE distribution increases for higher energies of the matched parton. Hence an

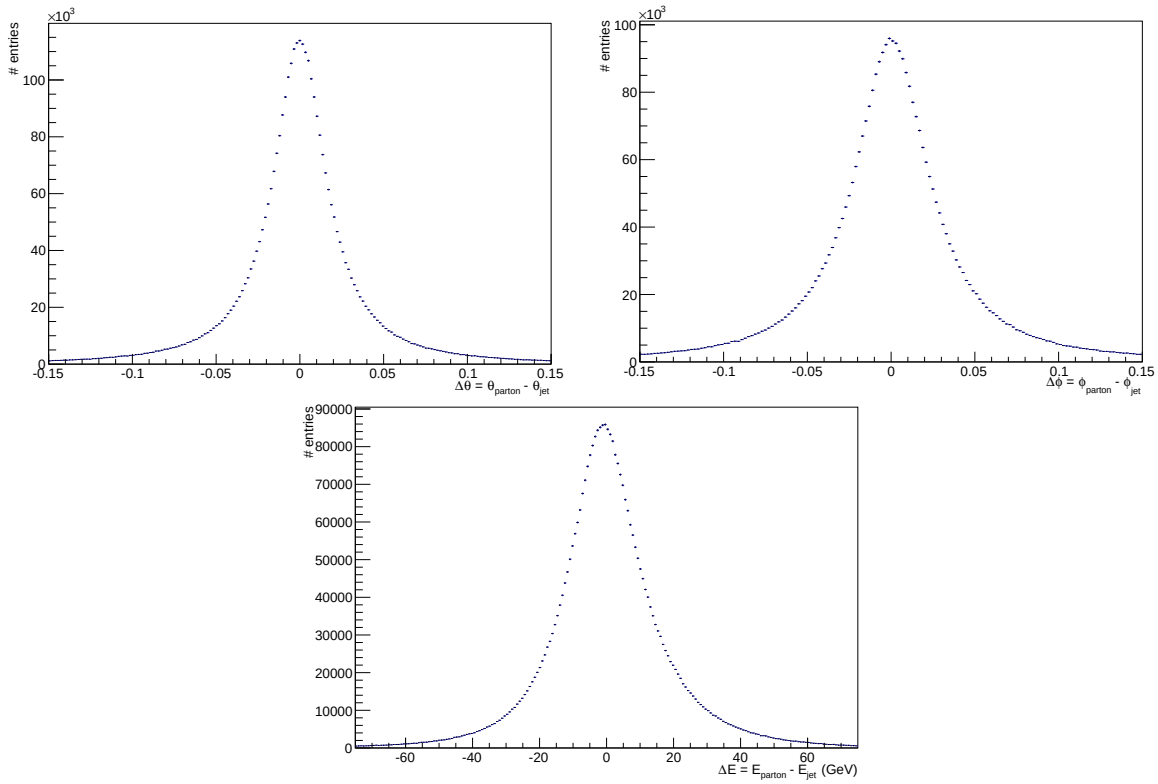

Figure 5.2: Distribution of the difference for θ (left), ϕ (middle) and E (right) obtained for jets, which indicates that the two object directions are more accurately measured.

accurate description of both the peak and the tail is necessary in order to ensure the transfer function remains valid for a wide energy regime.

$$W^E(parton, jet) = \frac{1}{\sqrt{2\pi}} \frac{1}{a_2 + a_3 a_5} \left(\exp \frac{-(\Delta E - a_1)^2}{2{a_2}^2} + a_3 \exp \frac{-(\Delta E - a_4)^2}{2{a_5}^2} \right) \quad (5.5)$$

where the parameters a_1 and a_4 represent the mean of the first and second Gaussian, respectively, while the parameters a_2 and a_5 corresponds to the width of those distributions. The remaining parameter a_3 takes into account the relative contribution of both distributions.

The actual determination of the two remaining transfer functions will be performed by applying a double-Gaussian fit on the obtained ΔE distribution for various E_{parton} values. For the light jets 16 bins are considered between 25 GeV and 160 GeV, while for the b-quark jets 18 bins are used between 30 GeV and 230 GeV. In order to ensure sufficient statistics is available throughout the entire energy-range, only the basic event-selection requirements have been applied. Hence the additional event-selection criteria discussed in Section 4.2 are not taken into account when determining the transfer functions.
The two-dimensional histogram showing the E_{parton} distribution with respect to the ΔE distribution for both the light- and heavy-flavoured jets can be found in Figure 5.4, respectively the left and right plot. Comparing the two distributions allows to conclude

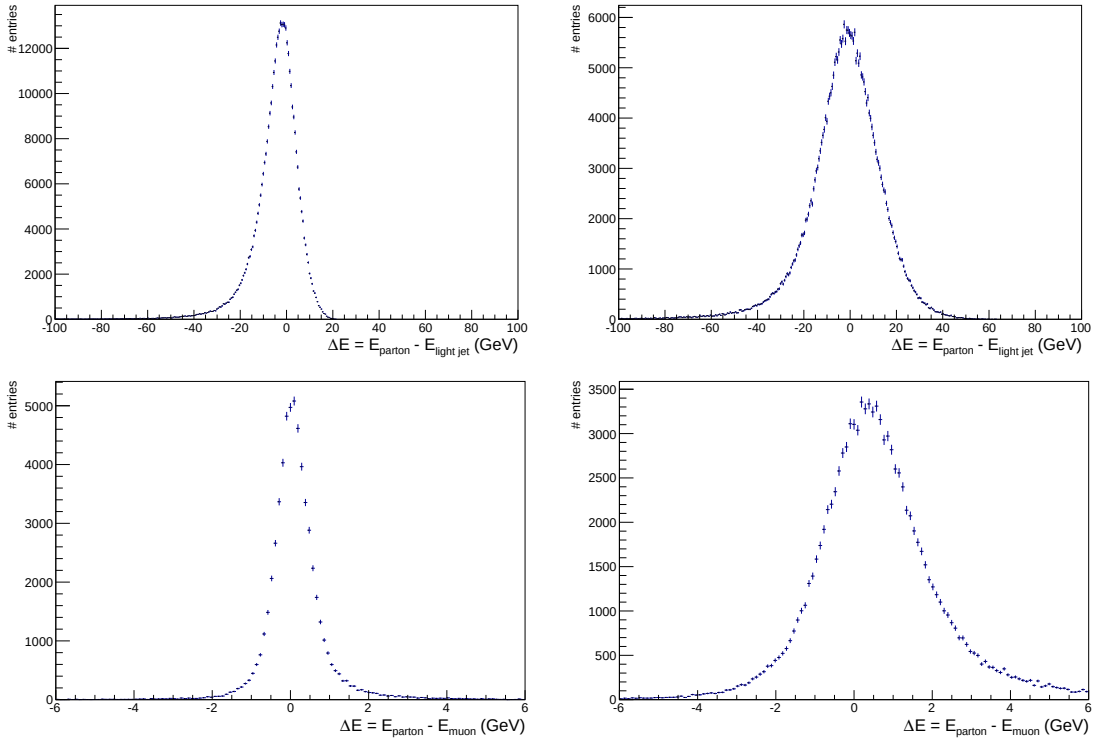

Figure 5.3: Obtained ΔE distributions for the light-flavoured jets (top) and the muons (bottom) for a randomly selected lower (left) and higher (right) E_{parton} value. The muons are clearly determined more accurately throughout the entire energy range.

that it is indeed beneficial to treat both types of jets independently since a wider ΔE distribution is clearly visible for the b-quark jets. The actual determination of the transfer functions will be done by fitting these histograms, using the specific range for each type as mentioned above, with a double Gaussian.

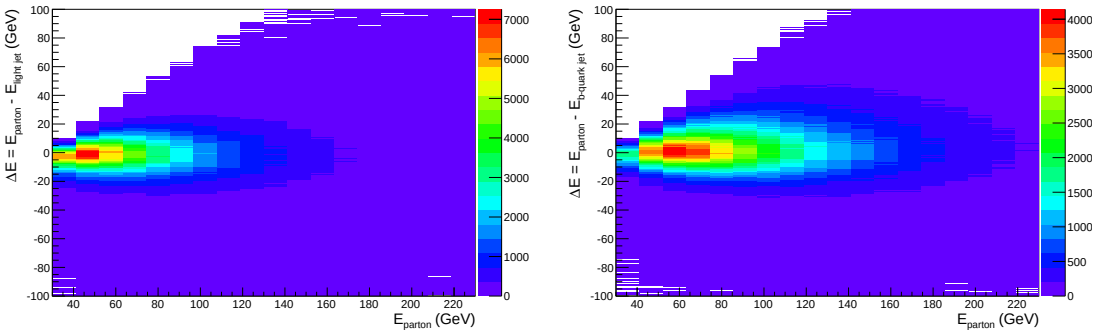

Figure 5.4: Two-dimensional histogram showing the parton energy E_{parton} with respect to the difference in energy with the matched jet, ΔE for the light jets (left) and the b-quark jets (right) in case the same range would be used for both.

For each of the considered energy ranges of the parton-level quarks or gluons, a measurement of the five E-dependent parameters of the double-Gaussian transfer function is obtained. The E-dependency of these transfer-function parameters is based on

the parametrisation of the calorimeter energy resolution, which corresponds to $y = a + b\sqrt{E} + cE$. However, in order to ensure the parameters are well described by this parameterisation, a quadratic term, or for some parameters even a cubic term, has been added. An overview of the imposed E-dependency for the different a_i parameters can be found in Table 5.1.

Table 5.1: Imposed E-dependency of the different transfer-function parameters. Only in three cases the additional cubic function was necessary to obtain a good description of the corresponding parameter.

Light-jet parameters	b-quark jet parameters
$x_{1,0} + x_{1,1}\sqrt{E} + x_{1,2}E + x_{1,3}E^2 + x_{1,4}E^3$	$x_{1,0} + x_{1,1}\sqrt{E} + x_{1,2}E + x_{1,3}E^2 + x_{1,4}E^3$
$x_{2,0} + x_{2,1}\sqrt{E} + x_{2,2}E + x_{2,3}E^2$	$x_{2,0} + x_{2,1}\sqrt{E} + x_{2,2}E + x_{2,3}E^2$
$x_{3,0} + x_{3,1}\sqrt{E} + x_{3,2}E + x_{3,3}E^2$	$x_{3,0} + x_{3,1}\sqrt{E} + x_{3,2}E + x_{3,3}E^2 + x_{3,4}E^3$
$x_{4,0} + x_{4,1}\sqrt{E} + x_{4,2}E + x_{4,3}E^2$	$x_{4,0} + x_{4,1}\sqrt{E} + x_{4,2}E + x_{4,3}E^2$
$x_{5,0} + x_{5,1}\sqrt{E} + x_{5,2}E + x_{5,3}E^2$	$x_{5,0} + x_{5,1}\sqrt{E} + x_{5,2}E + x_{5,3}E^2$

Figure 5.5 contains some examples of the ΔE distribution obtained for two of the considered parton-level energy bins. The two upper plots correspond to the light-flavoured jets while the two lower ones depict the situation for the heavy-flavoured ones. In order to demonstrate the benefit of using a double-Gaussian description for the transfer functions of the jet energies, the left distribution gives the distribution for some relatively low parton energies while the right corresponds to one of the more outer bins of the considered energy range. For the wider ones, this double-Gaussian description allows to nicely describe the shape of the tail. As can be seen, the range where the double-Gaussian fit is applied has been optimised for each of the considered bins of E_{parton}.

Finally, the obtained E-dependent shape of the five parameters describing the double-Gaussian transfer function is given in Figure 5.6 and 5.7 for the light and b-quark jets, respectively. In case the considered ΔE distribution is not well described due to a lack of statistics, it is combined with one of the surrounding ones in order to ensure sufficient information is available to perform the double-Gaussian fit. In each parameter overview the two upper parameters correspond to the narrow Gaussian while the two middle ones are related to the broad Gaussian, for which the mean is consistently given on the left and the width on the right. The bottom parameter depicts the relative contribution of the two Gaussian distributions. The blue fitted curve uses the parametrisation given in Table 5.1.
From the obtained shapes for these parameters can be concluded that the light-flavoured jets are in general better described by the imposed parametrisations while for the b-quark jets the interplay between the narrow and wide Gaussian is more challenging

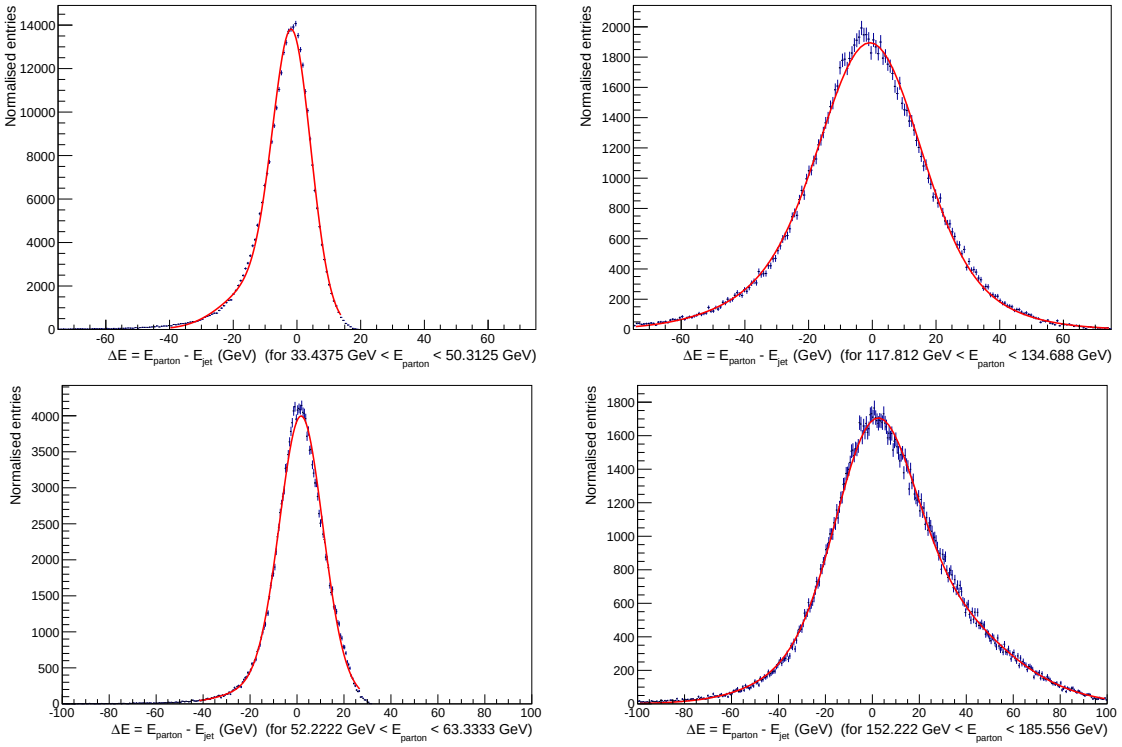

Figure 5.5: Distribution of the difference in energy between the parton-level and reconstructed object for both the light jets (upper two) and the b-quark jets (lower two) fitted with a double-Gaussian function. For both the light-flavoured and b-flavoured jets a random E_{parton} bin with lower (left) and higher energy (right) has been chosen to represent the need of a double-Gaussian function which is capable of both describing the peak and the tail.

to correctly represent. Nevertheless, it can be observed that for both types of jets the narrow Gaussian is the dominating one at low E_{parton} energies while the broad Gaussian takes over at higher values. For the b-quark jets, depicted in Figure 5.7, this broad Gaussian is more important than for light jets, Figure 5.6, since the energy of the b-quark jets is more difficult to measure.

