

Vrije Universiteit Brussel



Faculteit Wetenschappen en Bio-ingenieurswetenschappen  
Departement Natuurkunde

---

# **Direct measurement of the top quark decay width in the muon+jets channel using the CMS experiment at the LHC**

---

**Lieselotte Moreels**

Promotor

Prof. Dr. Petra Van Mulders

Co-promotor

Prof. Dr. Jorgen D'Hondt

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

April 2018

Doctoral examination commission:

Prof. Dr. Nick van Eijndhoven (Vrije Universiteit Brussel, *chair*)

Prof. Dr. Alberto Mariotti (Vrije Universiteit Brussel, *secretary*)

Prof. Dr. Petra Van Mulders (Vrije Universiteit Brussel, *supervisor*)

Prof. Dr. Jorgen D'Hondt (Vrije Universtiteit Brussel, *advisor*)

Prof. Dr. Dirk Ryckbosch (Universiteit Gent)

Dr. Ivan Marchesini (Vrije Universiteit Brussel)

Dr. Rebeca Gonzalez Suarez (University of Nebraska, USA)

Prof. Dr. Philippe Claeys (Vrije Universiteit Brussel)

Cover illustration: Fisheye view of the CMS detector. The barrel (front) and one of the endcaps (back) are connected by the beam pipe. Picture taken from CMS-doc-13419-v1.

Cover kindly designed by Isis Van Parijs.

ISBN: 978 94 923 1271 6

NUR: 924, 926

© 2018 Lieselotte Moreels

All rights reserved. No parts of this book may be reproduced or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the author.

# Contents

---

<b>Introduction</b>	<b>1</b>
<b>1 The Particle Universe</b>	<b>3</b>
1.1 The Standard Model . . . . .	4
1.1.1 Particles and forces . . . . .	4
1.1.2 Mathematical framework . . . . .	7
1.1.3 Open questions . . . . .	10
1.2 The Top Quark Sector . . . . .	12
1.2.1 Production and decay . . . . .	13
1.2.2 Top quark mass . . . . .	15
1.2.3 Top quark decay width . . . . .	15
<b>2 The CMS Experiment and the LHC</b>	<b>23</b>
2.1 The Large Hadron Collider . . . . .	23
2.2 The CMS Experiment . . . . .	26
2.2.1 The silicon tracker . . . . .	28
2.2.2 The electromagnetic calorimeter . . . . .	29
2.2.3 The hadron calorimeter . . . . .	31
2.2.4 The muon detectors . . . . .	33
2.2.5 Trigger and data acquisition . . . . .	34
2.2.6 The CMS experiment during the 2016 proton run . . . . .	38
<b>3 Event Generation and Simulation</b>	<b>41</b>
3.1 Simulation of Events at Proton Colliders . . . . .	41
3.1.1 The hard-scattering process . . . . .	42
3.1.2 Parton shower . . . . .	44
3.1.3 Hadronisation and decay . . . . .	46
3.1.4 Underlying event . . . . .	47
3.1.5 Simulation of the CMS detector . . . . .	49
3.2 Simulated Events for Top Quark Studies . . . . .	49
3.3 Reweighting Procedure for Simulated Samples . . . . .	53
3.3.1 Reweighting of the simulated top quark width . . . . .	53
3.3.2 Reweighting of the simulated top quark mass . . . . .	56

<b>4</b>	<b>Reconstruction of Proton Collisions</b>	<b>61</b>
4.1	Track Reconstruction . . . . .	62
4.2	Calorimetry . . . . .	64
4.3	Reconstruction and Identification of Particles and Jets . . . . .	64
4.3.1	Muon reconstruction . . . . .	65
4.3.2	Electron reconstruction . . . . .	69
4.3.3	Jet reconstruction . . . . .	70
4.3.4	Identification of jets originating from b quarks . . . . .	74
4.3.5	Missing transverse energy reconstruction . . . . .	76
<b>5</b>	<b>Event Selection &amp; Reconstruction</b>	<b>79</b>
5.1	Trigger and Basic Event Requirements . . . . .	79
5.2	Event Selection . . . . .	80
5.3	Event Reconstruction . . . . .	86
5.3.1	Reconstruction of generated top quark events . . . . .	86
5.3.1.1	Jet-parton matching . . . . .	87
5.3.1.2	Resolution functions . . . . .	87
5.3.2	Reconstruction of the top quark pair system . . . . .	89
5.3.2.1	Categorisation of events . . . . .	91
5.3.2.2	Kinematic fit . . . . .	91
5.3.2.3	Further selection requirements . . . . .	95
<b>6</b>	<b>Measurement of the Top Quark Width</b>	<b>101</b>
6.1	Measurement Procedure using One Variable . . . . .	101
6.1.1	Construction of a likelihood function . . . . .	102
6.1.2	Evaluation of the likelihood function . . . . .	107
6.1.3	Calibration curve . . . . .	108
6.1.4	Pull . . . . .	110
6.2	Combination of Several Sensitive Variables . . . . .	111
6.2.1	Template construction for another variable . . . . .	111
6.2.2	Combination of both variables . . . . .	113
6.2.2.1	Calibration curve . . . . .	114
6.2.2.2	Pull . . . . .	114
6.3	Systematic Uncertainties . . . . .	116
6.3.1	Theoretical uncertainties . . . . .	116
6.3.2	Experimental uncertainties . . . . .	118
6.3.3	Summary of systematic uncertainties . . . . .	120
6.4	Measurement of the Top Quark Decay Width . . . . .	123
<b>7</b>	<b>Conclusions &amp; Prospects</b>	<b>125</b>
	<b>Summary</b>	<b>131</b>
	<b>Samenvatting</b>	<b>133</b>
	<b>Bibliography</b>	<b>135</b>

---

<b>Appendix A Breakdown of Systematic Uncertainties</b>	<b>151</b>
<b>Appendix B Effect of the Systematic Uncertainties on the Sensitive Variable Distributions</b>	<b>157</b>
<b>Author's Contributions</b>	<b>161</b>
<b>List of Abbreviations</b>	<b>163</b>
<b>Glossary</b>	<b>167</b>
<b>Acknowledgements</b>	<b>169</b>



# Introduction

---

Complete knowledge about everything in the universe is the dream of physicists and evil masterminds trying to take over the world. The standard model (SM) of particle physics embodies much of our current understanding of what matter consists of and how it interacts. Yet, some puzzles remain. It is unclear how gravity can be incorporated into the SM and none of the known particles can account for the dark matter that is observed in the universe. Being the heaviest particle in the SM, the top quark is expected to be most sensitive to beyond the SM effects, such as interactions with as yet unknown particles. Besides direct searches for new physics, precise measurements of the properties of SM particles and interactions can be compared to their predictions, thus favouring or excluding certain beyond the SM theories. Further, precision measurements allow to perform consistency checks of the SM itself. Using the measurements of certain parameters as input, the values of other quantities can be predicted and compared to their direct measurements. It is therefore important that the uncertainties on all measurements are as small as possible, since this leads to more stringent tests of the SM.

This thesis investigates the decay properties of the top quark. In the SM the top quark decays almost exclusively into a W boson and a b quark. The probability for this process to happen is reflected in the top quark decay width, which is predicted to have a value around  $1.33 \text{ GeV}$ . If the top quark is able to decay into other particles as well, as is possible in several extensions of the SM, the top quark decay width will be larger than the SM prediction. This is investigated by performing a direct measurement of the top quark decay width using proton collisions produced by the Large Hadron Collider at a centre-of-mass energy of  $13 \text{ TeV}$ . The data were recorded by the CMS experiment in 2016 and correspond to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ .

At first, a concise overview of the standard model is presented in Chapter 1. The production and decay of the top quark is described in more detail and the current status of the top quark mass and decay width measurements is given. Top quarks are typically studied by producing them in particle colliders. Chapter 2 describes the Large Hadron Collider and the Compact Muon Solenoid experiment, which were used to produce and record the top quark events analysed in this thesis.

In order to understand how the observations in a particle detector correspond to the theoretical interactions between colliding particles, collision events are simulated. A step-by-step overview is given in Chapter 3. In addition, the simulated samples used in this thesis are summarised and a reweighting technique is presented to acquire simulated top quark pair events with a distribution characterised by an arbitrary top quark mass

and decay width.

Chapter 4 expands on how electronic detector signals are processed and combined to reconstruct particles. These are used in Chapter 5 to set up a basic event selection and reconstruct higher-order objects, such as the top quark. Based on the reconstructed event topology, additional requirements are posed on the selected events.

The measurement procedure itself is based on a maximum likelihood technique and is described in Chapter 6. It is set up using simulated events only, so as not to let the result of the measurement influence the procedure. Probability density functions are constructed using the distributions of variables that are sensitive to changes in the top quark decay width. These changes are simulated using the reweighting technique introduced in Chapter 3. Further, the measurement procedure is calibrated and the systematic uncertainties affecting the measurement are discussed. In addition, the likelihood functions of several sensitive variables are combined to get a more precise measurement.

The results of the two top quark decay width measurements presented in this thesis are summarised in Chapter 7 and possible improvements to obtain an even more precise measurement are discussed.

As is common in high-energy physics, particle masses and momenta will be expressed in  $\text{GeV}/c^2$  and  $\text{GeV}/c$ , resp., where  $1 \text{ eV} \simeq 1.6 \times 10^{-19} \text{ J}$  is equal to the energy an electron acquires when it is accelerated over a potential difference of 1 V. Further, natural units are used, i.e.  $\hbar = 1 = c$ , which allows to express both mass and momentum in units of GeV.

# 1

## The Particle Universe

---

All throughout history people have been intrigued by nature. Beholding its awe-inspiring beauty spawned a great number of questions, such as how life originated and what is the meaning of life. These questions are beyond the scope of this thesis. Other questions enquired about the fundamental building blocks that everything that can be seen consists of. Over the centuries, many theories have been put forward, evolving from a simple classification such as the elements water, earth, wind, and fire to the standard model of particle physics. Only the latter will be discussed in this chapter.

Deconstructing matter, the concept of molecules and atoms is well-known to most. Delving deeper into the atomic structure, the nucleus, only a fraction of the size of the atom itself, is surrounded by a cloud of electrons. The nucleus has further substructure and the amount of nucleons (protons and neutrons) it contains determines the chemical properties of the atom. Even smaller particles called quarks constitute the nucleons. This is visualised in Figure 1.1 (not to scale). Up till now there is no indication that quarks and/or electrons have further substructure. They are therefore considered to be elementary particles.

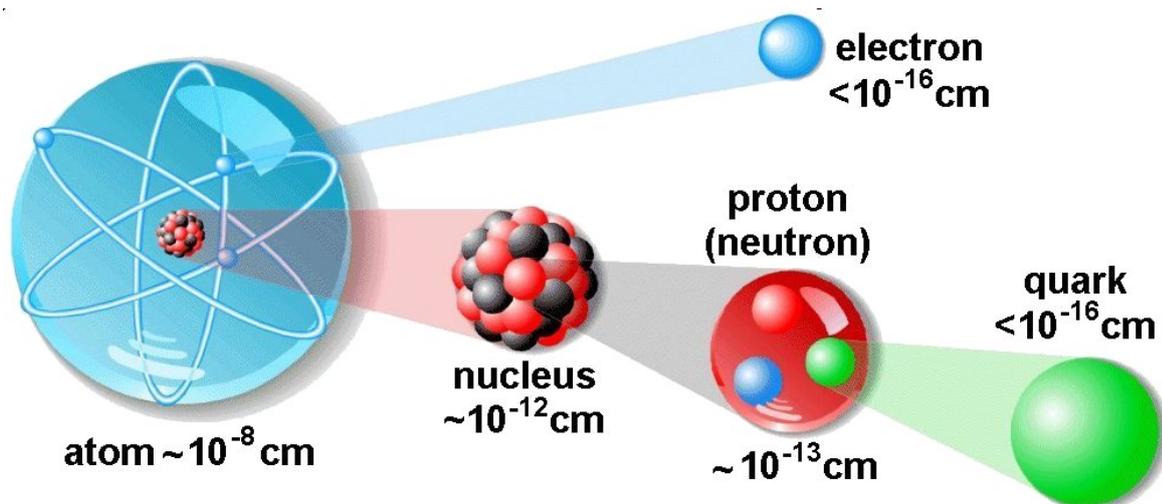


Figure 1.1: Impression of the inner structure of an atom, including an indication of the size of each component [1].

Below, Section 1.1 discusses the main aspects of the standard model of particle physics, which describes the collection of elementary particles and their interactions. One of these particles, the top quark, is emphasised in Section 1.2.

## 1.1 The Standard Model

Many of the ideas outlined below are described in [2–5].

### 1.1.1 Particles and forces

As in Mendeleev’s table for chemical elements, the elementary particles can be categorised according to their properties. In addition to energy, momentum, mass, and electrical charge, they also have a property called spin. For a macroscopic particle, such as the earth, this can be seen as a rotation around an internal axis. An elementary particle, however, is so small that it is perceived as a point particle. The term spin is therefore used to describe its intrinsic angular momentum, as opposed to the angular momentum that describes motion in space. Being an intrinsic property, the spin has a fixed value for a certain particle type and does not change when the particle is perturbed by the outside world. Due to the quantum mechanical rules that apply to small particles, only integer and half-integer spin values are allowed [6]. Particles with integer spin are called bosons and particles with half-integer spin fermions. The fermions of the standard model (SM) are summarised in Table 1.1. They are subdivided into quarks and leptons. There are six types or ‘flavours’ of quarks and six types of leptons. Each vertical group has the same electrical charge, which is indicated at the bottom of the table as a function of the elementary charge  $e \simeq 1.6 \times 10^{-19} \text{ C}$  [7]. Further, leptons and quarks are subdivided into three generations, where each generation is heavier than the previous one.

Table 1.1: Overview of the three generations of fermions in the standard model and their electrical charge.

Generation	Quarks				Leptons			
1	u	up	d	down	$e$	electron	$\nu_e$	electron neutrino
2	c	charm	s	strange	$\mu$	muon	$\nu_\mu$	muon neutrino
3	t	top	b	bottom	$\tau$	tau lepton	$\nu_\tau$	tau neutrino
El. charge ( $e$ )	$+2/3$		$-1/3$		$-1$			$0$

In addition to the particles in Table 1.1, each particle also has an antiparticle. Antiparticles have the same properties as their corresponding particle, except for their electrical charge, which is opposite, and other quantum numbers. Antiparticles are indicated with a bar above the particle symbol, e.g. the antiparticle of a quark  $q$  is the antiquark  $\bar{q}$ , except for a charged lepton  $l^-$ , where the sign of the electrical charge is

reversed for the antiparticle  $l^+$ . Throughout this text, the names in Table 1.1 shall be used to denote both particle and antiparticle. E.g. when speaking of electrons, both the electron and the antielectron or positron are considered, unless specified otherwise. Further, the term lepton will often be used to mean charged lepton, whereas neutral leptons will be called neutrinos, irrespective of their type.

The SM only accommodates three out of four fundamental forces, i.e. the electromagnetic force and the strong and weak nuclear forces. Due to difficulties in combining general relativity with quantum mechanics, gravity is not included. Since gravity is about  $10^{40}$  times weaker than the electromagnetic force, its effect on the interactions described by the SM is expected to be negligible [4]. The SM forces are mediated by the exchange of spin-1 bosons. A summary of the SM bosons, including these force-carrying particles, can be found in Table 1.2.

Table 1.2: Overview of the SM bosons, their mass, electrical charge and spin [8].

Boson	Mass (GeV)	El. charge ( $e$ )	Spin
$\gamma$ (photon)	0	0	1
$g$ (gluon)	0	0	1
$W^+, W^-$	$80.385 \pm 0.015$	$\pm 1$	1
$Z$	$91.188 \pm 0.002$	0	1
$H$	$125.09 \pm 0.24$	0	0

The electromagnetic force acts between electrically charged particles and is mediated by the photon. It governs all electrical and magnetic interactions, including chemical reactions. Whereas the electromagnetic force pervades all of space, scaling inversely with the distance squared, the strong and weak nuclear forces have a limited range. The strong force, mediated by the gluon, has a range of about  $\sim 10^{-15}$  m and the range of the weak force is even shorter due to the large mass of its mediators, the  $W$  and  $Z$  bosons [9]. The term  $W$  boson indicates both the  $W^+$  and the  $W^-$  boson.