5.4 Normalisation of the Matrix Element likelihood

A second important aspect of the Matrix Element method is the normalisation of the event probability using the cross-section and acceptance, which might vary significantly for the different values of the considered model parameter. This normalisation should incorporate both the different cross-section and event-selection acceptance conditions, which are likely to be dependent on the value of the theoretical parameter that is studied. In order to avoid significantly influencing the outcome of the Matrix Element estimator, the cross-section and acceptance should be determined with great care.
For generator-level events these values can be determined in a straightforward manner using the MadGraph generator. For the reconstructed events on the other hand,

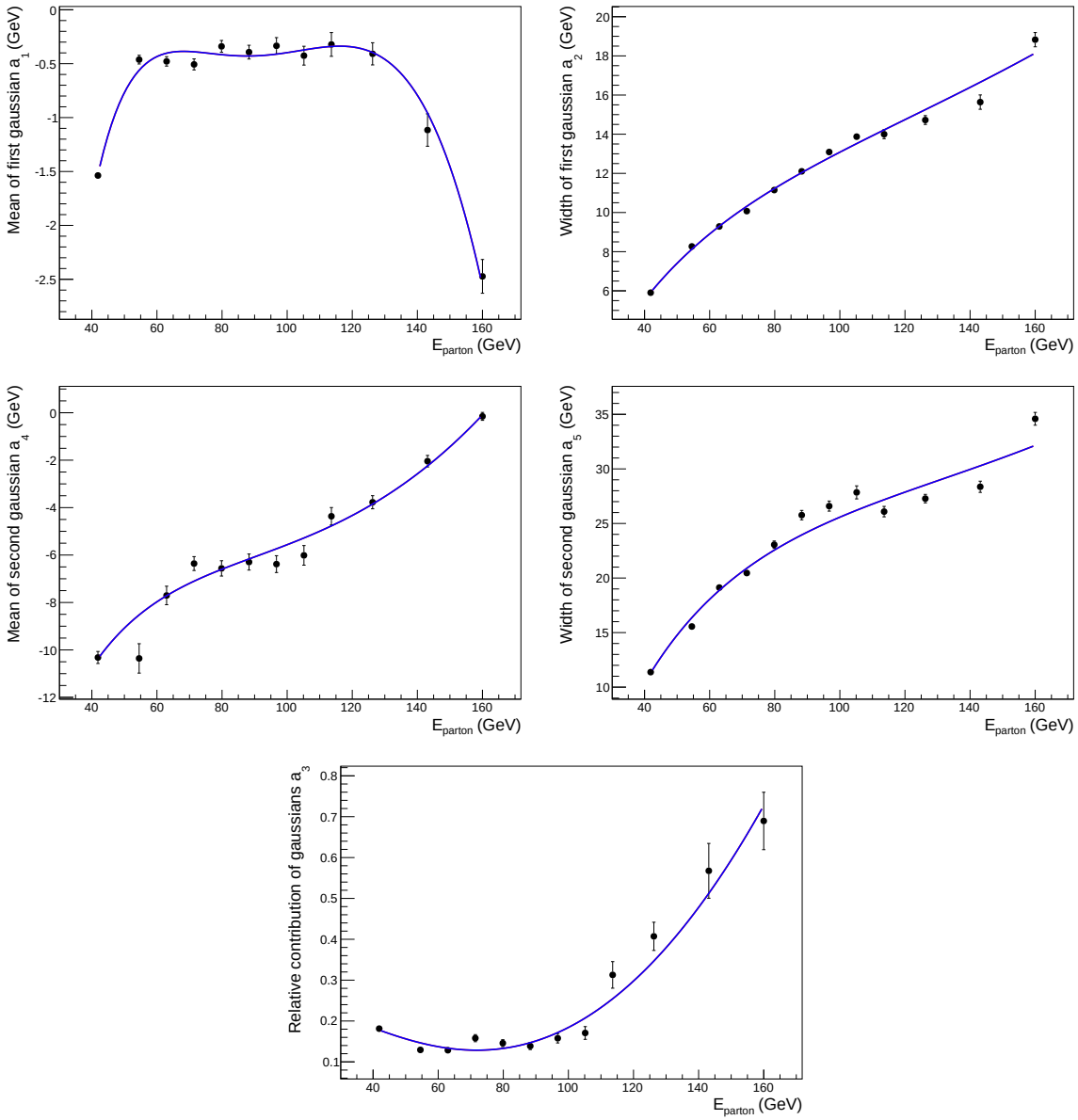

Figure 5.6: Obtained shape for the five E-dependent parameters describing the double-Gaussian transfer function for the light-flavoured jets. The two upper parameters correspond to the narrow Gaussian while the two middle ones are related to the broad Gaussian, for which the mean is consistently given on the left and the width on the right. The bottom parameter depicts the relative contribution of the two Gaussian distributions. Each parameter has been fitted with the corresponding parametrisation given in Table 5.1.

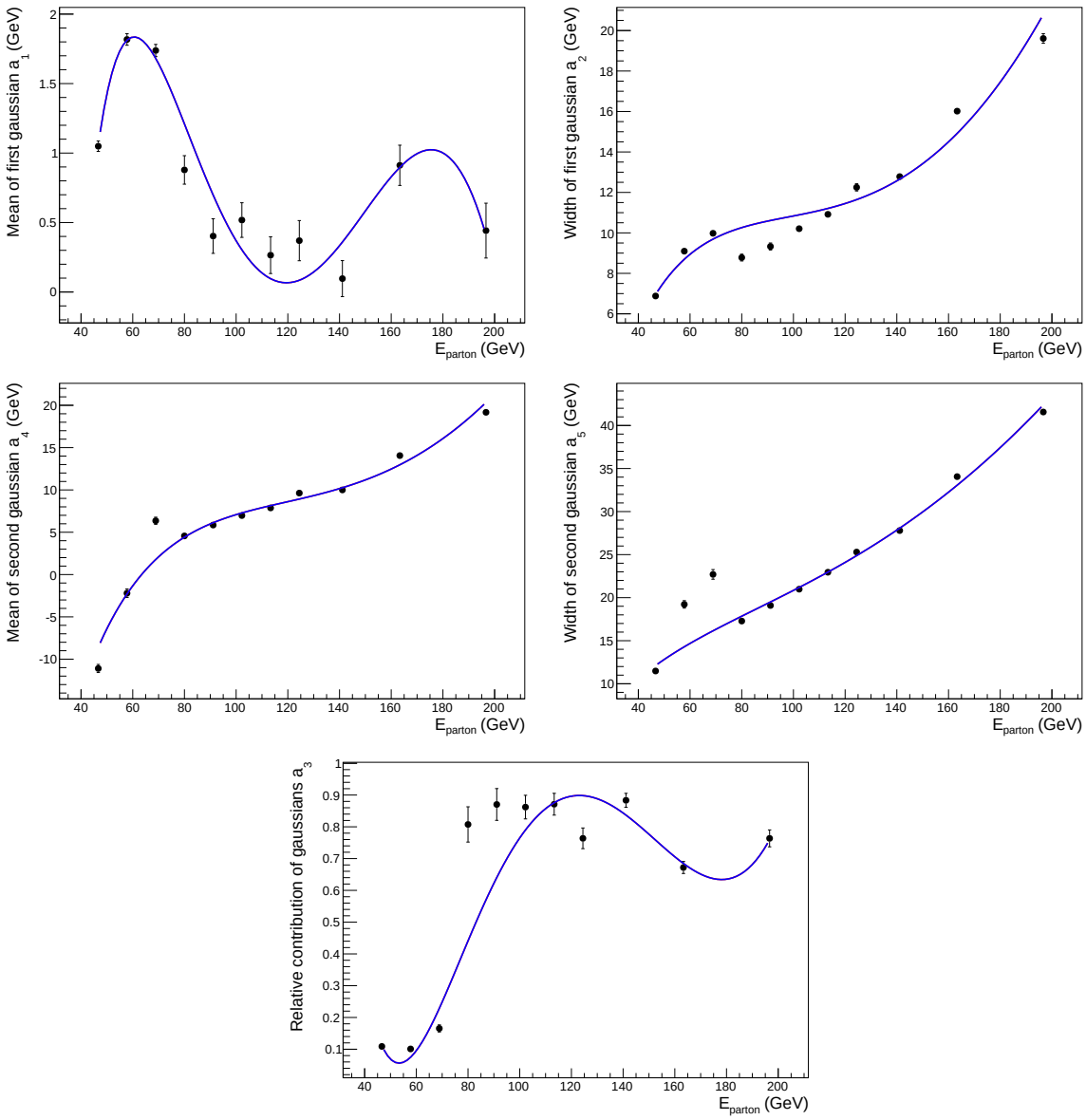

Figure 5.7: Obtained shape for the five E-dependent parameters describing the double-Gaussian jet transfer function for the b-flavoured jets. The two upper parameters correspond to the narrow Gaussian while the two middle ones are related to the broad Gaussian, for which the mean is consistently given on the left and the width on the right. The bottom parameter depicts the relative contribution of the two Gaussian distributions. Each parameter has been fitted with the corresponding parametrisation given in Table 5.1.

this normalisation procedure is much more challenging since generating samples with different values of the theoretical parameter of interest is rather difficult and time-consuming. Hence it has been opted for in this thesis to derive the cross-section values for the reconstructed events from the generator-level ones.
In order to ensure that the generator-level cross-sections can easily be related to the reconstructed ones, the conditions present for the reconstructed collision events will be mimicked as closely as possible during the generation process. Hence the generator events have to fulfil the basic event selection requirements[3] listed in Table 5.2. The remaining event-selection criteria can not be applied during the generation process, but are on the other hand not expected to be as sensitive to the value of the coupling coefficient. By applying these constraints on the kinematics of the generated events a simplified, but nevertheless sufficiently accurate, acceptance determination will be obtained.

Table 5.2: Basic event selection applied to the generator-level events in order to partially mimic the situation existing for the reconstructed collision events. The mentioned ΔR distance corresponds to the minimum distance of the considered generator-level parton and all the other partons in the event topology.

Parton	p_{T} value	$\lvert\eta\rvert$ value	ΔR distance
Quark and gluon	> 30 GeV	< 2.5	> 0.3
Muon	> 26 GeV	< 2.1	> 0.3
Neutrino	> 25 GeV	< 2.5	> 0.3

Besides applying a significant fraction of the event-selection criteria, the processes considered for the generation process have also been selected in order to obtain an event signature comparable to data. Hence the cross-section values have been determined using a combination of top-quark pair decay processes surrounded by additional jets. The actual number of considered processes has been limited to the $t\bar{t}$ decay with none, one and two additional jets since the contribution of the processes with more additional jets is negligible.

Even with these two optimisations applied, the cross-section normalisation determination using generator-level events will not result in exactly the same outcome as when reconstructed events are used. Since it is simply not possible to include every aspect of the full event-selection chain in exactly the same way when generating the different processes, the obtained cross-section values need to be scaled in order to incorporate the effect of these non-included event-selection criteria. For this scaling an identical behaviour throughout the entire g_R range is assumed, and each cross-section value is multiplied with the factor $\sigma_{SM}^{reco}/\sigma_{SM}^{gen}$.

[3]Important to note here is that once these selection criteria are applied to the generated events, the obtained cross-section will actually be a combination of the cross-section of the underlying physics process and the acceptance of the considered event selection. Hence the term "cross-section normalisation" will implicitly imply the combined normalisation.

The final result of the cross-section calculation can be found in Figure 5.8, where the red curve represents the cross-section values obtained for the generator-level events using the approach discussed above. The cross-section values for the selected events are given by the blue curve and have been obtained by multiplying each of the generator-level cross-section values with the above-mentioned scaling factor $\sigma_{SM}^{reco}/\sigma_{SM}^{gen}$.

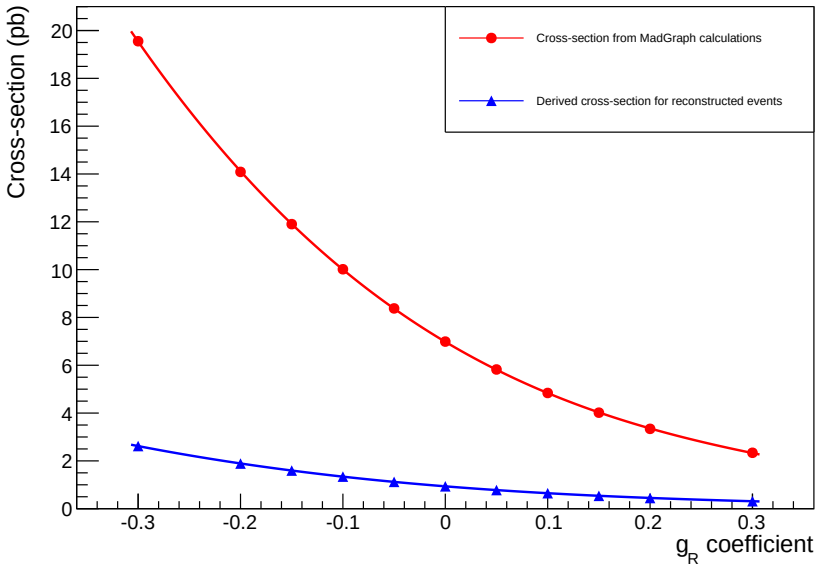

Figure 5.8: Overview of the obtained generator-level cross-sections for different g_R values and the reconstructed ones derived from them by applying the ratio $\sigma_{SM}^{reco}/\sigma_{SM}^{gen}$.

5.5 The Matrix Element estimator

Once all the different aspects of the Matrix Element method have been clearly established, such as the theoretical model that will be considered, the resolution functions that will be considered and the cross-section normalisation that should be applied, this technique can finally be applied onto experimental events. The integration procedure, performed by MadWeight in this thesis, then calculates for every individual event an event probability which represents how well the available kinematic information corresponds with the theoretical assumption. Through a likelihood maximisation, or equivalently a $\Delta\chi^2_{\text{MEM}}$-minimisation, the best estimate of the considered theoretical parameter is obtained, also denoted as the Matrix Element estimator.

In order to perform the MadWeight integration procedure, the kinematic information of the reconstructed events has to be provided in a predefined format. For each particle in the event topology, including the missing energy representing the neutrino, the transverse momentum, the pseudo-rapidity, the azimuthal angle and the mass should be given. In addition, the values of the theoretical parameter that need to be considered by the integration procedure have to be specified and an event probability will be determined for each of these values.

The practical application of the Matrix Element method will be demonstrated by measuring the top-quark mass using this advanced analysis technique. The Matrix Element method has been applied both for generator-level events (generated $m_{top} = 173\,\text{GeV}$) and for centrally produced reconstructed $t\bar{t}$ events fulfilling the full list of event-selection criteria discussed before. These samples have been generated with a top-quark mass of, respectively 173 GeV and 172.5 GeV.

Since this measurement serves merely as an illustrative example, the study of the reconstructed events will be restricted to 4000 $t\bar{t}$ events for which each jet in the reconstructed event topology has been correctly matched with the corresponding parton. The resolution functions applied for the generator-level events have been significantly simplified by restricting all of them to a Dirac-δ function while for the reconstructed events the ones discussed in Section 5.3 will be applied. Six different values of the top-quark mass have been scanned over between 170 GeV and 175 GeV.
Before the actual measurement of the top-quark mass can be performed, the cross-section values for the reconstructed events have to be determined following the same procedure as explained in Section 5.4. The obtained results is shown in Figure 5.9 and clearly indicates that the top-quark mass does not depend heavily on the cross-section.

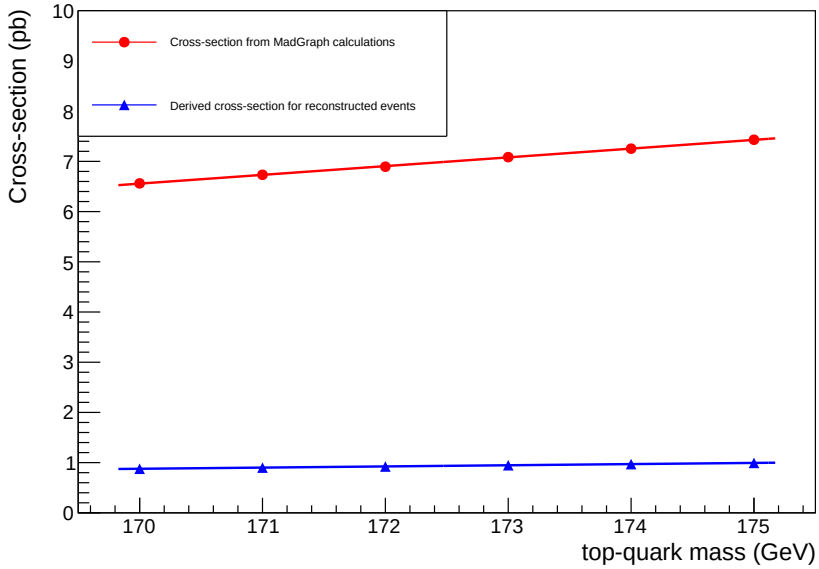

Figure 5.9: Overview of the obtained generator-level cross-sections for different top-quark mass values and the reconstructed ones derived from them by applying the ratio $\sigma_{SM}^{reco}/\sigma_{SM}^{gen}$.