The strong nuclear force only affects particles that are colour-charged, i.e. quarks. Contrary to the photon, the gluon is also charged under the strong force, so it can interact with itself. There are three colours—called red, green, and blue—and three anticolours. Quarks (antiquarks) carry one colour (anticolour) and, because colour charge is conserved, gluons carry both a colour and an anticolour. Due to colour confinement, only colour-neutral particles are observed in nature. This implies that quarks and gluons are never observed as bare particles. Colour-neutral particles are formed in two ways. A colour can get cancelled by its anticolour or, analogous to red, green, and blue light jointly appearing as white light, a combination of all colours makes for a colour-neutral particle. In the latter case three quarks combine to form a baryon, while in the former two quarks make a meson. Baryons and mesons are also called

hadrons. An example of a hadron is the proton (neutron), which contains two (one) up quark(s) and one (two) down quark(s). This is only a net effect, however, since a gluon can form a quark-antiquark pair and a quark and its antiquark can recombine to form a gluon. The inner structure of a proton thus looks a bit like in Figure 1.2 [10]. Three ‘valence’ quarks determine the properties of the proton, while ‘sea’ quarks are created through pair production before they recombine again.

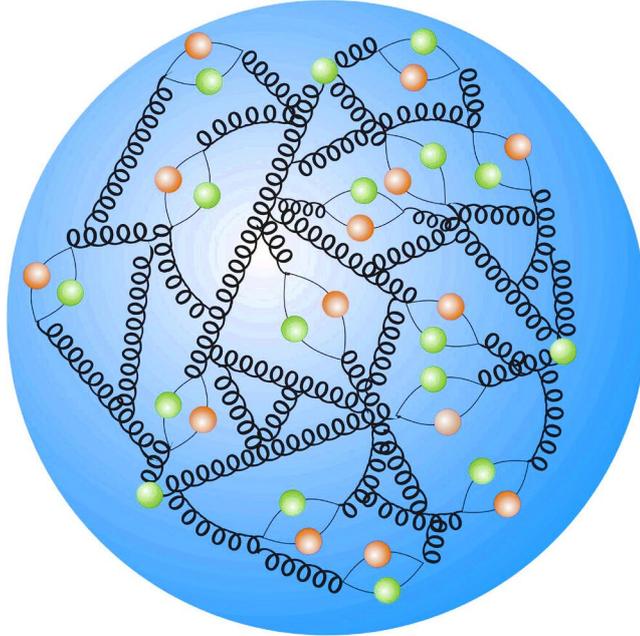


Figure 1.2: Artist’s impression of the inner structure of a proton [10]. Quarks (green) and antiquarks (orange) interact with each other by exchanging gluons (springs), which also interact amongst themselves. There are three more quarks than antiquarks.

The weak nuclear force is responsible for the decay of particles, as is seen in e.g. radioactive processes. Contrary to the other two forces, the weak force acts on all SM fermions. Weak interactions are subdivided into neutral and charged current interactions, depending on whether the neutral Z boson or one of the charged W bosons mediate the process. Neutral current interactions are very similar to electromagnetic interactions mediated by the photon. It is important to note, though, that neutral interactions involving a neutrino can only be mediated by the Z boson. Due to the conservation of charge during the interaction, a W boson will always link two particles that have an electrical charge difference of  $1e$ , such as a charged lepton and a neutrino or an up and a down quark.

The last particle of the SM is the recently discovered H boson [11–13], which is responsible for giving mass to the fermions, the W and Z bosons, and itself through the Brout–Englert–Higgs mechanism [14–16].

An overview of the interactions that are possible in the SM and which particles interact with each other is given in Figure 1.3.

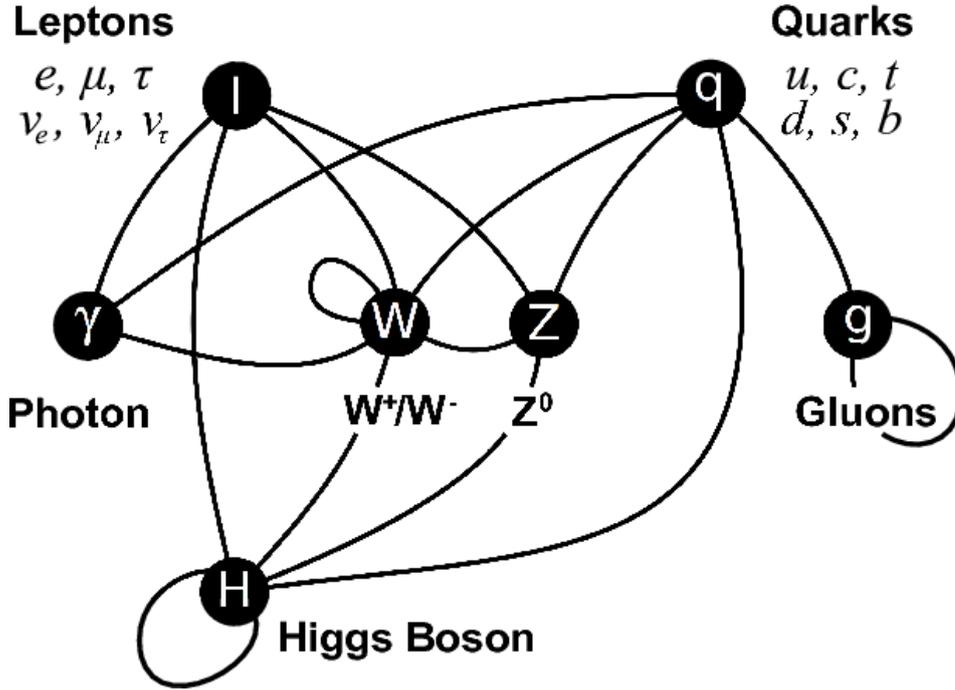


Figure 1.3: Overview of the particle interactions described by the SM [17].

### 1.1.2 Mathematical framework

The world of small, fast-moving particles is described by quantum field theory (QFT), which combines quantum mechanics and special relativity. Therefore, spacetime coordinates  $x = x^\mu = (t, \vec{x})$ , following the Minkowski metric  $g^{\mu\nu}$ , will be used throughout this section. Also the Einstein convention, where a summation over repeated indices is implicit, is adopted.

In QFT, a fermion is described by a Dirac field  $\psi(x)$ . The dynamics of the fermion and its antifermion, whose field corresponds to the hermitian conjugate  $\psi^\dagger(x)$  of that of the fermion, is described by the Lagrangian density

$$\mathcal{L}_\psi = \mathcal{L}(\psi(x), \partial_\mu \psi(x)) = \mathcal{L}(\psi(x), \partial_\mu \psi(x), \psi^\dagger(x), \partial_\mu \psi^\dagger(x)), \quad (1.1)$$

where  $\partial_\mu = \partial/\partial x^\mu$ . When a fermion does not interact with other particles, its Lagrangian contains only a kinetic term and a mass term,

$$\mathcal{L}_\psi = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \quad (1.2)$$

where  $m$  is the mass of the fermion,  $\bar{\psi} = \psi^\dagger\gamma^0$  the adjoint fermion field and  $\gamma^\mu$  are the Dirac matrices, defined by  $\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$ .