The strength of the Matrix Element method finds its origin in the fact that it analyses each event individually, assigns a probability to correspond with the presumed hypothesis and then combines this information into one overall likelihood. Hence events for which the reconstructed event topology and kinematic information corresponds well with the considered process will therefore contain the most relevant information and are supposed to contribute on average the most to the overall result.

The difficulty of extracting information on an event-by-event basis can be visualised in Figure 5.10 where four event probabilities obtained for the generator-level study are shown for the considered top-quark masses. The observed shapes vary significantly for the considered events and the overall likelihood combining all these individual event likelihood is only relevant in case sufficient events are considered.

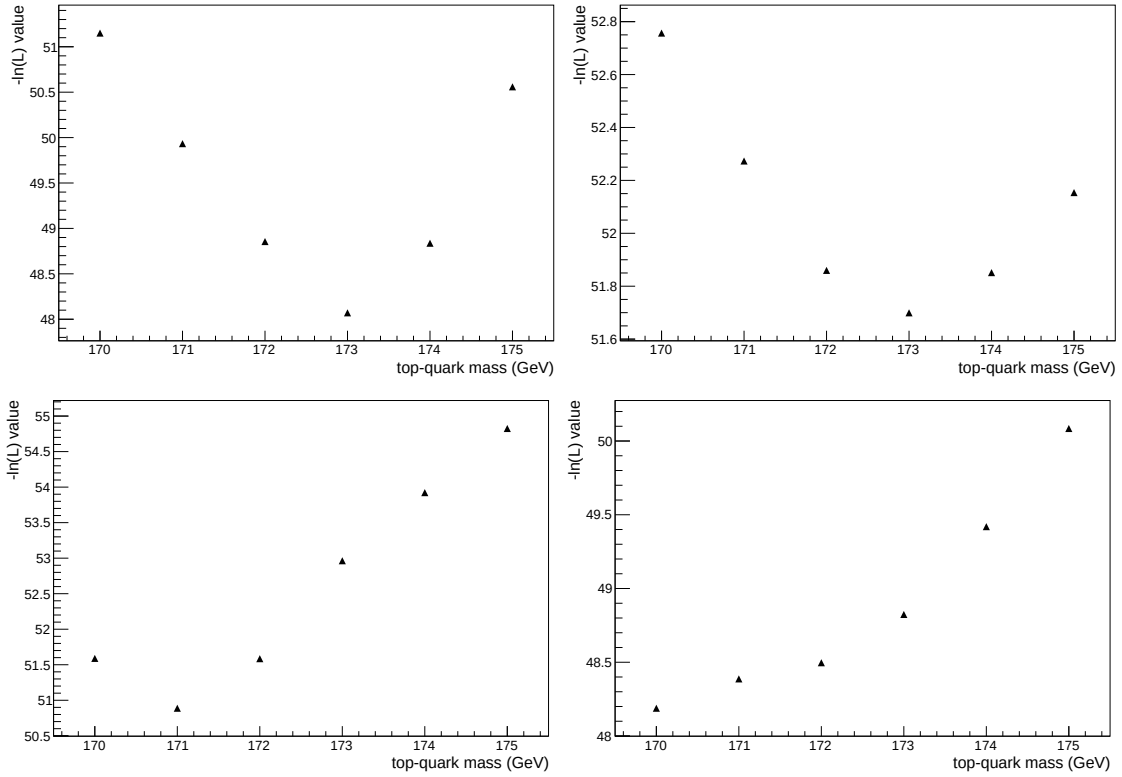

Figure 5.10: Individual event probabilities for the measurement of the top-quark mass using generator-level events.

The most optimal value of the considered theoretical parameter is obtained by performing a minimisation method on the negative logarithmic likelihood values of the full collection of experimental events. This Matrix Element estimator is obtained by fitting the $\Delta\chi^2_{\mathrm{MEM}}$ values with a quadratic function on a predefined range and the obtained minimum value then corresponds to the best estimate of the theoretical parameter.
For the top-quark measurement using generator-level events, for which the range of this quadratic fit has been restricted between 171 GeV and 175 GeV, the obtained best estimate corresponds to a value of 172.97 ± 0.02 GeV. This fit function together with the overall $\Delta\chi^2_{\mathrm{MEM}}$ values is given in Figure 5.11. Since these events have been generated with MadGraph by assuming a top-quark mass of 173 GeV this procedure results in a very nice agreement.

With the procedure applied for generator-level events, this proof of concept now contin-

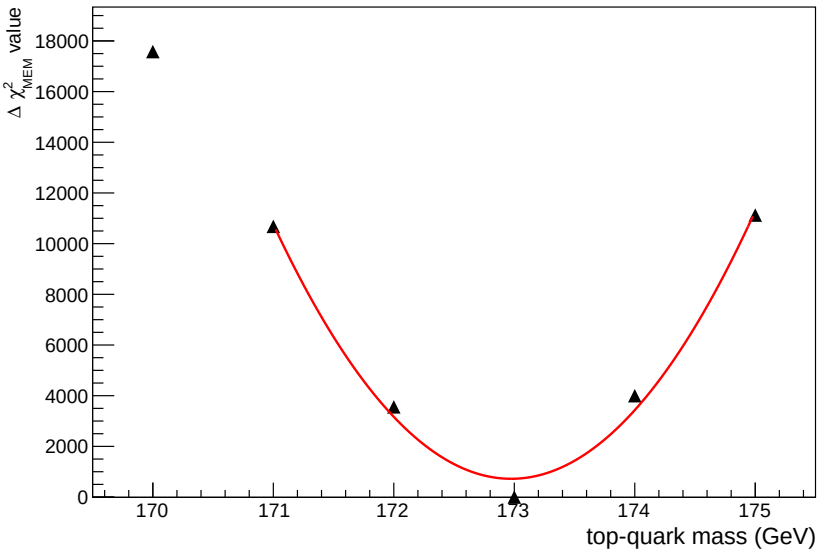

Figure 5.11: $\Delta\chi^2_{\mathrm{MEM}}$ curve obtained for 4000 generator-level events created in MadGraph with $m_t = 173\,\mathrm{GeV}$ fitted with a quadratic function. Minimising this function results in a top-quark mass of $172.97 \pm 0.02\,\mathrm{GeV}$.

ues by determining the top-quark mass using reconstructed events. As was the case for the generator-level events, the individual event probabilities also vary significantly as can be seen from Figure 5.12. Again the two upper plots show a shape consistent with the expectation while the two middle ones seem to contain less relevant information. Nevertheless the minimisation procedure results in a best estimate of of $173.8 \pm 0.2\,\mathrm{GeV}$, as can be seen from the bottom plot in Figure 5.12 containing both the quadratic fit function and the overall $\Delta\chi^2_{\mathrm{MEM}}$ values.
The obtained result deviates slightly from the top-quark mass used in the simulated events, 172.5 GeV, which can be explained by both the limited statistics and the more challenging conditions existing for reconstructed events. Hence in order to ensure the value obtained for the Matrix Element estimator can be trusted when considering reconstructed events, a dedicated calibration procedure should be performed as will be discussed in Chapter 6 for the measurement of the right-handed tensor coupling. Nevertheless, taking into account the large statistical uncertainty and the different systematic effects likely to influence the obtained top-quark mass measurement, this feasibility study has proven that the Matrix Element method is capable of providing very precise and accurate results.

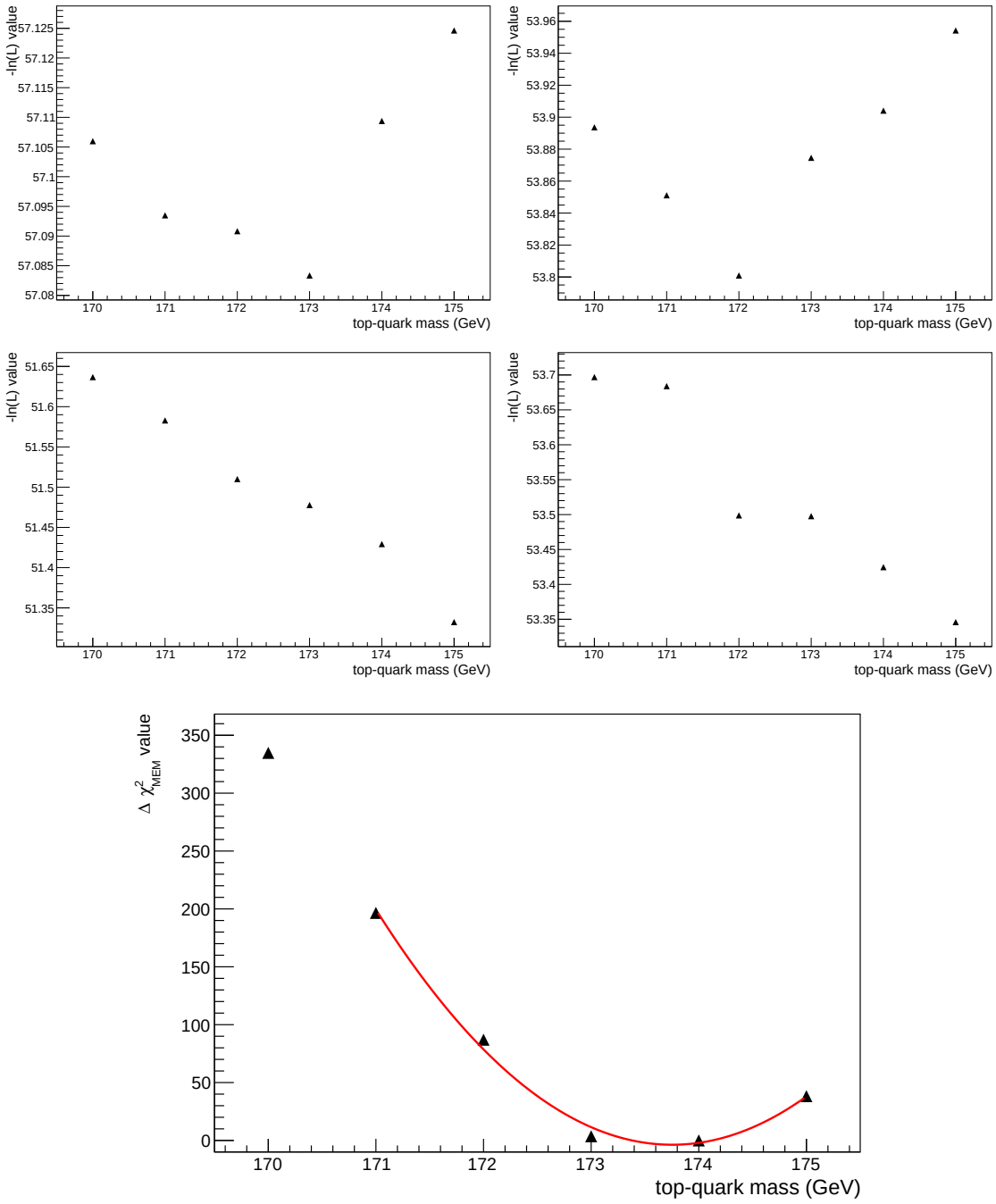

Figure 5.12: Individual event probabilities for the measurement of the top-quark mass (upper two rows) using reconstructed events, which get combined in the overall likelihood (last row) from which the most optimal value of the top-quark mass is extracted.

Chapter 6

Measurement of anomalous couplings in top-quark pair decays

The technicalities of the Matrix Element method have been described in detail in Chapter 5 and will now be put to use in order to measure the anomalous couplings in the Wtb interaction vertex. The main idea behind this advanced technique consists of combining individual event probabilities into an overall $\Delta\chi^2_{\mathrm{MEM}}$ curve from which the considered theoretical parameter can be extracted. This Matrix Element estimator is obtained by minimising this curve, and is denoted as $\hat{g_R}$ for the measurement of the right-handed tensor coupling. However, before the actual measurement can be performed, a number of tests have to be carried out in order to ensure the Matrix Element estimator is behaving properly.

The performance of the Matrix Element estimator has to be studied carefully since the Matrix Element method is likely to be influenced by the various assumptions and simplifications that have been introduced. These might influence the outcome of this estimator in such a way that it significantly deviates from the correct result. As a result, a detailed investigation of the correspondence between the estimator and expectation has been performed and will be discussed in Section 6.1. Once the required calibrations have been identified, the developed procedure to determine the Matrix Element estimator can be applied on the collision events collected by the CMS detector. The final results obtained for the measurement of the right-handed tensor coupling of the Wtb interaction will be given in Section 6.2.

6.1 Performance of the Matrix Element estimator

In order to ensure that the Matrix Element method behaves accordingly and results in the correct outcome, the properties of estimator will first be determined using simulated events. For these types of events, which have been generated using the Standard Model configuration, the expected outcome should be fixed by definition. Hence in case any deviation from the Standard Model expectations is observed for the outcome of the Matrix Element estimator, the developed procedure has to be calibrated.
Such a discrepancy does not hamper the applicability of the Matrix Element method,

but simply implies that its outcome should be corrected for the observed non-optimal behaviour. Possible reasons for such a deviation to occur are most likely caused by the assumptions in the model describing the anomalous couplings or by the simplifications applied in the developed procedure.
Hence the performance of this Matrix Element estimator will be studied in detail and various aspects will be considered. The first two performance tests focus on the correspondence between the outcome of the Matrix Element estimator $\hat{g_R}$ and the outcome expected from the simulated samples g_R^{MC}, which are supposed to be identical. The possible deviations from this ideal situation are described in Equation (6.1) where the slope a is the subject of the first test, the so-called linearity test, and is supposed to be equal to 1. The bias b for the developed procedure will be studied when determining the offset of the estimator. The third and last performance test which has been considered will focus on the statistical properties of the variance of the Matrix Element estimator.

$$\hat{g_R} = a \cdot g_R^{MC} + b \tag{6.1}$$

6.1.1 Linearity test

In this thesis it has been opted for to perform the linearity test independently from the offset determination, because for the first one several samples generated with different values of the right-handed tensor coupling are required. Since this is a rather challenging procedure for reconstructed events, especially compared to the ease with which generator-level samples can be created, it has been decided to use generator-level events for this study. The created samples will contain 20 000 events in order to be comparable in size to the considered data sample. The influence of the reconstructed events will be incorporated afterwards in the offset-determination study, which will thus only look at the outcome obtained for the Standard Model configuration.

The goal of the linearity test is to ensure that the outcome of the Matrix Element estimator for simulated events is directly related to the input value of the g_R coefficient. Hence the measurement will be repeated for various generator-level samples, all created by imposing a different g_R^{MC}-value during the generation process.
It is necessary to perform this linearity test since a deviation might occur in case the model describing the anomalous couplings in the Wtb interaction vertex is affected by the various assumptions that have been introduced. A second possible explanation for observing a difference between the value of the Matrix Element estimator and the g_R^{MC} value imposed during the generation process can be found in the applied event-selection criteria. If the varying detector acceptance conditions for different g_R^{MC} values are not perfectly described by the cross-section normalisation, the outcome of the estimator can be significantly influenced. However, as was discussed in detail in Section 5.4, the full event-selection chain for reconstructed events cannot be applied for generator-level events and will have to be mimicked by applying a limited set of selection criteria. These can be found in Table 5.2 and do not include the additional analysis-specific criteria discussed in Section 4.2.