Interactions between particles arise as a consequence of the symmetries that are imposed on the Lagrangian. If  $\psi(x)$  undergoes a local phase transformation,

$$\psi(x) \rightarrow \psi'(x) = e^{i\alpha(x)}\psi(x), \quad (1.3)$$

where  $\alpha(x)$  is an arbitrary function of the local coordinates  $x$ , the Lagrangian acquires an extra term due to the derivation in Eq. (1.2),

$$\mathcal{L}_{\psi'} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - \bar{\psi}\gamma^\mu(\partial_\mu\alpha)\psi - m\bar{\psi}\psi. \quad (1.4)$$

If the transformation described by  $\alpha(x)$  can be written as a function of the generators of a Lie group [5],  $\alpha(x) = \epsilon(x) \cdot \vec{T} = \epsilon^i T_i$ , the Lagrangian can be made invariant under the transformation in Eq. (1.3) by introducing a vector boson field  $A_\mu^i$  for each generator  $T_i$ . Defining the covariant derivative as

$$D_\mu = \partial_\mu + ig \frac{\vec{T}}{2} \cdot \vec{A}_\mu, \quad (1.5)$$

where  $g$  is the coupling strength between  $\vec{A}_\mu$  and  $\psi$ , Eq. (1.4) can be rewritten as

$$\mathcal{L}_\psi = i\bar{\psi}\gamma^\mu D_\mu\psi - m\bar{\psi}\psi. \quad (1.6)$$

The transformation in Eq. (1.3) combined with a simultaneous transformation of the fields  $A_\mu^i$  according to

$$A_\mu \rightarrow A'_\mu = A_\mu^i - \partial_\mu\alpha, \quad (1.7)$$

leaves the Lagrangian invariant, i.e.  $\mathcal{L}_{\psi'} = \mathcal{L}_\psi$ .

Thus requiring the Lagrangian to be invariant under Lie group transformations, which are also called gauge transformations, brings about an interaction between a fermion and a vector field. This principle can be extended to the SM Lagrangian, which is invariant under the symmetry group

$$G_{SM} = SU(3) \times SU(2) \times U(1). \quad (1.8)$$

This is a direct product of unitary groups<sup>1</sup>, where  $SU(2) \times U(1)$  describes the symmetry of the electroweak theory, which combines the interactions of the electromagnetic and the weak nuclear force, and  $SU(3)$  represents the symmetry of the theory describing the strong nuclear force, quantum chromodynamics (QCD). The covariant derivative that leaves the SM Lagrangian invariant under a  $G_{SM}$  transformation is equal to

$$D_\mu = \partial_\mu + ig_s \frac{\lambda_a}{2} G_\mu^a + ig \frac{\sigma_i}{2} W_\mu^i + ig' \frac{Y}{2} B_\mu, \quad (1.9)$$

where  $g$ ,  $g'$ , and  $g_s$  are the respective coupling strengths of the vector fields in each term. The rest of the notation will be explained below.

## Quantum chromodynamics

The second term in Eq. (1.9) ensures that the SM Lagrangian is invariant under  $SU(3)$  transformations. The generators of the  $SU(3)$  group that describes the symmetry of QCD are the Gell-Mann matrices  $\lambda_a$ , where  $a = 1 \dots 8$  [4]. Therefore, eight gluon fields  $G_\mu^a$  are introduced, whose excitations are the massless gluons. Since  $SU(3)$  is a non-commutative group, gluons can interact amongst themselves.

<sup>1</sup>The unitary group  $U(n)$  contains all  $n \times n$  matrices whose inverse is equal to their hermitian conjugate,  $U^\dagger U = 1$ . In addition, all matrices in the special unitary group  $SU(n)$ , which is a subgroup of  $U(n)$ , have a matrix determinant equal to one.

## Electroweak theory

The fields responsible for the electroweak interactions are introduced by a local phase transformation under  $SU(2) \times U(1)$ . In order to account for the electroweak parity violation that was observed in some experiments [18], only left-handed particles, i.e. particles with a spin opposite to their direction of motion, are allowed to couple to  $SU(2)$  [5]. The generators of the electroweak symmetry group are the Pauli matrices  $\sigma_i$  ( $i = 1 \dots 3$ ) and the hypercharge  $Y$  [4], thus three  $W_\mu^i$  fields are introduced by the transformation under  $SU(2)$  and one field  $B_\mu$  for  $U(1)$ . Contrary to the gluon fields, these cannot be directly related to any of the bosons in Table 1.2. Using the coupling strengths of the fields, the weak mixing angle is defined,

$$\tan \theta_W = \frac{g'}{g}, \quad (1.10)$$

such that the photon and the Z and  $W^\pm$  bosons can be retrieved as

$$A_\mu = W_\mu^3 \sin \theta_W + B_\mu \cos \theta_W, \quad (1.11a)$$

$$Z_\mu = W_\mu^3 \cos \theta_W - B_\mu \sin \theta_W, \quad (1.11b)$$

$$W_\mu^\pm = \sqrt{\frac{1}{2}}(W_\mu^1 \mp iW_\mu^2), \quad (1.11c)$$

respectively.

## Electroweak symmetry breaking: The Brout–Englert–Higgs mechanism

All the vector fields introduced above are massless. As this does not correspond with observations, vector boson mass terms have to be added to the SM Lagrangian in order for the W and Z bosons to acquire mass. However, this breaks the gauge invariance of the Lagrangian. The Brout–Englert–Higgs mechanism shows that this does not need to be a problem, as the electroweak symmetry can be broken spontaneously under certain conditions. The key concept is the introduction of a scalar field,

$$\Phi = \begin{pmatrix} \phi_1^a + i\phi_2^a \\ \phi_1^b + i\phi_2^b \end{pmatrix}, \quad (1.12)$$

which restores the invariance of the Lagrangian by breaking the symmetry of the vacuum state. The potential of this scalar field is equal to

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \quad (1.13)$$

where  $\mu^2$  is a mass parameter and  $\lambda > 0$  a measure for the field's self-interaction. If  $\mu^2 < 0$ , the potential will be shaped as in Figure 1.4. Instead of having a single minimum, as would be the case if  $\mu^2$  was positive, an equidistant ring of minima with a value of

$$v = \sqrt{\frac{-\mu^2}{2\lambda}} \simeq 246 \text{ GeV}, \quad (1.14)$$

where  $v$  is called the vacuum expectation value, is distributed around a local maximum.

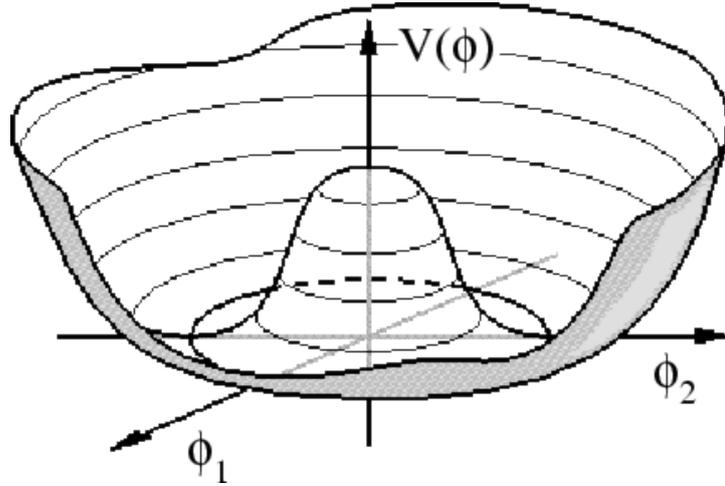


Figure 1.4: Distribution of the potential of a scalar field when  $\mu^2 < 0$  [19].

The minimum of the scalar potential is thus not uniquely defined and the symmetry is broken by choosing a particular minimum. The scalar field can be re-written as an oscillation around this minimum,

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \sigma(x) \end{pmatrix}, \quad (1.15)$$

where  $\sigma(x)$  is the Brout–Englert–Higgs field, which has a spin-0 particle with a mass of  $\sqrt{2\lambda}v$  associated to it, the H boson. The degrees of freedom of the other fields in Eq. (1.12) have been used to give mass to the W and Z bosons,

$$m_W = \frac{1}{2}vg, \quad m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}, \quad \text{and} \quad \frac{m_W}{m_Z} = \cos\theta_W. \quad (1.16)$$

Fermions, on the other hand, acquire mass by adding gauge invariant terms of the form  $g_y\bar{\psi}\Phi\psi$  to the SM Lagrangian, where  $g_y$  is the coupling strength of the scalar field to the fermion, which acquires a mass

$$m_f = g_y v / \sqrt{2}. \quad (1.17)$$

### 1.1.3 Open questions

Far from being ‘just another abstract theory’, the SM presents an extraordinarily precise interpretation of experimental observations. One of its greatest successes was its ability to predict the existence of the H boson about 50 years before its actual discovery in 2012. However, despite its predictive capacities, the SM cannot explain all current observations. Some of the open questions that still need addressing are briefly discussed below.