The g_R^{MC} value imposed during the generation process is then compared with the g_R values estimated with the Matrix Element method using generator-level events with the basic event-selection criteria applied. In the ideal case both values should be identical, which is indeed the case for this analysis as can be concluded from Figure 6.1. The linearity test clearly indicates that the dependency of the estimated values of g_R on the values g_R^{MC} can be described by a straight line with slope close to 1 (a = 0.97). Even though the actual bias of the method will be determined afterwards using reconstructed events, the offset obtained here (b = -0.005) corresponds well with the expectation of an unbiased estimate. Hence it can be concluded that the Matrix Element method behaves properly when considering generator-level events.
The linearity test has been limited to the range $[-0.17, 0.17]$ since this is the relevant region where the Standard Model configuration should be recovered. The deviation from the expected shape outside the region of interest is most likely caused by the simplifications applied in the constructed theoretical model describing the Wtb interaction vertex or the assumptions made while developing the analysis procedure and the estimator definition.

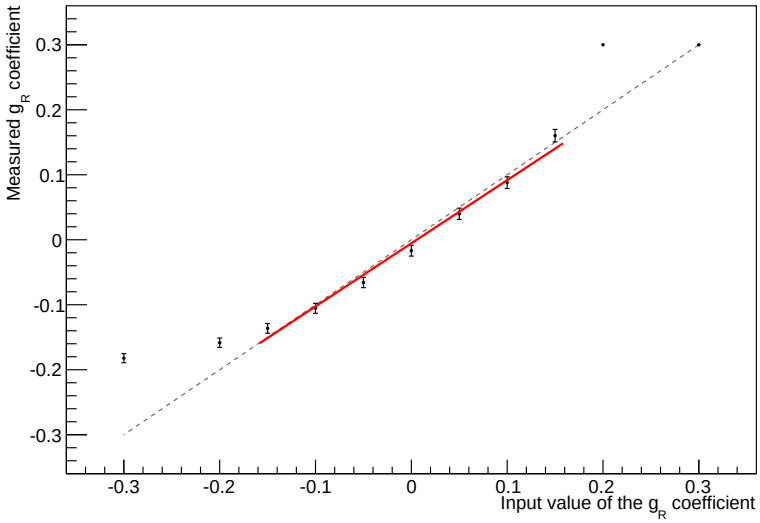

Figure 6.1: Outcome of the linearity test based on generator-level events, for which the obtained curve is described by a straight line. Hence both the developed model and method behave as expected and no calibration is required.

6.1.2 Offset calibration

The aim of the second performance test is to determine whether the Matrix Element estimator has a bias. Even though the linearity test has proven that an excellent agreement exists when analysing generator-level events; both for the slope and the bias; this cannot be generalised immediately to reconstructed events. Since the bias might be affected by the different nature of reconstructed events, it should be determined using the reconstructed events fulfilling the entire event-selection chain introduced in Chapter 4. Since these events have been created by imposing $g_R^{MC} = 0$, the outcome of the Matrix Element estimator should be zero as well. In case a deviation from the

expectation would be observed, the bias will be taken into account and the final result will be calibrated accordingly.

Although on generator-level the event kinematics are correct by definition as well as the jet combination, this is not always the case when performing the analysis on reconstructed objects. The variance on the event kinematics can be large and the wrong jet combination can be used as input for the Matrix Element method. This might result in difficulties for the numerical integrations performed in the Matrix Element method. Before studying the bias on the g_R estimator, some additional event cleaning criteria will need to be applied in order to filter out these problematic events.

Understanding the nature of reconstructed events

Applying the analysis procedure on reconstructed events does not result in the expected outcome of the Matrix Element estimator due to a small fraction of events for which the event probability appears to be wrongly calculated. This is not completely unexpected since reconstructed events are likely to be influenced by detector inefficiencies and ill-determined event kinematics. In addition, the Matrix Element treats all events as semi-leptonic top-quark pair decays such that any deviation from the expected topology can result in an incorrect event probability.

The distribution of the $-\ln(\mathcal{L}_{\mathrm{MEM}})$ value obtained for $g_R = 0$, shown in Figure 6.2 for both generator-level and reconstructed events, clearly indicates a difference in shape between both types of events. The right distribution is obtained using reconstructed $t\bar{t}$ events for which the four jets have been correctly matched with the generator-level parton since this allows to exclude the contribution of events for which the wrong jet combination is used. Hence the significant difference in shape for the reconstructed and generator-level events, for which the obtained distribution is shown on the left, implies that this difference is definitely caused by the different nature of reconstructed events and not by badly reconstructed event topologies.
The presence of the tail for the reconstructed events seems to suggest that for a small fraction of events a significantly lower event probability is obtained from the Matrix Element integration procedure. In order to determine whether this behaviour changes when the wrong jet combinations are used during this integration, the same procedure has been applied on a sample of $t\bar{t}$ events for which at least one jet has not been matched with the correct generator-level parton[1].
The normalised distributions for these $-\ln(\mathcal{L}_{\mathrm{MEM}})$ values obtained for $g_R = 0$ of both considered $t\bar{t}$ samples are given in Figure 6.3. Since the distribution obtained for the wrong event topologies is rather similar it can be concluded that using the wrong jet combination does not deteriorate the observed difference between generator-level and reconstructed events. Hence this proves that this feature is truly associated with applying this advanced analysis technique on reconstructed collision events.

[1]Since the two jets assigned to the hadronic decay of the W-boson can not be distinguished and will therefore both be considered by the Matrix Element integration procedure, this jet-assignment is allowed to be switched.

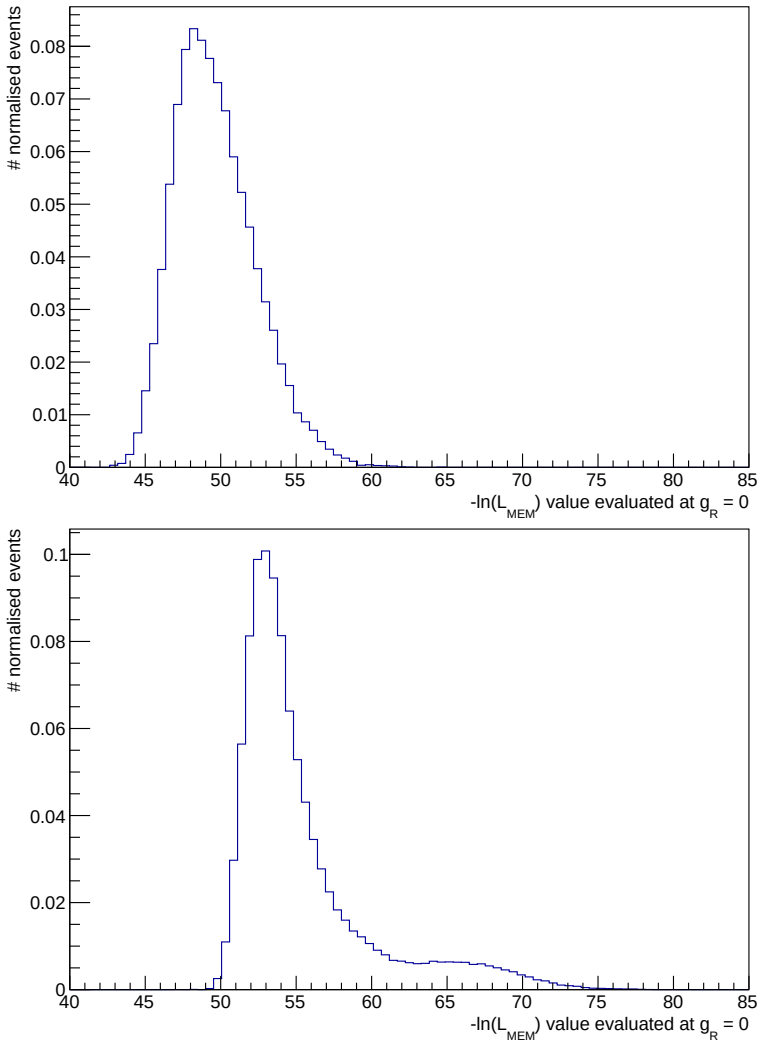

Figure 6.2: Normalised distribution of the $-\ln(\mathcal{L}_{\mathrm{MEM}})$ value obtained at $g_R = 0.0$ for both the generator- (top) and reconstructed-level (bottom) events.

A final attempt has been made at finding an explanation for this peculiar behaviour observed for the reconstructed events by investigating the dependence of other event variables on this $-\ln(\mathcal{L}_{\mathrm{MEM}})$ one. The goal is to find a distinguishing feature that could be used to reject the events negatively affecting the output of the Matrix Element method. In order to be unaffected by other influences, only the $t\bar{t}$ sample containing events for which the topology is correctly reconstructed has been considered.
This study will focus on the shape of the event probabilities calculated by the Matrix Element method. For each event this event probability is converted into a negative log likelihood according to Equation (5.2), for which the evolution as a function of the g_R value is expected to be described by a parabola.
For each event it can be determined whether this expected shape is recovered by looking at the value obtained for the second derivative evaluated at $g_R = 0$. In case the minimum of the parabola corresponds to the expected value of $g_R = 0$, this second derivative should be positive. The value given on the y-axis of the top histogram

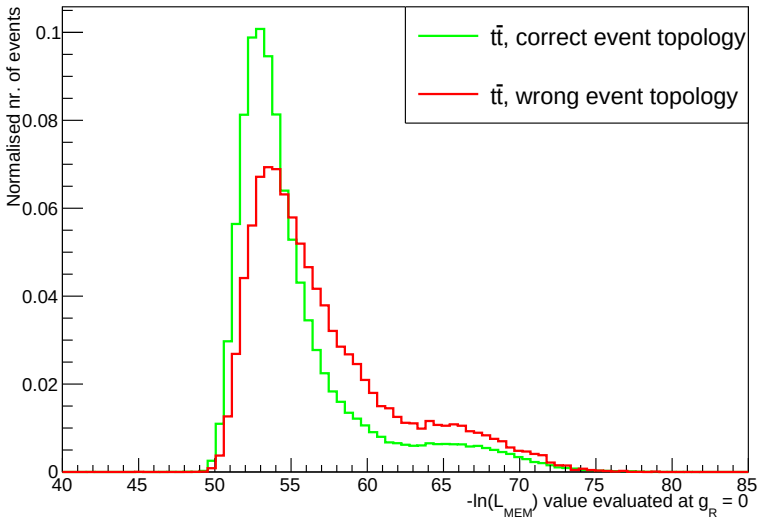

Figure 6.3: Normalised distribution of the $-\ln(\mathcal{L}_{\mathrm{MEM}})$-variable obtained at $g_R = 0$ for correctly reconstructed (green) and wrongly reconstructed $t\bar{t}$ (red) event topologies.

in Figure 6.4 has been calculated using the g_R points -0.1, 0.0 and 0.1 and has an average value of 0.82, which corresponds with the expectation of an average positive second derivative. Nevertheless a rather strange behaviour can be observed if this variable on the x-axis, the $-\ln(\mathcal{L}_{\mathrm{MEM}})$ value obtained at $g_R = 0.0$, becomes larger than approximately 65. Events with a very large, positive or negative, second derivative correspond to events containing a lot of information and are therefore not expected to be located in the tail of this distribution shown in Figure 6.2.
The second variable which has been considered represents the maximal variation in event-probability observed for each event, which is in general an indication of the steepness of the quadratic function fitted through the $-\ln(\mathcal{L}_{\mathrm{MEM}})$ values. This maximal variation has been determined by subtracting the negative likelihood value from the lowest one and is depicted on the y-axis of the bottom histogram in Figure 6.4. From this can be concluded that most events have a rather flat shape, while the events residing in the tail of the distribution of this $-\ln(\mathcal{L}_{\mathrm{MEM}})$ variable obtained at $g_R = 0.0$ have a significantly higher maximal variation for the event-probabilities. Hence these events have not only a higher negative log likelihood value assigned but the differences between the different $-\ln(\mathcal{L}_{\mathrm{MEM}})$ values are also significantly larger, again implying that this peculiar behaviour for reconstructed events is most likely caused by events for which the integration of the Matrix Element method did not manage to converge.

Unfortunately, the two histograms that have been considered did not allow to determine which specific event-characteristic is responsible for this peculiar behaviour when reconstructed events are analysed by the Matrix Element method. However they still allow to conclude that it starts to occur once the $-\ln(\mathcal{L}_{\mathrm{MEM}})$ variable obtained at $g_R = 0.0$ becomes larger than about 65. Hence the value of this variable will need to be limited in order to discard the events located in the tail of the distribution such that the result obtained from the Matrix Element method is no longer biased.

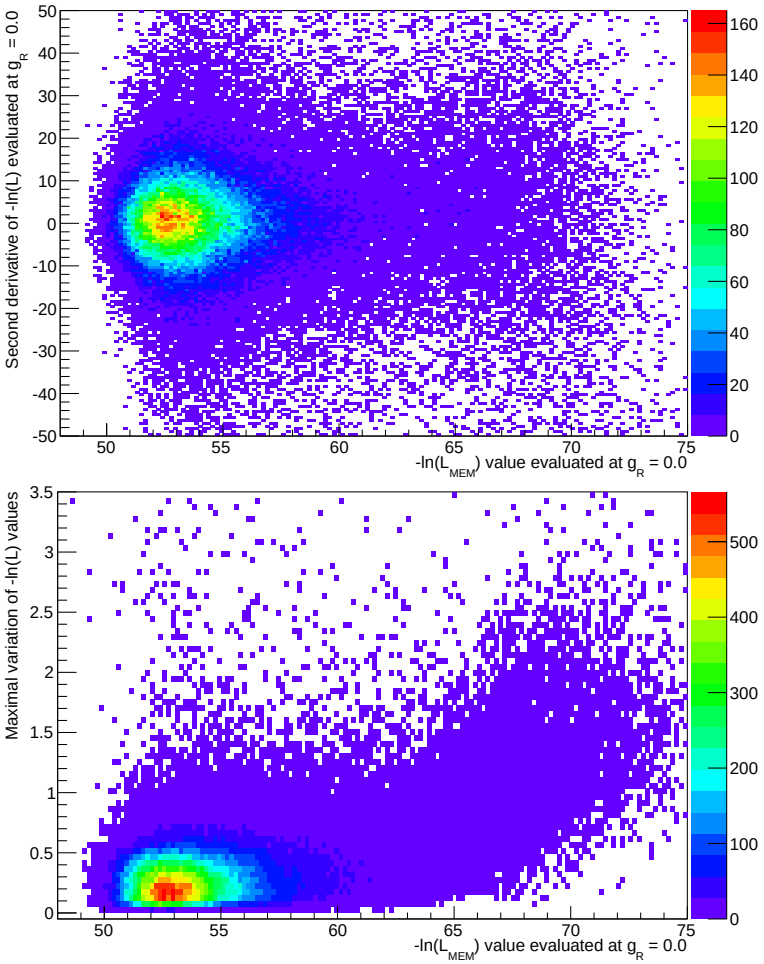

Figure 6.4: The second derivative of the negative likelihood values evaluated at $g_R = 0$ (top) and the difference of the maximal and minimal $-\ln(\mathcal{L}_{\mathrm{MEM}})$ value (bottom), both as a function of the value of the negative likelihood value for $g_R = 0$.

Matrix Element event-cleaning procedure

A dedicated event-cleaning procedure will be developed in order to ensure the events residing in the tail of the distribution shown in Figures 6.2 and 6.3 are excluded. No additional attempt is made to reduce the contribution of badly reconstructed event topologies since this is already taken care of by the event-selection criteria formulated in Chapter 4. An accurate estimation of the most optimal cut-value will now be performed by applying the analysis for several cut-values and determining for which value the Matrix Element estimator corresponds to the expectation of $g_R^{MC} = 0$.
Since the most optimal cut-value is likely to be located around 65, the range of interest where to apply the cut on has been restricted between 60 and 70. For each considered cut-value all events for which this $-\ln(\mathcal{L}_{\mathrm{MEM}})$ value obtained at $g_R = 0$ is smaller than the cut-value will be selected and the measurement of g_R will be performed using the full collection of simulated events surviving this cut. The obtained estimator value for each cut has been added in Figure 6.5, which clearly shows the sensitivity of the Matrix

Element estimator on the applied cut-value. The different points are then fitted using a polynomial of degree 3 such that the cut-value where $\hat{g_R} = 0$ can be derived. This corresponds to requiring this $-\ln(\mathcal{L}_{\mathrm{MEM}})$ value obtained at $g_R = 0$ to be smaller than 63.87.