#### What about gravity?

Whereas the effect of gravity is negligible at the energy scales that are currently under investigation, it is expected that gravity will play a more important role

at higher energy scales, such as the Planck scale,  $\Lambda_p \sim 10^{19}$  GeV [4]. In order to describe gravity, a consistent quantum theory of general relativity must be constructed. Despite much effort, it is not yet clear how this will be realised. At the moment of writing, one of the most promising paths is string theory [20].

### **What are dark matter and dark energy?**

The SM provides an extremely good description of the visible universe. Unfortunately, the visible part amounts to a mere  $\sim 5\%$  of the mass-energy content of the universe. Cosmological observations, such as the rotation curves of galaxies and the structure of the cosmic microwave background, suggest that there is much more matter than can be seen. The accelerated expansion of the universe, on the other hand, can only be possible if there is “something” that pushes the galaxies apart. It is estimated that the mass-energy content of the universe contains about 26% of dark matter and 69% of the elusive dark energy. None of the particles or processes described by the SM can account for these observations.

### **How do neutrinos acquire mass?**

Neutrinos are massless in the original formulation of the SM. However, when measuring solar neutrinos, the SNO experiment [21] observed only a third of the amount of neutrinos that was expected considering the nuclear activity of the sun. As the detector is optimised to measure electron neutrinos that are produced in nuclear fusion reactions, this deficit can be explained by the concept of neutrino oscillations, where neutrinos of one kind, say electron neutrinos, transform into neutrinos of another kind, muon or tau neutrinos. In order for this to happen, the neutrinos must have a different mass. For this discovery, the SNO experiment received the physics Nobel prize in 2015 together with the Super-Kamiokande [22] experiment, which observed the same thing for atmospheric neutrinos. Several models have been proposed to explain how neutrinos can obtain mass [23], but it is unclear which of these, if any, provides the best description of nature.

### **Why are there three generations of matter?**

All visible matter can be constructed using the fermions of the first generation. Yet three generations of matter are observed. It is unclear what the purpose of the other two generations is. Due to the apparent unpredictability of their comparative masses, the question whether there are more than three generations is raised. Despite indications that there are indeed only three generations of matter [24], many searches for fourth generation particles are currently ongoing, as these cannot be excluded by the theoretical framework of the SM.

### **What about the hierarchy problem?**

The mass of a particle, e.g. the H boson, is influenced by its interactions with other particles. The larger the mass of the interacting particle, the larger its influence on the H boson mass. Separating the bare particle mass and its corrections, which are proportional to the energy scale  $\Lambda$  at which the interaction process takes place, the H boson mass can be expressed as  $m_H^2 = (m_{H,\text{bare}})^2 + \mathcal{O}(\Lambda^2)$ . At the electroweak scale ( $\sim 10^2$  GeV) the corrections are of the same order of magnitude as the bare mass. When moving to higher energy scales, however, an unnatural amount of

fine-tuning is necessary to keep the H boson mass at its observed value of about 125 GeV. This is also called the naturalness problem. One possible solution is to extend the SM with more particles that interact with the H boson in such a way that the correction terms by SM particles get (partially) cancelled. One model that provides this is supersymmetry [25, 26], where each fermion gets a bosonic partner and each boson an associated fermion. Some of the additional particles introduced by supersymmetry can also be considered as dark matter candidates.

## 1.2 The Top Quark Sector

Since its discovery by the CDF [27] and D0 [28] experiments at the Tevatron collider [29] in 1995, many studies have been performed to try and measure the top quark's properties with the utmost precision. Especially after the discovery of the H boson by the ATLAS [30] and CMS [31] experiments at the Large Hadron Collider (LHC) [32], which fixed the last unknown parameters of the SM, precision measurements provide an excellent consistency check of the most tested theory in particle physics. One of these tests is a simultaneous indirect measurement of the top quark and W boson masses using measurements of electroweak variables, such as  $\theta_W$  [33]. The results are visualised in Figure 1.5. The contours of the direct mass measurements are indicated in green and the result of the indirect measurement is indicated by a blue or grey ellipse, depending on whether the measured H boson mass was used as a constraint or not. Both ellipses overlap with the direct measurements. If the uncertainties on the masses of the top quark, W and H bosons can be further reduced, this would provide an even

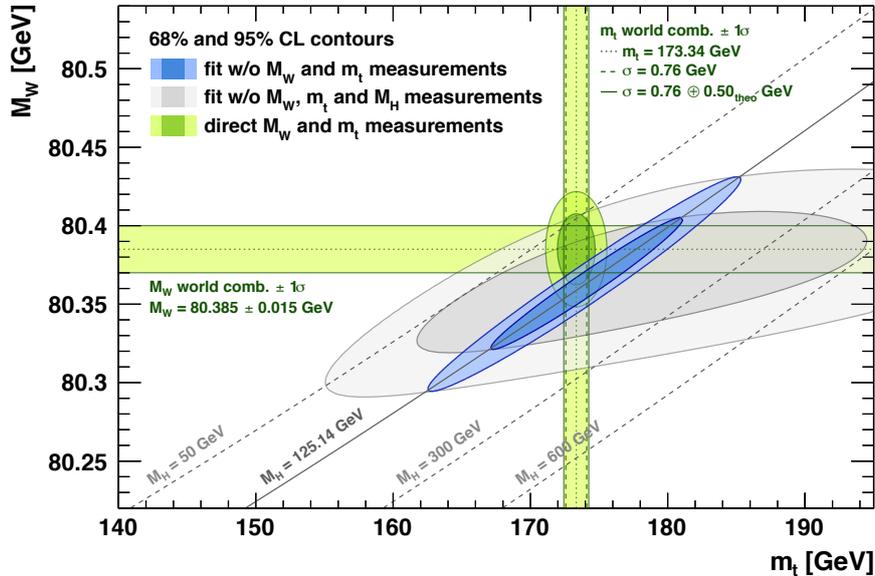


Figure 1.5: Comparison of the measured W boson and top quark masses (green) with the prediction of the SM for these masses using the measurements of other electroweak observables [33]. The blue (grey) ellipse represents the measurement where the H boson mass has (not) been used as a constraint in the fit.

more stringent test of the SM. Amongst these three masses, the top quark mass is known with the least precision.

Not only the mass of the top quark can be used to test the SM, also properties such as its decay width or polarisation can give valuable information about physics beyond the SM (BSM). If precision measurements of the top quark's properties deviate from the SM predictions, this might be an indication of interactions with particles that have not been discovered yet. As there is no indication of new particles at current energies, it is expected that BSM particles have large masses. The top quark being the heaviest of the known particles is therefore an excellent candidate to interact with these particles.

Because of its large mass, the top quark has a lifetime of only  $5 \times 10^{-25}$  s [8]. This means that the top quark decays before it can hadronise (see more in Section 3.1.3), which enables it to transfer some of its properties, such as its direction of spin, to its decay products. Investigating the angular distributions of the top quark decay products can thus give an indication of, amongst others, the top quark polarisation and possible anomalous couplings, which are suppressed in the SM, influencing the top quark production and decay.

### 1.2.1 Production and decay

Top quarks can be produced singly or in quark-antiquark pairs. The latter process is governed by the strong interaction, while the former is an electroweak process. As top quark pair production is the dominant production mode, this will be the focal point of this section.

In proton-proton colliders there are two mechanisms to produce a top quark pair, i.e. gluon fusion and quark-antiquark annihilation. These processes are visualised in Figure 1.6. At a centre-of-mass energy of 14 TeV, about 90% of top quark pairs will be produced via gluon fusion [8].

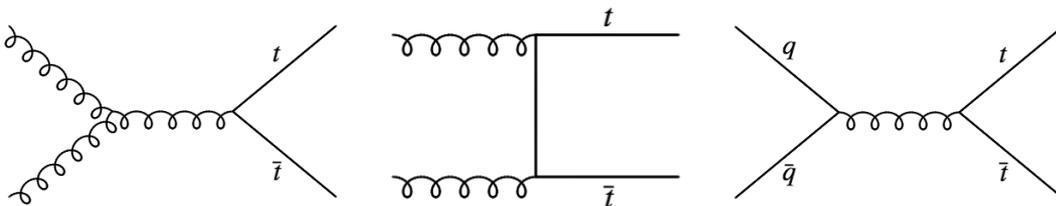


Figure 1.6: Production of a top quark pair.