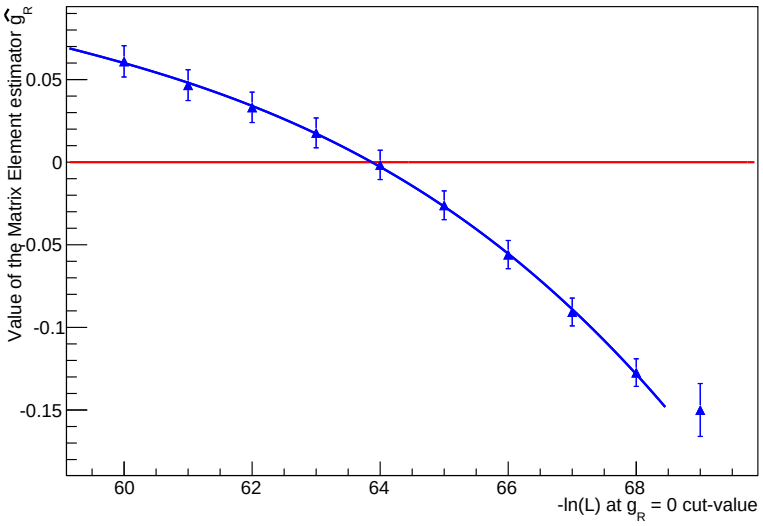

Figure 6.5: Obtained values of the Matrix Element estimator $\hat{g_R}$ when applying different restrictions on the $-\ln(\mathcal{L}_{\mathrm{MEM}})$ value obtained at $g_R = 0$. The optimal cut-value is determined using a 3^{rd} order polynomial.

After this event cleaning is performed, the bias is corrected for and the $\Delta\chi^2_{\mathrm{MEM}}$ obtained for the full collection of simulated samples corresponds reasonably well with the expected Gaussian behaviour around $g_R = 0$. As has been explained in Section 5.5, the minimisation procedure is applied by fitting these values with a quadratic function. It has been chosen for to restrict the range of this fit between $g_R = -0.15$ and $g_R = 0.15$ in order to avoid that the Matrix Element estimator is significantly influenced by deviations from this Gaussian description observed for higher g_R values. This minimisation procedure results in a g_R value of 0.0015 ± 0.0026, which is statistically compatible to 0.

Applying the Matrix Element method on reconstructed collision events has clearly indicated that these type of events are influenced by inefficiencies non-existing for generator-level events. Moreover, since the Matrix Element method treats every event as if it is a perfectly described semi-leptonic top-quark pair decay, any deviation from the expected topology is likely to result in the event probability being badly calculated due to a failing phase-space integration. Nonetheless, it has been established that the origin of this different behaviour is not caused by the applied event-selection procedure but is a true feature of any analysis deploying the Matrix Element method in a realistic collider environment.
Hence a detailed event-cleaning procedure needs to be applied in order to exclude the events causing this peculiar behaviour observed for reconstructed events. This has been accomplished by requiring the value of $-\ln(\mathcal{L}_{\mathrm{MEM}})$ obtained at $g_R = 0$ to be lower

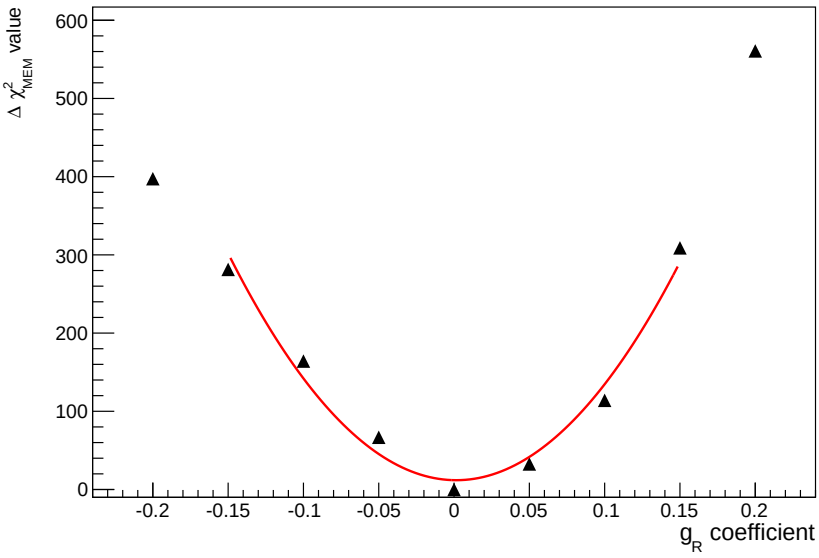

Figure 6.6: Obtained $\Delta\chi^2_{\mathrm{MEM}}$ curve using all simulated samples after requiring the $-\ln(\mathcal{L}_{\mathrm{MEM}})$ variable obtained at $g_R = 0$ to be smaller than 63.87.

than 63.87 for each event and will be applied in the remainder of this analysis. The remaining bias is statistically compatible with 0 and thus confirms that the presence of an offset is taken care of by the applied cut-value.

6.1.3 Statistical properties

The third and final test which has been performed in order to ensure the Matrix Element estimator $\hat{g_R}$ behaves properly is related to its statistical properties. This is a necessary aspect to study since it allows to understand whether the uncertainty obtained from the estimator is correct. In case the uncertainty of the estimator would be over- or underestimated, the uncertainty of the final result will need to be calibrated accordingly.
The statistical properties of this estimator are evaluated using a resampling technique which generates a set of samples by randomly selecting events from the full collection of simulated samples until the data luminosity of 19.6 fb^{-1} is reached. The considered events are all required to pass the full set of event-selection criteria, including the $-\ln(\mathcal{L}_{\mathrm{MEM}})$ cut-value discussed earlier.

The considerable amount of statistics available for the simulated samples allows to create 1000 of these samples, so-called pseudo-experiments, without introducing a significant correlation between the different samples. Each of these pseudo-experiments is a representation of the data sample and will be treated as such by the developed procedure to obtain the value of the Matrix Element estimator. The obtained minimum $\hat{g}_{R,i}$ and corresponding uncertainty $\hat{\sigma}_i$ for the considered pseudo-experiments are shown in Figure 6.7.

The mean value of the right-handed tensor coupling $\langle g_R \rangle$ for the different pseudo-

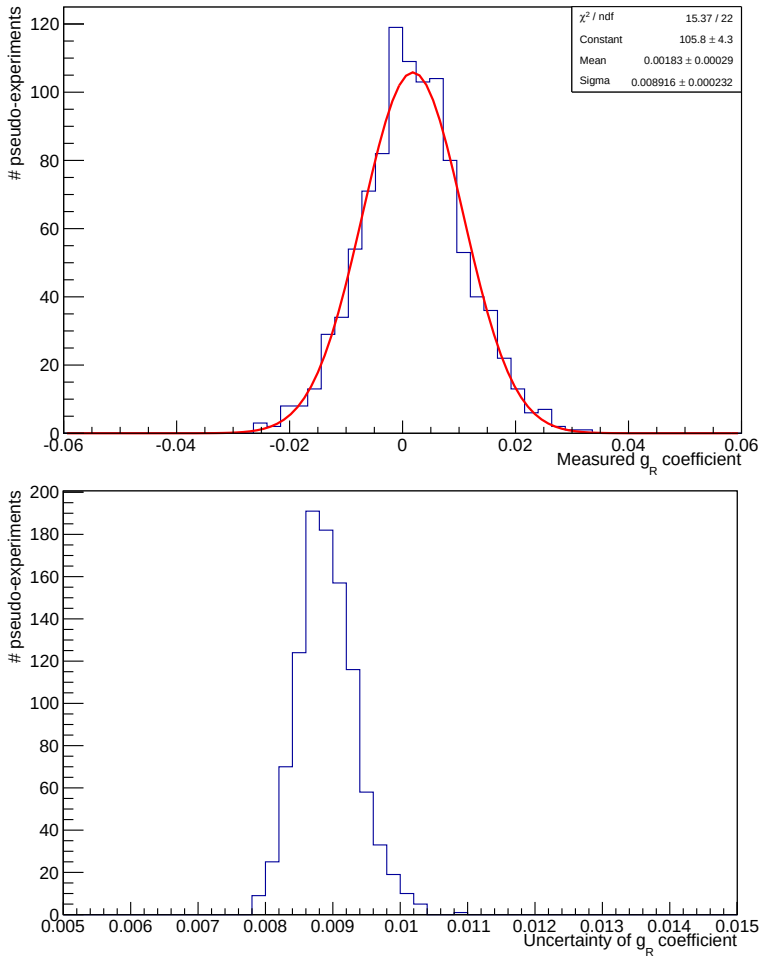

Figure 6.7: Distribution of the measured g_R value (top) and its uncertainty (bottom) for the 1000 considered pseudo-experiments.

experiments can be determined by applying a Gaussian fit on the distribution of the measured g_R coefficients. The pull, for which the distribution should correspond to a Gaussian function with mean of 0 and width of 1, can then be determined using:

$$\text{pull}_i = \frac{g_{R,i} - \langle g_R \rangle}{\sigma_i} \tag{6.2}$$

The obtained distribution for the pull is given in Figure 6.8, which clearly corresponds with the expected behaviour. The value of the width is equal to 0.979 ± 0.025, implying a perfect agreement is observed and thus no correction is needed for the estimated uncertainty on the Matrix Element estimator $\hat{g_R}$.

6.2 Measurement of g_R with the Matrix Element method

The different performance tests discussed in the previous section have clearly indicated that the Matrix Element estimator $\hat{g_R}$ is behaving adequately. Hence the full analysis

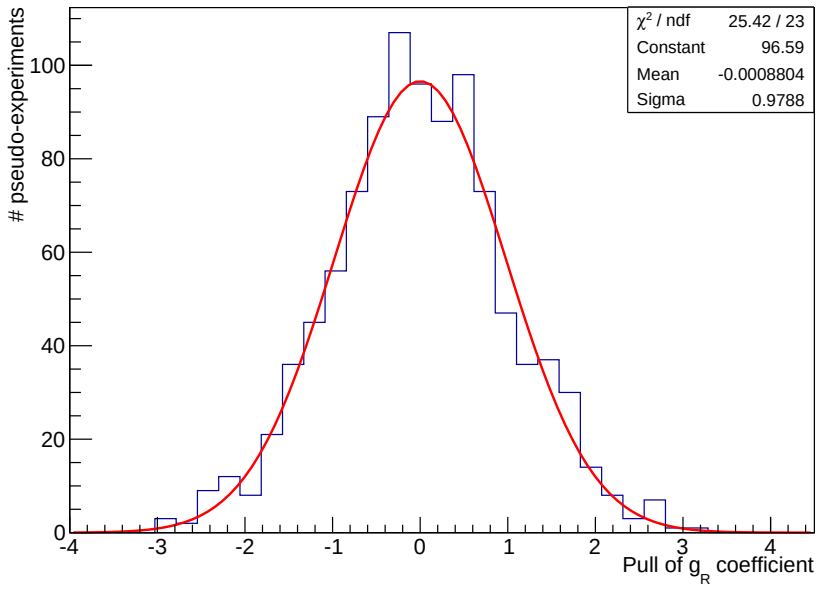

Figure 6.8: Pull distribution obtained for the 1000 considered pseudo-experiments, which can be described by a Gaussian function with mean = -0.001 ± 0.032 and width = 0.979 ± 0.025.

procedure can now be applied on the data events collected by the CMS experiment in order to determine the value of the right-handed tensor coupling of the Wtb interaction.

6.2.1 Results on data

Applying the full analysis procedure on the data events results in the $\Delta\chi^2_{\mathrm{MEM}}$ curve, obtained by converting the event-probabilities using Equation (5.3), shown in Figure 6.9. The outcome of the Matrix Element estimator is then determined by fitting this curve with a 2^{nd} degree polynomial, since a Gaussian behaviour is assumed.

Applying the minimisation method on this fitted function results in a g_R-value for the selected data events of:

$$g_R = -0.0071 \pm 0.0083 \tag{6.3}$$

This value of the right-handed tensor coupling of the Wtb interaction vertex indicates an excellent agreement with the Standard Model, indicating that this measurement does not observe any influence of anomalous couplings in the decay of top-quark pairs. The statistical uncertainty obtained for the data events can be compared with the uncertainty expected from the study of the statistical properties of the Matrix Element estimator, which have been discussed in Section 6.1.3. Also here an excellent agreement is obtained.

Comparing Figure 6.6 and 6.9 indicates that the assumed Gaussian behaviour is not perfectly achieved and especially for g_R values further away from zero the deviation becomes rather significant. Nevertheless, this does not hamper the applicability of the Matrix Element method since a similar behaviour is observed for data and simulation.

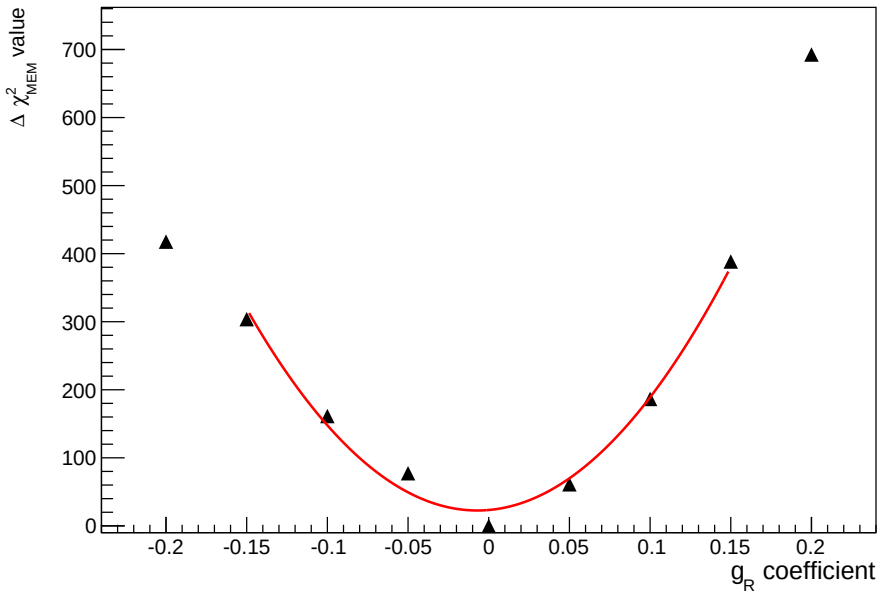

Figure 6.9: Obtained $\Delta\chi^2_{\mathrm{MEM}}$ curve for the data-sample at 8 TeV.

This can be seen from Figure 6.10, where the added lines are merely connecting the dots.

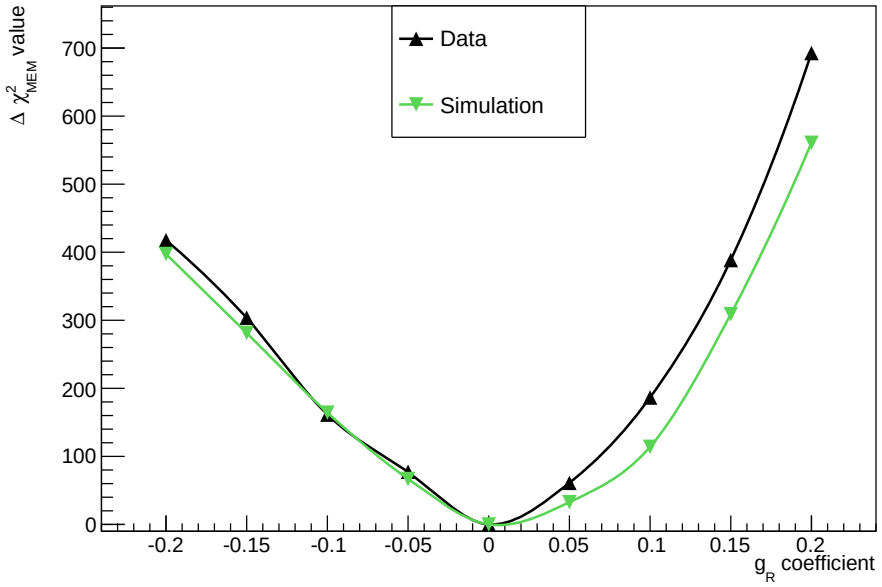

Figure 6.10: Obtained $\Delta\chi^2_{\mathrm{MEM}}$ curve for both the data and full simulation sample at 8 TeV.

6.2.2 Systematic uncertainties

Several systematic effects might change the outcome of the Matrix Element estimator and their influence should be carefully investigated. Such a systematic uncertainty can be caused by several sources: an incomplete understanding of the detector

performance, an incorrect modelling of the simulation, the assumptions made in the likelihood-extraction procedure, or even by the algorithms used during the reconstruction process.
Their effect can in general be determined by varying the parameter responsible for the considered systematic uncertainty up- and downwards with one standard deviation for the entire collection of considered simulation samples. The entire analysis procedure used to obtain the g_R measurement on the data sample is applied on these simulated samples, resulting in a value of the Matrix Element estimator for the considered systematic shift. The systematic uncertainty used in this analysis is given in Equation 6.4, except if the corresponding statistical uncertainty is larger than the actual systematic uncertainty the former one will be used.

$$\Delta g_R = \frac{|g_R^{\text{down}} - g_R^{\text{up}}|}{2} \tag{6.4}$$

The complexity of the Matrix Element method has an important consequence for the way the systematic uncertainties will be evaluated. Since for some of the considered systematic effects this up- and downwards shift results in altered event kinematics, the full integration procedure of the Matrix Element method should be repeated. In order to reduce the required processing time, it has been opted for in this analysis to only redo the calculation for the relevant background samples.
This is an acceptable approach since the contribution of the considered background samples is severely restricted by the stringent event selection. Hence they are not expected to significantly influence the obtained measurement of the g_R coefficient, as can be seen from Table 6.1. Here the value of the Matrix Element estimator is given when the main $t\bar{t}$ signal sample is combined with each of the different background samples that have been considered. An individual measurement cannot be performed for these background samples due to the limited statistics available after the full event-selection chain is applied. Hence the contribution of each separate background sample can be derived from this table by comparing the combined result with the value obtained using all the available simulated samples, denoted as "Total simulation" in this table.

Table 6.1: Obtained result of the Matrix Element estimator when combining the signal sample ($t\bar{t}$) with the different background samples considered.

Sample	Obtained $\hat{g_R}$-value
$t\bar{t}$	-0.0013 ± 0.0027
$t\bar{t}$ + Single top	0.0010 ± 0.0026
$t\bar{t}$ + W-jets	-0.0009 ± 0.0026
$t\bar{t}$ + Z-jets	-0.0013 ± 0.0026
Total simulation	0.0015 ± 0.0026

As was already clear from Table 4.1, Table 6.1 confirms that the most important

background sample corresponds to the single-top quark decay while the remaining ones only slightly alter the value obtained for the $t\bar{t}$ signal sample. Hence the single-top background will be taken into account for the computation of the systematic effect of the jet energy scale uncertainty, which is expected to be one of the dominating systematic effects of this measurement. The systematic uncertainty corresponding to the jet energy resolution on the other hand is presumed to be much smaller and will therefore be determined using only the signal $t\bar{t}$ sample. Nevertheless, since the up- and downwards shift of the relevant parameter are compared to each other, the obtained overall systematic uncertainty is not supposed to be influenced in case some of the background samples are not considered.
The remaining systematic uncertainties that have been studied do not change the kinematics of the event, but simply alter the relative contribution of each event to the overall $-\ln(\mathcal{L}_{\mathrm{MEM}})$ value. Hence the full set of simulation samples will be considered for these systematic effects since no additional computing time is required. This has as advantage that in case the considered systematic is dominated by the corresponding statistical uncertainty this uncertainty can be determined using all the available statistics.
However it should be noted that this is a rather conservative approach since the correlation between the different systematic samples implies that the statistical uncertainty on these systematic shifts is smaller. Hence more accurate results could be obtained in case a dedicated resampling technique would be applied. Nevertheless for this measurement the overall uncertainty is dominated by the systematic effects, implying that only a minor improvement on the total uncertainty would be achieved by applying this technique.

Once the method to assess the different systematic uncertainties has been established, their effect has been propagated to the outcome of the Matrix Element estimator. The influence of each considered systematic uncertainty can be found in Table 6.2, where the values highlighted using boldface font represent the actual uncertainty used to determine the total systematic effect. From this summarised overview can directly be concluded that the majority of the considered systematic uncertainties is dominated by the corresponding statistical uncertainty and thus have a very small influence. A more detailed evaluation of the different systematic uncertainties can be found below.

Jet Energy scale
The energy of the reconstructed PF jets is calibrated using dedicated p_{T}- and η-dependent jet energy scale calibrations, as was explained in Section 3.3.4. Since the Matrix Element method evaluates in a direct manner the kinematics of the different final state particles in the event, the uncertainty on the jet energy scale should be propagated to the Matrix Element estimator. The importance of this systematic effect, one of the leading systematic uncertainties for this measurement, can be understood by the presence of four jets in the topology of top-quark pair events.

Jet Energy resolution
The different energy resolution observed in data and simulation requires the

Table 6.2: Overview of the different systematic uncertainties considered for the measurement of the right-handed tensor coupling g_R. For each contribution the larger among the estimated shift and its statistical uncertainty is quoted, as indicated by the bold script.

Source	Estimated effect on g_R
Jet energy scale	**0.0056** $\pm$ 0.0018
Jet energy resolution	0.0010 $\pm$ **0.0019**
b-tagging efficiency	0.00001 $\pm$ **0.00185**
mis-tagging efficiency	0.00004 $\pm$ **0.00185**
Pileup reweighting	0.0008 $\pm$ **0.0019**
Background composition	0.0028 $\pm$ **0.0032**
Offset calibration	0.0015 $\pm$ **0.0026**
Minimum-extraction	**0.0039**
Q^2-scale	0.00004 $\pm$ **0.00396**
ME-PS matching	**0.0097** $\pm$ 0.0059
Total	0.0137

energy of the simulated events to be smeared using the JER correction factor. In order to evaluate the possible effect on the Matrix Element estimator, the systematic uncertainty originating from this η-dependent correction has been calculated but appears to be dominated by the corresponding statistical uncertainty.

Efficiency of the b-jet identification

The different efficiency of the b-jet identification algorithm in data and simulation has been discussed in detail in Section 4.3 and necessitated the introduction of p_{T}-dependent scale-factors, determined separately for light-flavoured and b- and c-flavoured jets. Even though the scale-factor itself is identical for the b- and c-quark jets, the uncertainty of the latter one is defined to be twice as large. Since the uncertainties of the light- and heavy-flavoured jets are assumed to be uncorrelated, they have been determined separately. In both cases, the obtained systematic uncertainty is negligible and is dominated by the statistical uncertainty.

Pileup reweighting

The number of additional pile-up interactions in simulation is obtained by reweighting the mean number of interactions in each event. In order to estimate the effect of this systematic uncertainty, this mean number of interactions

has to be shifted with ± 5%. This takes into account the luminosity uncertainty, the uncertainty on the total inelastic cross-section and even an additional uncertainty to cover the pileup modelling. However, the performed measurement is insensitive to this reweighting procedure and the obtained systematic uncertainty is again dominated by the statistical uncertainty.

Background composition

A different composition of the background samples might alter the obtained g_R value since it will change the relative contribution of each background process to the overall $\Delta\chi^2_{\text{MEM}}$ shape. The corresponding uncertainty is calculated by comparing the nominal value obtained for the g_R coefficient with the one obtained if no background events are taken into account, which is a very conservative approach. Hence this systematic uncertainty is not determined using the formula given in Equation (6.4), but is calculated as the absolute value of the difference between the two g_R values. Even with this conservative approach, this is not one of the main systematic uncertainties since the effect of the different background samples is very small due to the stringent event-selection criteria applied.

Offset calibration

The offset calibration discussed in detail in Section 6.1.2 is limited in precision by the statistics of the simulated event samples. Hence the bias obtained for the full collection of simulated samples is quoted as systematic uncertainty.

Minimum-extraction method

Since the obtained result relies heavily on the fitting procedure developed to extract the outcome of the right-handed tensor coupling from the Matrix Element output, the influence of applying a fit range different from $[-0.15, 0.15]$ has been taken into account. Hence four different fit ranges have been considered, and for each of these it has been studied whether the obtained outcome is shifted.

In order to include this effect in the correct way, the shift observed for the data measurement should be compared to the shift obtained using the full collection of simulated samples since only the overall shift is relevant to evaluate whether a systematic effect exists. This comparison can be found in Table 6.3 where the results for the four considered fit ranges are given.

A small discrepancy between the different fit-ranges can be observed and thus, following a rather conservative approach, the largest of the overall shift values will be quoted as the systematic uncertainty associated with the method developed to extract the minimum from the fit on the $\Delta\chi^2_{\text{MEM}}$ values.

Hadronisation and factorisation scale

This systematic effect takes into account the uncertainty on the amount of initial- and final-state radiation and on the choice of the Q^2-scale during the event generation. This uncertainty is evaluated using dedicated $t\bar{t}$ samples for which the Q^2-scale is varied up and down by a factor 4 while simultaneously the amount of initial- and final-state radiation is increased and decreased.

However, this systematic uncertainty can not be determined in the similar

Table 6.3: Observed shift of the g_R measurement when comparing the different fit-ranges to the standard adopted one of $[-0.15, 0.15]$. The overall shift is defined as the absolute value of the difference between the shift obtained for the data and simulation measurement.

Fit range		Data-shift	Simulation-shift	Overall shift
Wide	$[-0.20, 0.20]$	0.0103	0.0100	0.0002
Narrow	$[-0.10, 0.10]$	0.00050	0.0088	0.0039
Zoomed	$[-0.05, 0.05]$	0.0100	0.0071	0.0029
Asymmetric	$[-0.05, 0.015]$	0.0117	0.0119	0.0003

straightforward manner as the other uncertainties but requires some significant adaptations. This because the shape of the $\Delta\chi^2_{\text{MEM}}$ values for both the upwards and downwards shift does not correspond with the expectation, but on the contrary exhibits a rather peculiar behaviour. The $\Delta\chi^2_{\text{MEM}}$ values associated with g_R values lower than -0.05 are about 20 times higher than the rest of the $\Delta\chi^2_{\text{MEM}}$ values related to higher g_R values, implying that no quadratic shape can be observed.
Such a unrealistic deviation seems to suggest that this effect is not an actual consequence of this systematic uncertainty but is caused by the Matrix Element method not being able to describe the introduced shift in parameter of this Q^2-scaling when $g_R >> 0$. This is not completely illogical since the shifts applied for this systematic uncertainty are extreme and are in addition determined with significantly lower statistics than the nominal $t\bar{t}$ sample.
Unfortunately this behaviour significantly complicates the determination of this systematic effect since the currently considered fitting range cannot be used. Hence in order to ensure the alternative fit-ranges considered before do not influence the systematic uncertainties, the total systematic uncertainty has been determined for each of these four fit-ranges. The obtained results can be found in Table 6.4, which clearly indicate an almost perfect agreement between the different fit ranges.

Table 6.4: Total systematic uncertainty, with the exception of the uncertainty associated with the Q^2-scale, for the various fit-ranges that have been considered.

Fit range		Total systematic uncertainty
Wide	$[-0.20, 0.20]$	0.0125
Narrow	$[-0.10, 0.10]$	0.0125
Zoomed	$[-0.05, 0.05]$	0.0133
Asymmetric	$[-0.05, 0.15]$	0.0129

As a result, this insensitivity of the outcome on the considered fit range will also be assumed for the systematic uncertainty associated with the factorisation

and hadronisation scale, which can only be determined using this asymmetric range due to the limited number of $\Delta\chi^2_{\mathrm{MEM}}$ values available. This approach resulted in the systematic uncertainty given in Table 6.2.

Matrix Element - Parton Shower Matching Threshold
The uncertainty on the threshold applied for the matching of tree-level matrix-elements with the parton showers is evaluated using dedicated $t\bar{t}$ samples. This threshold has been scaled up and down with a factor 2 in these samples.

Besides the systematic uncertainties summarised in Table 6.2, this measurement might be sensitive to an additional systematic effect that should be investigated in detail. This corresponds to the application of the $-\ln(\mathcal{L}_{\mathrm{MEM}})$ cut-value, for which the optimal value is determined using simulation events as was discussed in Section 6.1.2. In case the conditions for this $-\ln(\mathcal{L}_{\mathrm{MEM}})$ variable in data deviate from those in simulation, implementing this cut might introduce a bias for the data events and thus result in a systematic effect that should be taken into account.
Hence it has been investigated whether the conditions are similar for data and simulation and whether the dependency of the g_R result for data corresponds with the one obtained from simulation around the chosen $-\ln(\mathcal{L}_{\mathrm{MEM}})$ cut-value. At first the $-\ln(\mathcal{L}_{\mathrm{MEM}})$ variable evaluated at the Standard Model configuration is compared, as can be seen from the top distribution in Figure 6.11. Taking into account the statistical uncertainties on the data sample, an reasonable agreement is observed. The scan on the different cut-values is shown as the bottom plot of Figure 6.11 and, within the statistical uncertainties, confirms this that the dependency on g_R of the cut-value is similar.
The very good agreement between data and simulation obtained for this cut-value determination indicates that no additional systematic uncertainty needs to be taken into account for this procedure other than the systematic uncertainty on the determination of the bias on the g_R estimator.

The combination of the considered systematic uncertainties with the g_R measurement from the data sample results in a final g_R value equal to:

$$g_R = -0.0071 \pm 0.0083\,(\mathrm{stat.}) \pm 0.0137\,(\mathrm{syst.}) \tag{6.5}$$

This is compatible with the Standard Model hypothesis, which corresponds to a right-handed tensor coupling of 0 such that the Wtb interaction can be uniquely represented by the left-handed vector coupling V_L.

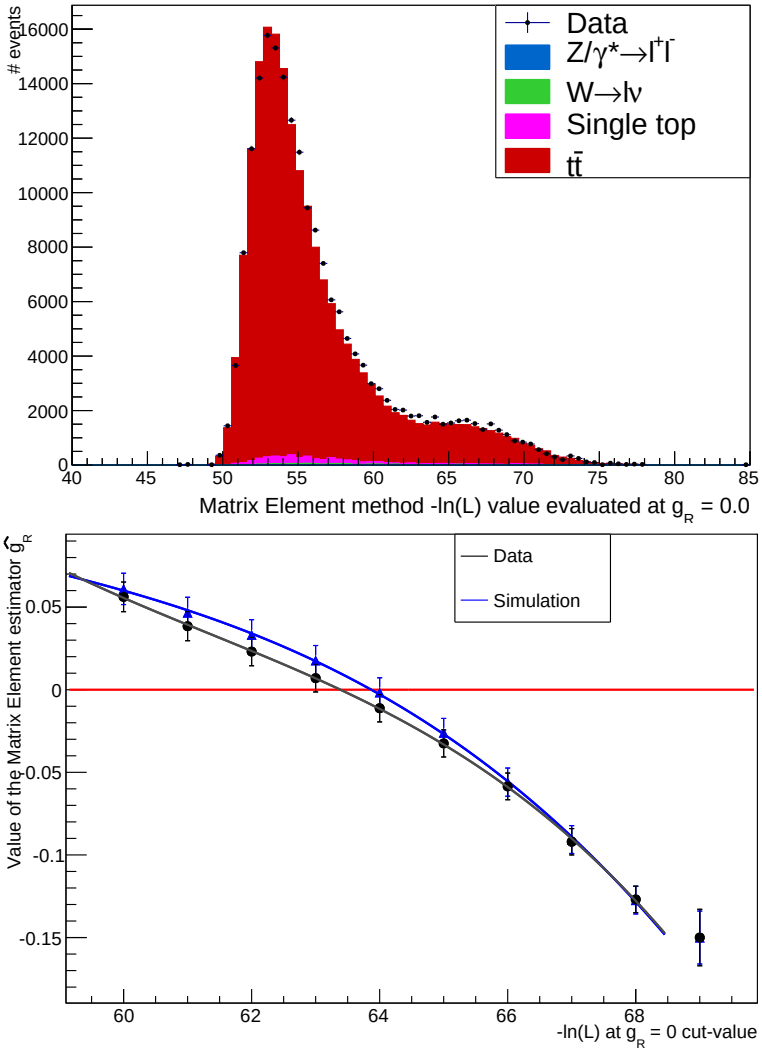

Figure 6.11: The distribution of the $-\ln(\mathcal{L}_{\mathrm{MEM}})$ variable (top) and the dependency of g_R on the cut-value (bottom).