The top quark decays primarily into a bottom quark and a W boson. The final state of the top quark decay thus depends on the decay of the W boson, which decays two out of three times into a quark-antiquark pair and one out of three times into a charged lepton and a neutrino [8]. As a top quark pair contains two decaying top quarks and thus also two decaying W bosons, three different final states are observed. Firstly, the all-hadronic final state, where both W bosons decay into a quark-antiquark pair, has a probability of  $4/9 \simeq 45\%$ . A similar probability is observed for semileptonic decays, where only one of the W bosons decays into a quark-antiquark pair. A dileptonic decay, where

both  $W$  bosons decay into a lepton-neutrino pair, occurs in only  $1/9 \simeq 11\%$  of the cases. In practice, the tau lepton also decays, either into hadrons or into a lepton-neutrino pair. Therefore, the all-hadronic final state is the most commonly observed decay mode for top quark pairs. Due to colour confinement, the quarks will hadronise and are observed as a stream of particles moving in the same general direction. This is called a ‘jet’.

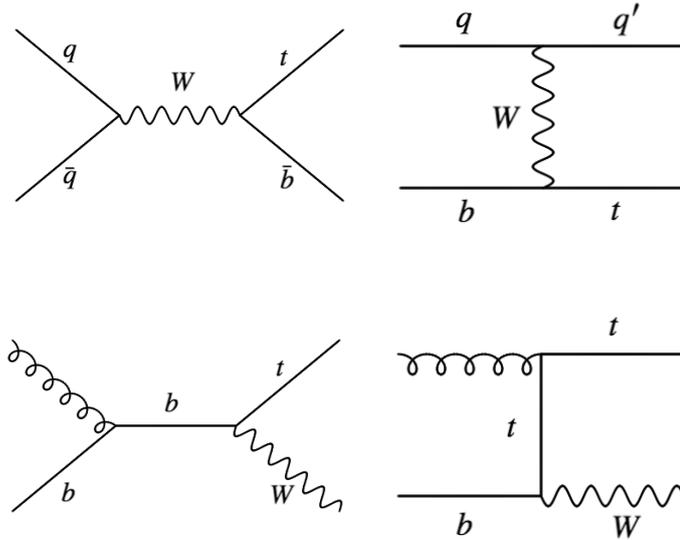


Figure 1.7: Production of a single top quark via the  $s$  channel (top left),  $t$  channel (top right), and in association with a  $W$  boson (bottom).

Top quarks are produced singly via the charged current electroweak processes that are visualised in Figure 1.7. The top left process, also called the  $s$  channel production mode, is the least common, because the intermediate  $W$  boson has to produce the heavier top quark. The  $t$  channel production in the top right corner, on the other hand, occurs the most often. A top quark can also be produced in association with a  $W$  boson, as is shown in the lower diagrams in Figure 1.7. For convenience, these processes are denoted together as the  $tW$  channel. The  $b$  quarks and antiquarks appearing in the initial states of the  $t$  and  $tW$  channels originate from a gluon splitting into a  $b\bar{b}$  pair, resulting in equal quantities of  $b$  quarks and antiquarks. Therefore, top quarks and antiquarks are produced in equal amounts via the  $tW$  channel, while for the  $t$  channel the ratio of produced top quarks to antitop quarks depends on the quark content of the proton. Considering only valence quarks, one would expect to produce the positively charged top quark twice as often as the negatively charged antitop quark. At higher energies, however, there are more sea quarks available that can partake in the process (see Section 3.1.1), thus reducing the ratio of positively-to-negatively charged quarks in the proton. As a consequence, only about 1.7 top quarks are produced for each antitop quark via the  $t$  channel at 13 TeV, whereas this is about 1.85 at 8 TeV [34].

## 1.2.2 Top quark mass

Precision measurements of the top quark mass have been performed by the CDF [35–38] and D0 [39, 40] experiments at the Tevatron collider and by the ATLAS [41–46] and CMS [47–53] experiments at the LHC. Most of these results were obtained using top quark pair events, treating each decay channel separately. The uncertainty on the individual measurements can be reduced by combining the measurements from several decay channels and different experiments. In the latter case, it is important to deconvolve the systematic effects induced by the detector itself, e.g. limitations due to granularity, from the measurements before the combination. The results obtained with data recorded by the LHC experiments at 7 TeV are combined in [54] and, together with the combination of the Tevatron results [55], a world average of  $(173.34 \pm 0.76)$  GeV is obtained for the top quark mass [56]. Since then, measurements using data collected at 8 TeV have further improved the resolution of the measured top quark mass. If the LHC measurements at 7 and 8 TeV are combined for each experiment separately, a top quark mass of  $(172.51 \pm 0.50)$  GeV is obtained for the ATLAS combination and  $(172.44 \pm 0.48)$  GeV for the CMS results, which is the most precise measurement to date. A combination of the results from both experiments is expected to further reduce the uncertainty on the measurement of the top quark mass. An overview of the measurements that were taken into account in the calculation of the world average top quark mass value can be found in Figure 1.8. Newer results are compared to the world average value in Figure 1.9.

The most precise individual top quark mass measurement to date is the one in the semileptonic decay channel using data recorded by the CMS experiment in 2012 at a centre-of-mass energy of 8 TeV [51]. The top quark mass is measured to be

$$m_t = 172.35 \pm 0.16 \text{ (stat.)} \pm 0.48 \text{ (syst.) GeV} = 172.35 \pm 0.51 \text{ GeV}, \quad (1.18)$$

using the ideogram method, which performs a two-dimensional measurement of the top quark mass and an energy scale factor. This allows to partially absorb the non-negligible uncertainty due to the limited energy resolution of the detector into the energy scale dimension, which leads to a smaller uncertainty on the top quark mass. Using simulated data, a likelihood is determined for several values of the top quark mass and the energy scale factor. The measurement result then corresponds to the maximum of the 2D likelihood function.

## 1.2.3 Top quark decay width

The decay width  $\Gamma$  of a particle is a measure for the probability that a particle decays into another (pair of) particle(s) in a certain time frame. It is inversely proportional to the particle's lifetime. If a particle has more than one way of decaying, the total decay width is the sum of the partial widths for each final state,  $\Gamma(X) = \sum_f \Gamma(X \rightarrow f)$ . As heavy particles in general have more options to decay into lighter particles, their decay width is larger than that of light particles.

The decay width of a particle can be visualised using its mass distribution. When performing a mass measurement, the observed width of the mass distribution is determined by the decay width of the particle and the statistical and systematic

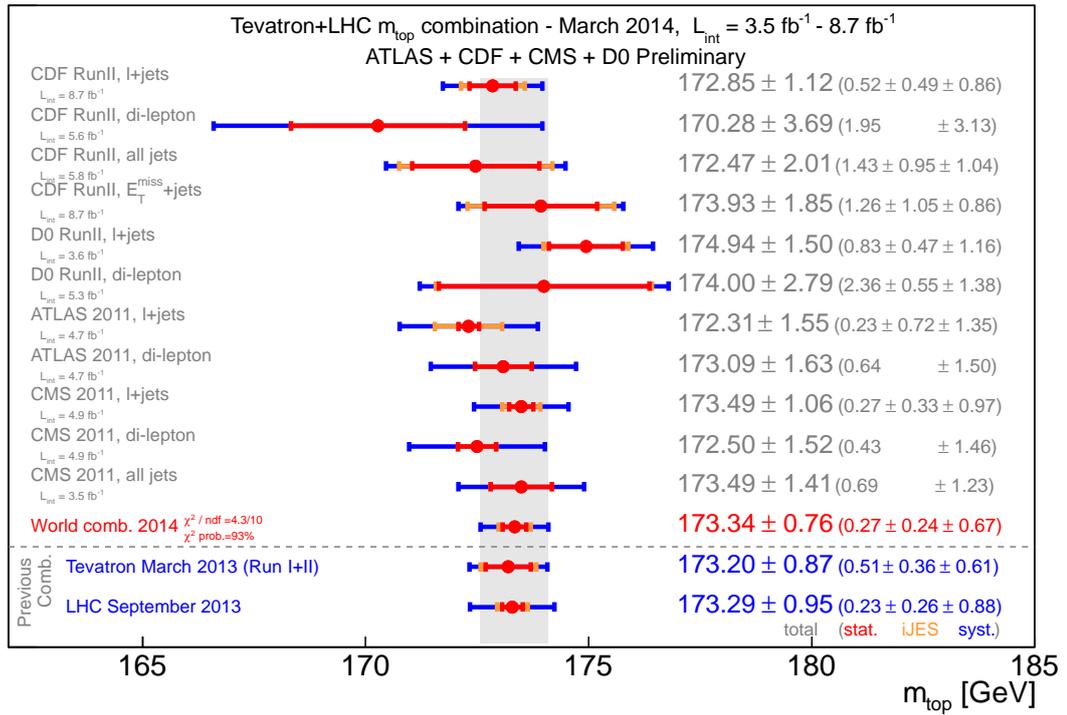


Figure 1.8: Overview of the top quark mass measurements that are included in the world combination [57]. The analyses are performed by the CDF and D0 experiments using proton-antiproton collisions at a centre-of-mass energy of 1.8 TeV and 1.96 TeV, and the ATLAS and CMS experiments using proton-proton collisions at 7 TeV.

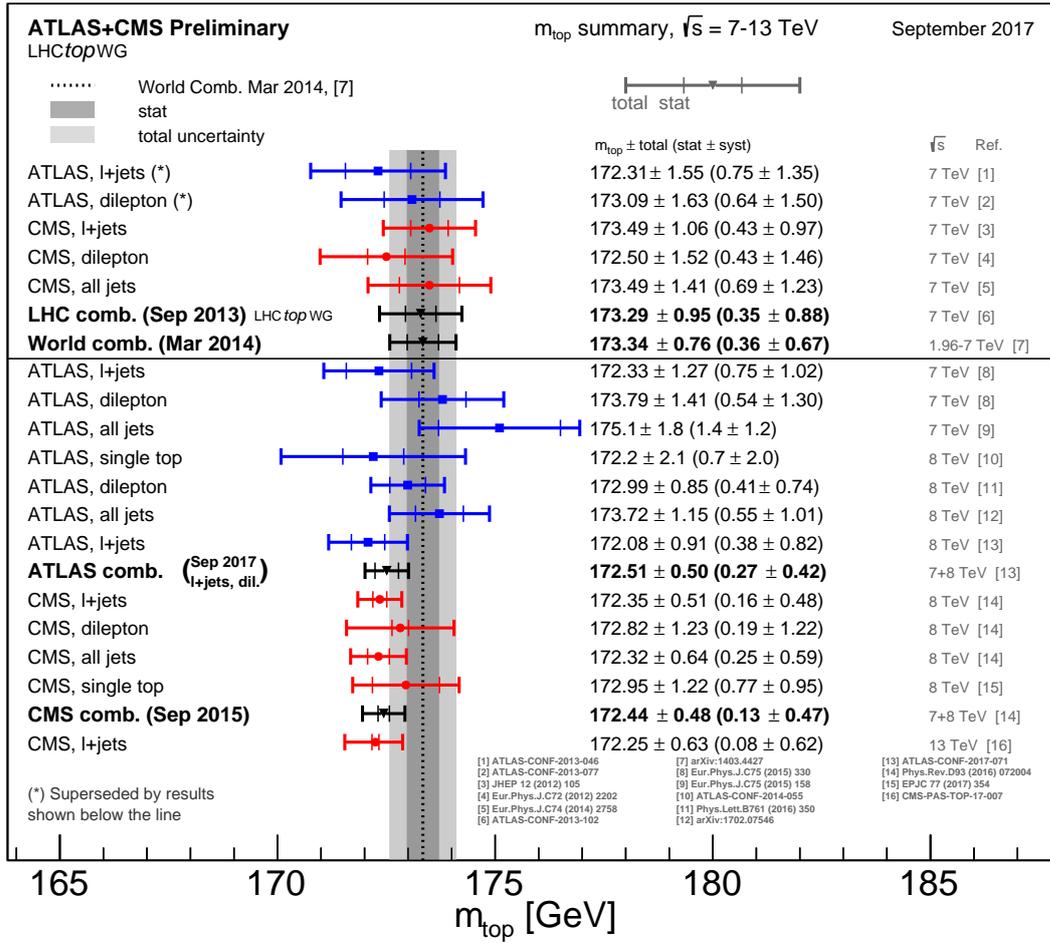


Figure 1.9: Summary of the top quark mass measurements performed by the ATLAS and CMS experiments, as well as combinations of the LHC results and the world average, combining the LHC and Tevatron results that are presented in Figure 1.8 [57].

uncertainties affecting the measurement. Whereas the latter can theoretically be reduced to zero, the decay width is ‘irreducible’. If the uncertainties are considered to be zero, the mass  $m$  of a particle is distributed according to a Breit-Wigner (BW) function,

$$\text{BW}(m) = \frac{1}{\pi} \frac{\frac{\Gamma}{2}}{(m - M)^2 + \left(\frac{\Gamma}{2}\right)^2}, \quad (1.19)$$

where  $M$  and  $\Gamma$  are resp. the ‘true’ mass and decay width of a particle. The decay width  $\Gamma$  represents the full width at half-maximum (FWHM) of the distribution, which is related to the standard deviation  $\sigma$  as  $\Gamma \simeq 2.35\sigma$ . Compared to a Gaussian function, a BW distribution has more pronounced tails, as can be seen in Figure 1.10. Further, it is better suited to describe particles with very small decay widths, even when  $\Gamma \rightarrow 0$ , as is the case for stable particles.

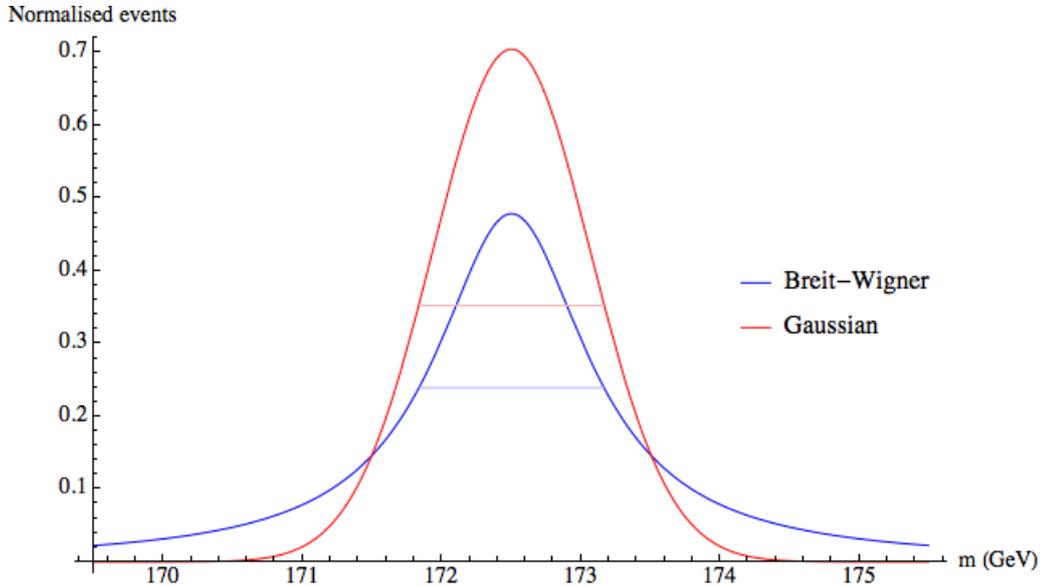


Figure 1.10: Comparison of a Breit–Wigner (blue) to a Gaussian distribution (red). Both functions are centred around  $M = 172.5$  GeV and have a FWHM of 1.33 GeV, which is indicated in light blue (red).