Chapter 7
Conclusions

The discovery of the Brout-Englert-Higgs boson in the summer of 2012 provided once more a confirmation of the correctness and the predictive power of the Standard Model. Nevertheless several questions remain unanswered and the possible presence of physics phenomena beyond the Standard Model should be investigated in detail. An excellent window for such studies has been opened with the discovery of the top-quark in 1995 by the CDF and DØ collaborations. The Large Hadron Collider is capable of producing enormeous amounts of top-quark pairs and is thus used extensively to ensure a detailed measurement of various top-quark properties.
In this thesis a thorough study of the interaction vertex describing the decay of the top-quark, almost exclusively resulting in a W-boson and bottom quark, was conducted. This Wtb vertex is purely characterised by a single left-handed vector coupling in the Standard Model, but could nevertheless be influenced by new-physics phenomena. Hence the presence of these so-called anomalous couplings, which are capable of altering the decay of the top-quark, has been studied in detail using the μ+jets $t\bar{t}$ events produced during the 2012 LHC run and collected by the CMS experiment at a centre-of-mass energy of 8 TeV.

The measurement discussed in detail in this thesis is performed using a Matrix Element method, of which the details are explained in Section 5.1, and is the first direct measurement of the right-handed tensor coupling coefficient of the Wtb vertex. Since the Matrix Element method analyses each event individually in order to extract the maximum available experimental information, the selected events should correspond as closely as possible with actual top-quark pair decays. This has been obtained by applying several analysis-specific event-selection criteria, with the double b-tagging requirement of the b-quark jets resulting in the most significant background separation and event reduction.
In order to apply the Matrix Element method a detailed theoretical description of the hard scattering process of interest, the decay of the top-quark for this analysis, is required. This has been achieved by developing a new FeynRules model, based on the Standard Model, that perfectly incorporates the full Wtb Lagrangian. Using this model, the Matrix Element method is used for the measurement of the right-handed tensor coupling g_R:

$$g_R = -0.0071 \pm 0.0083\,(\text{stat.}) \pm 0.0137\,(\text{syst.}) = -0.0071 \pm 0.0160, \qquad (7.1)$$

which is in perfect agreement with the Standard Model prediction ($g_R = 0$). In this measurement the other coupling coefficients in the Wtb vertex have been taken equal to their Standard Model value ($V_L = 1$ and $V_R = g_L = 0$).

The obtained measurement should also be compared with the current experimental limits existing for this right-handed tensor coupling, given in Section 1.2.2. These have been obtained by combining the latest information available on the W-boson helicity fractions from both the ATLAS and CMS experiments. An accurate determination of these physical quantities signficantly limits the coupling coefficients of the Wtb interaction. Figure 7.1 contains the exclusion limits for the tensor coupling coefficients obtained by the CMS experiment using 7 TeV of semi-leptonically decaying ($\ell = \mu$ or e) top-quark pair events.
From the 1σ and 2σ contours it can be concluded that our g_R measurement is in perfect agreement with the previous measurements and even significantly improves the precision of this coupling coefficient.

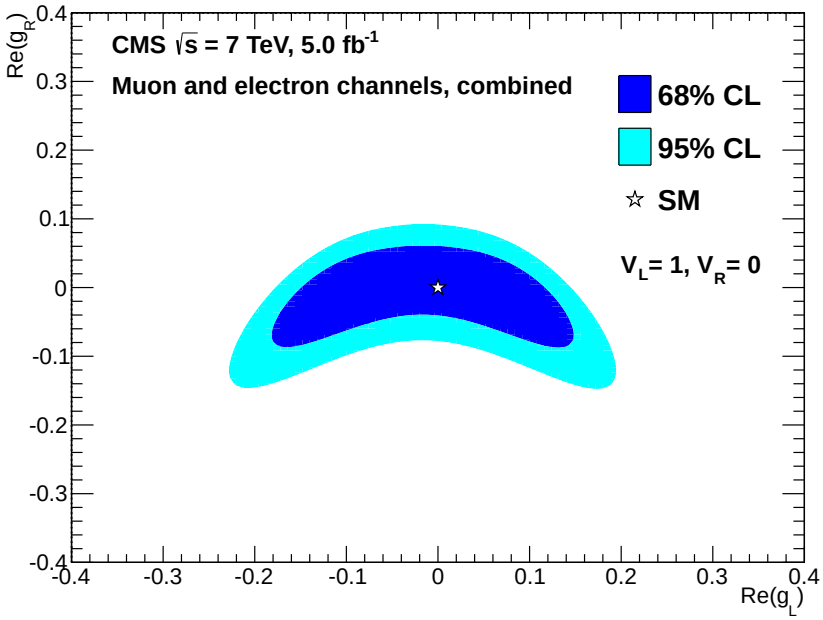

Figure 7.1: Limits on the real components of the anomalous couplings g_R and g_L at 68% and 95% CL [1].

Perspectives

The uncertainty of this measurement is clearly dominated by the systematic uncertainty, for which the leading ones correspond to the Matrix-Element and Parton-Shower matching and the jet energy scale uncertainties, as summarised in Table 6.2. Hence a significant improvement for the total uncertainty can be achieved if both systematic effects can be reduced.
The leading systematic, related to the ME-PS matching, is determined using dedicated

$t\bar{t}$ samples but with rather limited statistics. Hence for this systematic effect the corresponding uncertainty is relatively large implying that the actual influence of altering this threshold for the matrix-element matching with the parton-shower cannot be determined accurately enough using the currently available samples. A significant gain could possibly be achieved for this systematic effect by for example using NLO generators and using constraints from data. In addition, the applied upwards and downwards scaling with a factor 2 is also overconservative and should be evaluated.
The second systematic, associated with the jet energy scale uncertainty, can be reduced partially by tightening the event-selection criteria for the transverse momentum p_{T} since these uncertainties become more significant for lower p_{T} values. In this thesis the applied p_{T} requirement is rather loose since all jets with p_{T} above 30 GeV have been considered.
The remainder of the systematic uncertainties have been determined in a rather conservative manner. Hence once the two leading systematics would be reduced, some minor improvements can still be obtained by looking into the details of the different considered systematics and, if possible, applying a less conservative approach. In addition, the strong correlation between the different systematic samples used to determine the effect of the background composition, the b- and mis-tagging efficiency and the pile-up reweighting implies their associated statistical uncertainty is in reality much lower than what is mentioned in Table 6.2. This can significantly be improved in case a dedicated resampling technique [112] is used to determine the statistical uncertainty on each of these systematic effects.
The statistical uncertainty obtained for this measurement can be improved in a rather straightforward manner by including the electron channel of the semi-leptonic top-quark pair decay. However this has not been done since including this channel would signficantly complicate the Matrix Element calculations. The worse energy resolution of electrons implies that the corresponding transfer function can not be described by a simplified Dirac-δ and would thus require a third transfer function to be determined. Hence due to the limited gain this would imply for the overall result, wich is nevertheless dominated by the systematic uncertainty, this has not been considered.

The full procedure developed in this thesis measure the right-handed tensor coupling has been constructed in general and is therefore perfectly applicable for determining the remaining anomalous coupling coefficients of the Wtb interaction vertex. However, these coupling coefficients are known to be less sensitive to for example the W-boson helicity fractions. Hence the measurements are expected to be less precise.

Bibliography

[1] CMS Collaboration, "Measurement of the W-boson helicity in top-quark decays from $t\bar{t}$ production in lepton+jets events in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **10** (2013) 167, doi:10.1007/JHEP10(2013)167, arXiv:1308.3879.

[2] F. Mandl and G. Shaw, "QUANTUM FIELD THEORY". 1985.

[3] M. E. Peskin and D. V. Schroeder, "An Introduction to quantum field theory". 1995.

[4] E. A. Paschos, "Electroweak theory". Cambridge University Press, 2010.

[5] Particle Data Group Collaboration, "Review of Particle Physics", *Chin. Phys.* **C38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.

[6] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett.* **B716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[7] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett.* **B716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[8] ATLAS, CMS Collaboration, "Combined Measurement of the Higgs Boson Mass in *pp* Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments", *Phys. Rev. Lett.* **114** (2015) 191803, doi:10.1103/PhysRevLett.114.191803, arXiv:1503.07589.

[9] CMS Collaboration, "Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV", *Phys. Rev.* **D92** (2015), no. 1, 012004, doi:10.1103/PhysRevD.92.012004, arXiv:1411.3441.

[10] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.* **13** (1964) 321–323, doi:10.1103/PhysRevLett.13.321.

[11] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.* **13** (1964) 508–509, doi:10.1103/PhysRevLett.13.508.

[12] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles", *Phys. Rev. Lett.* **13** (1964) 585–587, doi:10.1103/PhysRevLett.13.585.

[13] N. Cabibbo, "Unitary Symmetry and Leptonic Decays", *Phys. Rev. Lett.* **10** (Jun, 1963) 531–533, doi:10.1103/PhysRevLett.10.531.

[14] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction", *Prog. Theor. Phys.* **49** (1973) 652–657, doi:10.1143/PTP.49.652.

[15] Planck Collaboration, "Planck 2015 results. I. Overview of products and scientific results", arXiv:1502.01582.

[16] SNO Collaboration, "The Sudbury neutrino observatory", *Nucl. Instrum. Meth.* **A449** (2000) 172–207, doi:10.1016/S0168-9002(99)01469-2, arXiv:nucl-ex/9910016.

[17] "Super Kamiokande". http://www-sk.icrr.u-tokyo.ac.jp/sk/index-e.html.

[18] B. Pontecorvo, "Inverse beta processes and nonconservation of lepton charge", *Sov. Phys. JETP* **7** (1958) 172–173. [Zh. Eksp. Teor. Fiz.34,247(1957)].

[19] Z. Maki, M. Nakagawa, and S. Sakata, "Remarks on the unified model of elementary particles", *Prog. Theor. Phys.* **28** (1962) 870–880, doi:10.1143/PTP.28.870.

[20] "The Tevatron Collider". http://www-bdnew.fnal.gov/tevatron/.

[21] "The CDF Experiment". http://www-cdf.fnal.gov/.

[22] "The DØ Experiment". http://www-d0.fnal.gov/.

[23] "CERN". http://home.web.cern.ch/.

[24] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$", *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

[25] M. Czakon et al., "Constraints on the gluon PDF from top quark pair production at hadron colliders", *JHEP* **07** (2013) 167, doi:10.1007/JHEP07(2013)167, arXiv:1303.7215.

[26] CDF, D0 Collaboration, "Combination of measurements of the top-quark pair production cross section from the Tevatron Collider", *Phys. Rev.* **D89** (2014), no. 7, 072001, doi:10.1103/PhysRevD.89.072001, arXiv:1309.7570.

[27] CMS Collaboration, "Legacy Run-I measurement of the ttbar cross section", *PAS-TOP-13-004* (Should go towards paper in 2015).

[28] CMS Collaboration Collaboration, "Combination of ATLAS and CMS top quark pair cross section measurements in the emu final state using proton-proton collisions at 8 TeV", Technical Report CMS-PAS-TOP-14-016, CERN, Geneva, (2014).

[29] "Measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s} = 13$ TeV using $e\mu$ events with b-tagged jets", Technical Report ATLAS-CONF-2015-033, CERN, Geneva, (Jul, 2015).

[30] CDF Collaboration, "First Observation of Electroweak Single Top Quark Production", *Phys. Rev. Lett.* **103** (2009) 092002, doi:10.1103/PhysRevLett.103.092002, arXiv:0903.0885.

[31] D0 Collaboration, "Observation of Single Top Quark Production", *Phys. Rev. Lett.* **103** (2009) 092001, doi:10.1103/PhysRevLett.103.092001, arXiv:0903.0850.

[32] ATLAS Collaboration, "Evidence for single top-quark production in the s-channel in proton-proton collisions at $\sqrt{s}$ =8 TeV with the ATLAS detector using the Matrix Element Method", arXiv:1511.05980.

[33] CMS Collaboration, "Measurement of the t-channel single-top-quark production cross section and of the $| V_{tb} |$ CKM matrix element in pp collisions at $\sqrt{s} = 8\,\text{TeV}$", *JHEP* **06** (2014) 090, doi:10.1007/JHEP06(2014)090, arXiv:1403.7366.

[34] CMS Collaboration, "Search for Flavor-Changing Neutral Currents in Top-Quark Decays $t \to Zq$ in pp Collisions at $\sqrt{s} = 8\,\text{TeV}$", *Phys. Rev. Lett.* **112** (2014), no. 17, 171802, doi:10.1103/PhysRevLett.112.171802, arXiv:1312.4194.

[35] CMS Collaboration Collaboration, "Search for top quark decays t $\to$ qH with H $\to \gamma\gamma$ in pp collisions at $\sqrt{s} = 8\,\text{TeV}$", Technical Report CMS-PAS-TOP-14-019, CERN, Geneva, (2015).

[36] CMS Collaboration, "Search for baryon number violation in top-quark decays", *Phys. Lett.* **B731** (2014) 173–196, doi:10.1016/j.physletb.2014.02.033, arXiv:1310.1618.

[37] Gfitter Group Collaboration, "The global electroweak fit at NNLO and prospects for the LHC and ILC", *Eur. Phys. J.* **C74** (2014) 3046, doi:10.1140/epjc/s10052-014-3046-5, arXiv:1407.3792.

[38] CMS Collaboration, "Measurement of the top quark mass using proton-proton data at $\sqrt{s} = 7$ and 8 TeV", arXiv:1509.04044.

[39] ATLAS Collaboration, "Measurement of Spin Correlation in Top-Antitop Quark Events and Search for Top Squark Pair Production in pp Collisions at $\sqrt{s} = 8\,\text{TeV}$ Using the ATLAS Detector", *Phys. Rev. Lett.* **114** (2015), no. 14, 142001, doi:10.1103/PhysRevLett.114.142001, arXiv:1412.4742.

[40] CMS Collaboration Collaboration, "Measurement of the W-boson helicity in top decays from ttbar production in lepton+jets events at the LHC at sqrt(s)=8 TeV", Technical Report CMS-PAS-TOP-13-008, CERN, Geneva, (2013).

[41] K. J. F. Gaemers and G. J. Gounaris, "Polarization Amplitudes for e+ e- —> W+ W- and e+ e- —> Z Z", *Z. Phys.* **C1** (1979) 259, doi:10.1007/BF01440226.

[42] C.-R. Chen, F. Larios, and C. P. Yuan, "General analysis of single top production and W helicity in top decay", *Phys. Lett.* **B631** (2005) 126–132, doi:10.1016/j.physletb.2005.10.002, arXiv:hep-ph/0503040.

[43] Q.-H. Cao et al., "A General Analysis of Wtb anomalous Couplings", arXiv:1504.03785.

[44] T. D. Lee and C.-N. Yang, "Question of Parity Conservation in Weak Interactions", *Phys. Rev.* **104** (1956) 254–258, doi:10.1103/PhysRev.104.254.

[45] C. S. Wu et al., "Experimental Test of Parity Conservation in Beta Decay", *Phys. Rev.* **105** (1957) 1413–1414, doi:10.1103/PhysRev.105.1413.

[46] ATLAS Collaboration, "Measurement of the W boson polarization in top quark decays with the ATLAS detector", *JHEP* **06** (2012) 088, doi:10.1007/JHEP06(2012)088, arXiv:1205.2484.

[47] CMS Collaboration, "Measurement of the W boson helicity in events with a single reconstructed top quark in pp collisions at $\sqrt{s} = 8$ TeV", *JHEP* **01** (2015) 053, doi:10.1007/JHEP01(2015)053, arXiv:1410.1154.

[48] J. A. Aguilar-Saavedra et al., "Probing anomalous Wtb couplings in top pair decays", *Eur. Phys. J.* **C50** (2007) 519–533, doi:10.1140/epjc/s10052-007-0289-4, arXiv:hep-ph/0605190.

[49] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.

[50] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[51] ALICE Collaboration, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002, doi:10.1088/1748-0221/3/08/S08002.

[52] LHCb Collaboration, "The LHCb Detector at the LHC", *JINST* **3** (2008) S08005, doi:10.1088/1748-0221/3/08/S08005.

[53] J.-L. Caron, "Cross section of LHC dipole.. Dipole LHC: coupe transversale.", (May, 1998). AC Collection. Legacy of AC. Pictures from 1992 to 2002.

[54] "CERN accelerator complex". http://www.stfc.ac.uk/research/particle-physics-and-particle-astrophysics/particle-physics/large-hadron-collider/cern-accelerator-complex/.

[55] TOTEM Collaboration, "The TOTEM experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08007, doi:10.1088/1748-0221/3/08/S08007.

[56] LHCf Collaboration, "The LHCf detector at the CERN Large Hadron Collider", *JINST* **3** (2008) S08006, doi:10.1088/1748-0221/3/08/S08006.

[57] MoEDAL Collaboration, "The Physics Programme Of The MoEDAL Experiment At The LHC", *Int. J. Mod. Phys.* **A29** (2014) 1430050, doi:10.1142/S0217751X14300506, arXiv:1405.7662.

[58] "CMS Luminosity: Public results". https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[59] CMS Collaboration, "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software". Technical Design Report CMS. CERN, Geneva, 2006.

[60] "CMS detector design". http://cms.web.cern.ch/news/cms-detector-design.

[61] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014), no. 10, P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[62] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl. Instrum. Meth.* **A262** (1987) 444–450, doi:10.1016/0168-9002(87)90887-4.

[63] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, "Adaptive vertex fitting", *J. Phys.* **G34** (2007) N343, doi:10.1088/0954-3899/34/12/N01.

[64] CMS Collaboration, "Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV", *JINST* **8** (2013) P09009, doi:10.1088/1748-0221/8/09/P09009, arXiv:1306.2016.

[65] CMS Collaboration, "Identification and Filtering of Uncharacteristic Noise in the CMS Hadron Calorimeter", *JINST* **5** (2010) T03014, doi:10.1088/1748-0221/5/03/T03014, arXiv:0911.4881.

[66] CMS Collaboration, "The performance of the CMS muon detector in proton-proton collisions at sqrt(s) = 7 TeV at the LHC", *JINST* **8** (2013) P11002, doi:10.1088/1748-0221/8/11/P11002, arXiv:1306.6905.

[67] P. Z. Skands, "QCD for Collider Physics", in *High-energy Physics. Proceedings, 18th European School, ESHEP 2010, Raseborg, Finland, 20 Jun. - 3 Jul., 2010.* 2011. arXiv:1104.2863.

[68] B. R. Webber, "Fragmentation and hadronization", *Int. J. Mod. Phys.* **A15S1** (2000) 577–606, doi:10.1142/S0217751X00005334, arXiv:hep-ph/9912292. [,577(1999)].

[69] T. Sjostrand, "Monte Carlo Tools", in *Proceedings, 65th Scottish Universities Summer School in Physics: LHC Physics (SUSSP65)*, pp. 309–339. 2009. arXiv:0911.5286.

[70] A. Buckley et al., "General-purpose event generators for LHC physics", *Phys. Rept.* **504** (2011) 145–233, doi:10.1016/j.physrep.2011.03.005, arXiv:1101.2599.

[71] G. P. Salam, "Elements of QCD for hadron colliders", in *High-energy physics. Proceedings, 17th European School, ESHEP 2009, Bautzen, Germany, June 14-27, 2009*. 2010. arXiv:1011.5131.

[72] M. A. Dobbs et al., "Les Houches guidebook to Monte Carlo generators for hadron collider physics", in *Physics at TeV colliders. Proceedings, Workshop, Les Houches, France, May 26-June 3, 2003*, pp. 411–459. 2004. arXiv:hep-ph/0403045.

[73] G. Corcella et al., "HERWIG 6.5 release note", arXiv:hep-ph/0210213.

[74] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[75] J. Alwall et al., "MadGraph 5 : Going Beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[76] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[77] S. Frixione and B. R. Webber, "The MC and NLO 3.4 Event Generator", arXiv:0812.0770.

[78] S. Höche, "Introduction to parton-shower event generators", in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014) Boulder, Colorado, June 2-27, 2014*. 2014. arXiv:1411.4085.

[79] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", *Nucl. Phys.* **B126** (1977) 298, doi:10.1016/0550-3213(77)90384-4.

[80] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory", *Sov. J. Nucl. Phys.* **15** (1972) 438–450. [Yad. Fiz.15,781(1972)].

[81] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.", *Sov. Phys. JETP* **46** (1977) 641–653. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[82] M. L. Mangano et al., "Matching matrix elements and shower evolution for top-quark production in hadronic collisions", *JHEP* **01** (2007) 013, doi:10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.

[83] B. Andersson, "The Lund model", *Nuclear Physics A* **461** (1987), no. 1, 513 – 520, doi:http://dx.doi.org/10.1016/0375-9474(87)90510-0.

[84] CDF Collaboration, "The Underlying event in hard scattering processes", *eConf* **C010630** (2001) P501, arXiv:hep-ph/0201192.

[85] J. Allison et al., "Geant4 developments and applications", *IEEE Trans. Nucl. Sci.* **53** (2006) 270, doi:10.1109/TNS.2006.869826.

[86] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.

[87] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[88] W. Adam et al., "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC",.

[89] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET",.

[90] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector",.

[91] CMS Collaboration, "Particle-flow commissioning with muons and electrons from J/Psi and W events at 7 TeV",.

[92] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[93] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[94] CMS Collaboration, "Jet Energy Scale performance in 2011",.

[95] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **8** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462.

[96] CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015), no. 02, P02006, doi:10.1088/1748-0221/10/02/P02006, arXiv:1411.0511.

[97] CMS Collaboration, "Single Muon efficiencies in 2012 Data", CMS Detector Performance Summaries CMS-DP-2013-009, (2013).

[98] CMS Collaboration, "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV",.

[99] CMS Collaboration, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015), no. 06, P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.

[100] C. D. White et al., "Isolating Wt production at the LHC", *JHEP* **11** (2009) 074, doi:10.1088/1126-6708/2009/11/074, arXiv:0908.0631.

[101] CMS Collaboration Collaboration, "Performance of b tagging at sqrt(s)=8 TeV in multijet, ttbar and boosted topology events",.

[102] CDF Collaboration, "Precise measurement of the top quark mass in the lepton+jets topology at CDF II", *Phys. Rev. Lett.* **99** (2007) 182002, doi:10.1103/PhysRevLett.99.182002, arXiv:hep-ex/0703045.

[103] CDF Collaboration, "Top Quark Mass Measurement in the Lepton plus Jets Channel Using a Modified Matrix Element Method", *Phys. Rev.* **D79** (2009) 072001, doi:10.1103/PhysRevD.79.072001, arXiv:0812.4469.

[104] D0 Collaboration, "Measurement of the top quark mass in the lepton+jets final state with the matrix element method", *Phys. Rev.* **D74** (2006) 092005, doi:10.1103/PhysRevD.74.092005, arXiv:hep-ex/0609053.

[105] CMS Collaboration, "Search for a Standard Model Higgs Boson Produced in Association with a Top-Quark Pair and Decaying to Bottom Quarks Using a Matrix Element Method", *Eur. Phys. J.* **C75** (2015), no. 6, 251, doi:10.1140/epjc/s10052-015-3454-1, arXiv:1502.02485.

[106] CMS Collaboration, "Search for ttH production using the Matrix Element Method",.

[107] ATLAS Collaboration, "Evidence for single top-quark production in the s-channel in proton-proton collisions at $\sqrt{s}$ =8 TeV with the ATLAS detector using the Matrix Element Method", arXiv:1511.05980.

[108] CMS Collaboration, "Measurement of spin correlations in t-tbar production using the matrix element method in the muon+jets final state in pp collisions at sqrt(s) = 8 TeV", arXiv:1511.06170.

[109] P. Artoisenet et al., "Automation of the matrix element reweighting method", *JHEP* **12** (2010) 068, doi:10.1007/JHEP12(2010)068, arXiv:1007.3300.

[110] G. P. Lepage, "VEGAS: AN ADAPTIVE MULTIDIMENSIONAL INTEGRATION PROGRAM", *Cornell preprint CLNS 80-447!TEST* (1980).

[111] A. Alloul et al., "FeynRules 2.0 - A complete toolbox for tree-level phenomenology", *Comput. Phys. Commun.* **185** (2014) 2250–2300, doi:10.1016/j.cpc.2014.04.012, arXiv:1310.1921.

[112] R. G. Miller, "The jackknife – a review", *Biometrika* **61** (1974) 1.

Summary

The Large Hadron Collider is the first particle accelerator which is capable of providing an abundant amount of top-quark pair events at record-breaking centre-of-mass energies, allowing for a thorough investigation of physics at the TeV scale. The data collected by the CMS detector during the 2012 run of the LHC is considered in this thesis and the event signature of interest corresponds to the semi-muonic decay of top-quark pair events. These events have been used to determine whether the top-quark decay vertex is described purely by a left-handed vector coupling, as predicted by the Standard Model of elementary particle physics, or whether additional couplings occur due to new-physics phenomena. In this thesis the focus was on the estimation of the right-handed tensor coupling, denoted as g_R.

The study discussed in this thesis is the first direct measurement of the right-handed anomalous tensor coupling of the top-quark decay vertex. It has been performed using a Matrix Element method, which is capable of extracting potentially the best estimate of any theoretical parameter from a sample of experimental events. This advanced analysis technique evaluates each event and calculate a corresponding event probability using a dedicated phase-space integration. It has been opted to apply a stringent event selection by for instance exploiting the specific characteristics of the b-quark jets in order to improve the reconstruction efficiency. Two such jets are expected in top-quark pair events since the top-quark decays almost exclusively into a b-quark and a W-boson.

With $19.6\,\mathrm{fb}^{-1}$ of collision data taken at a centre-of-mass energy of 8 TeV a right-handed tensor coupling value of

$$g_R = -0.0071 \pm 0.0083\,(\text{stat.}) \pm 0.0137\,(\text{syst.}) = -0.0071 \pm 0.0160$$

was measured, which is consistent with the prediction of the Standard Model ($g_R = 0$). Comparing the obtained result with the currently existing exclusion limits for this right-handed tensor coupling also indicates an excellent agreement. Hence it can be concluded that the Matrix Element method has been successfully applied on the reconstructed collision events recorded by the CMS experiment and has resulted in a rather accurate and first direct measurement of this anomalous coupling coefficient g_R.

Samenvatting

Meting van de anomale koppelingen in de Wtb vertex gebruik makende van de Matrix Element methode aan de LHC

De Large Hadron Collider is de eerste deeltjesversneller ter wereld die in staat is om zo'n enorme massamiddelpuntsenergie te op te wekken. Hierdoor is het mogelijk om een massieve hoeveelheid van top-quark paren te produceren tijdens de proton-proton botsingen. Deze botsingen zullen gebruikt worden voor een nauwkeurig onderzoek van fysica bij de TeV energie-schaal. In deze thesis zal gebruik gemaakt worden van de dataset die in 2012 werd gededecteerd door het CMS experiment. Uit deze dataset zal specifiek gekeken worden naar botsingen waarbij twee top-quarks geproduceerd werden in het μ+jets kanaal ($t\bar{t} \to b\bar{b}W^+W^- \to b\bar{b}q\bar{q}\mu\nu_\mu$).
Deze botsingen zullen gebruikt worden om na te gaan of de eigenschappen van de Wtb vertex, die het verval van top-quarks beschrijft, overeenkomt met de voorspellingen in het Standaard Model. Afwijkingen van deze voorspelling zou kunnen betekenen dat dit verval beïnvloedt wordt door extra koppelingen veroorzaakt door nieuwe fysica fenomenen. In de analyse besproken in deze thesis zal specifiek nagegaan worden of de rechtshandige tensor koppeling, genoteerd als g_R, daadwerkelijk gelijk is aan 0. Deze meting is de eerste rechtstreekse meting van deze fysische parameter en zal uitgevoerd worden gebruik makende van een Matrix Element methode. Dit is een geavanceerde analyse techniek die in staat is om de meest optimale theoretische parameter te bepalen door de kinematische informatie van experimentele botsingen te bestuderen. De kracht van deze techniek schuilt in het feit dat elk event afzonderlijk bestudeerd wordt en aanleiding geeft tot een individuele waarschijnlijkheid.
Aangezien deze Matrix Element methode een lange berekeningstijd nodig heeft is ervoor gekozen in deze analyse om het aantal te bestuderen botsingen tot een minimum te houden. Dit betekent dat er een zeer strikte botsingsselectie zal toegepast worden om te verzekeren dat de geanalyseerde botsingen zo goed mogelijk overeenkomen met top-quark paren in het μ+jets kanaal. Een zeer efficiente methode om dit soort botsingen te onderscheiden van de enorme hoeveelheid van andere botsingen die geproduceerd worden in de LHC is door de specifieke eigenschappen van b-quark jets uit te buiten.
De meting van de rechtshandige tensor koppeling g_R, uitgevoerd met $19.6\,\text{fb}^{-1}$ aan botsingsgegevens bij een energie van 8 TeV, gaf een waarde van:

$$g_R = -0.0071 \pm 0.0083\,(\text{stat.}) \pm 0.0137\,(\text{syst.}) = -0.0071 \pm 0.0160$$

Dit resultaat komt uitstekend overeen met zowel de voorspelling in het Standaard Model ($g_R = 0$) als met de huidige limieten voor deze koppeling coefficient.

Acknowledgements

All good things must come to an end; a saying that definitely holds for a PhD. My four years at the IIHE was a wonderful experience in the company of amazing colleagues, which were always in for both a good laugh and a serious physics discussion.

First of all I would like to thank my promotor Jorgen D'Hondt for providing me the opportunity to conduct a PhD, and for always taking the time to discuss and provide the necessary insight to tackle the various problems encountered during my research. In addition, I would like to thank the members of my jury to go through the entire manuscript and provide both conceptual and textual comments, this truly improved the quality of my manuscript. I also want to thank Pierre Artoisenet and Olivier Mattelaer for explaining me all the technicalities of the Matrix Element method.

I started my PhD in a small, partially hidden office; more a small dungeon; but which had the benefit that I could switch off the heating during the winter months. Even after four years, people who had to visit my office still feel the cold air running over their back ... After less than a year this advantage was traded for an even better one: a big office filled with many nice colleagues. Over the last three years the faces have changed continuously but the good atmosphere never left. Thank you Petra, Quentin, Kevin, Lieselotte, Isis, Isabelle, Shimae, Nadir, Jarne, Leonidas, Seth, Dominic, Denys and everyone who has already left the lab for all the good times we had in our sunny office! A special dedication goes out to the amazing Elementary FCNC indoor soccer team, it was a true honour to be part of the best (experimental physics) soccer team ever existing (at the VUB). When I was not at the office, I truly enjoyed teaching the lab sessions together with Jan. It always felt like a day off when we were chatting and laughing in the lab.

This pleasant atmosphere continues throughout the entire IIHE where everyone does its utter best to help each other. In my case I want to explicitly thank the technical support of the computing admins for letting me fully exploit the capacities of the cluster, otherwise I would never had been able to finish this analysis. Clearly Marleen should not be forgotten: without your administrative support the IIHE would just not be able to function properly!

Last but not least I want to thank my family and friends for their continuous support. First of all my parents for always encouraging me to study Physics and to follow my dreams to become a PhD researcher and whatever comes next. Also my friends to provide the necessary distraction when needed during the less pleasant times of a PhD. Finally I would like to thank Stijn for the textual corrections, but more importantly for the unconditional support and encouragements given during the difficult moments of my PhD.