In the SM, the top quark decays almost exclusively into a W boson and a b quark. Its total decay width is thus equal to the partial width of this process,

$$\Gamma_t = \Gamma(t) \simeq \Gamma(t \rightarrow Wb). \quad (1.20)$$

Assuming  $m_b \ll m_W < m_t$ , the SM prediction for the top quark width is equal to

$$\Gamma_t = \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{m_W^2}{m_t^2}\right)^2 \left(1 + 2 \frac{m_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3\pi} \left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right], \quad (1.21)$$

where the last terms are next-to-leading-order (NLO) corrections [8, 58, 59],  $G_F$  is the Fermi constant [7] and  $\alpha_s = g_s^2/4\pi$  the coupling constant of the strong force. From this

expression it is clear that the top quark decay width is strongly dependent on the top quark mass. Therefore, any quoted value for the top quark decay width needs to be accompanied by the top quark mass used in the measurement. In this thesis, a top quark mass of 172.5 GeV is assumed, unless specified otherwise. The value of the top quark decay width evaluated for several top quark masses can be found in Table 1.3. It can be seen that the width value changes by roughly 2% for a change of 1 GeV in the mass value. Considering a top quark mass of 172.5 GeV, the top quark decay width is equal to 1.33 GeV at NLO.

Table 1.3: The top quark decay width as a function of the top quark mass according to Eq. (1.21).

Mass (GeV)	Width (GeV)
171.5	1.303
172	1.316
172.5	1.329
173	1.343
173.5	1.357
174	1.370

It is expected that second order QCD corrections have an effect of about 2% on the value of the top quark decay width [60–62]. As these are only determined numerically and no precise analytical expressions are publicly available, the precision of this correction depends strongly on the values of e.g.  $m_t$  and  $\alpha_s$  that were used for the calculation. Most of these calculations were performed in the late nineties and are thus not up to date [60, 61]. In addition, the last term in Eq. 1.21 is only an approximate correction, ignoring terms of the order  $\alpha_s (m_W/m_t)^2$  [59]. As a result the influence of first-order QCD corrections is overestimated by about 1.5%. This effect falls within the uncertainties of the second-order QCD corrections, but the full expression in  $\alpha_s$  needs to be considered when these are added. The most recent top quark width calculation, which was published in 2013, predicts a value of 1.35 GeV for a top quark mass of 173.5 GeV [62]. Extrapolation to  $m_t = 172.5$  GeV gives  $\Gamma_t = 1.32$  GeV [62].

Contrary to the top quark mass, only few measurements of the top quark decay width have been performed. There are two approaches. Direct measurements [63–67] use observables for which the distributions are influenced by the top quark decay width. When the observable is the reconstructed top quark mass, the width of the distribution is directly influenced by varying  $\Gamma_t$ . In general, these changes are rather small, so most direct measurements up till now have determined a confidence interval without quoting a central value. Indirect measurements [68–70], on the other hand, aim to measure quantities that are related to the top quark production and decay, such as the production cross section of the top quark in the t channel or the branching ratio of the top quark to a particular other quark,  $B(t \rightarrow Wq)$ , where  $q = d, s, b$ . These

measurements can be combined in such a way that the top quark decay width can be determined, which results in much more precise measurements than the direct ones. A disadvantage of this approach is that it implicitly assumes that all top quark decays follow a  $t \rightarrow Wq$  pattern, thus excluding more exotic decays that can appear in BSM models. A precise direct measurement of the top quark decay width, as is presented in this thesis, is in that respect an important tool to investigate the decay of the top quark in search for new physics phenomena. Examples of these are flavour-changing neutral currents [71], such as  $t \rightarrow Zq$  and  $t \rightarrow Hq$ , which are heavily suppressed in the SM, but whose cross section is enhanced by new physics.

Current measurements of the top quark decay width, in the remainder of this text also referred to as the top quark width, are summarised in Figure 1.11. The direct measurements can be found above and the indirect measurements below the dashed line.

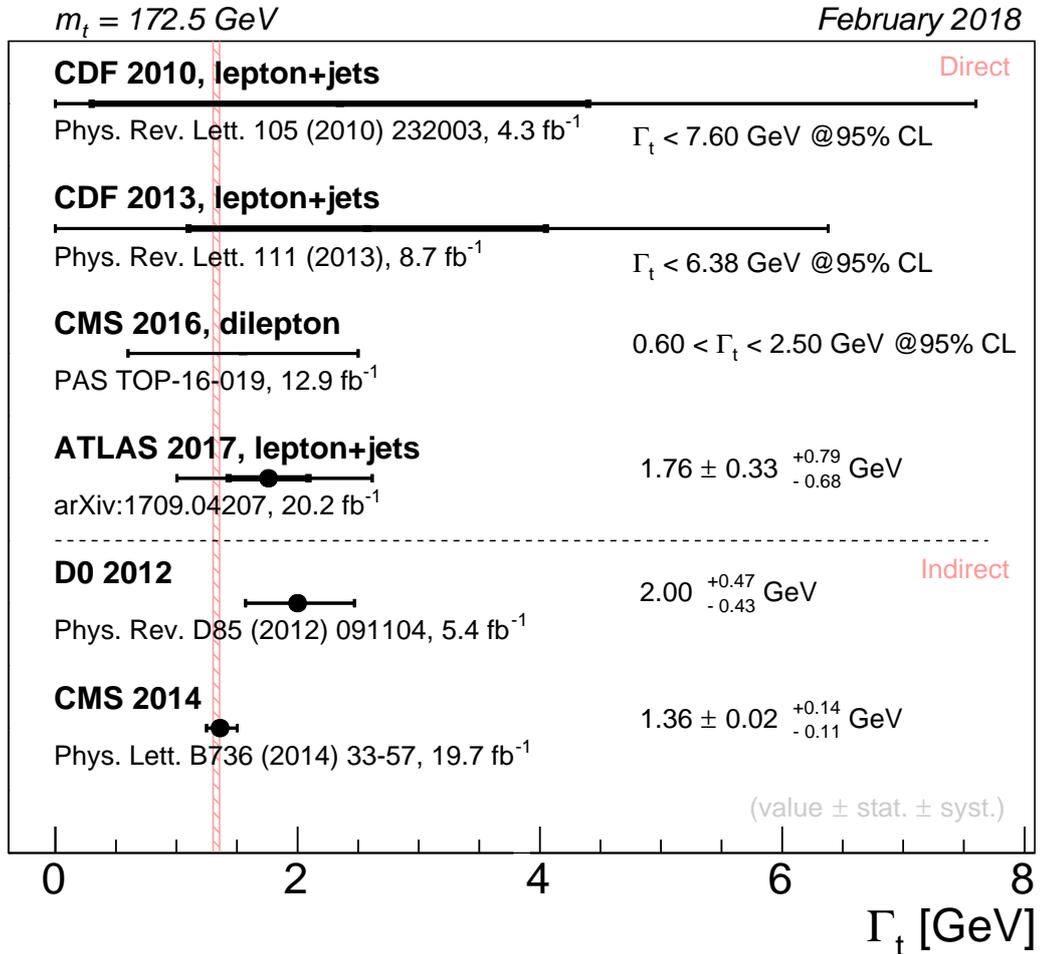


Figure 1.11: Summary of the top quark decay width measurements performed by the ATLAS, CDF, CMS, and D0 experiments, all considering a top quark mass of 172.5 GeV. Direct measurements are grouped above the dashed line and indirect measurements below. The reference interval (red) indicates the theoretical prediction for the top quark decay width at NLO when considering  $m_t = 172 - 173$  GeV.

Whereas the statistical uncertainty was still a limiting factor for the Tevatron measurements, more recent measurements are dominated by systematic uncertainties.

The indirect measurements of the D0 [69] and CMS collaborations [70] are based on the same principle. Both determine the cross section for top quark production in the  $t$  channel and the branching fraction for a top quark decaying into a  $W$  boson and a  $b$  quark, i.e.

$$\mathcal{B}(t \rightarrow Wb) = \frac{\Gamma(t \rightarrow Wb)}{\Gamma_t}. \quad (1.22)$$

If  $\sum_q \mathcal{B}(t \rightarrow Wq)$ , where  $q$  is a down-type quark, is assumed to be equal to one, the top quark decay width can be determined as

$$\Gamma_t = \frac{\Gamma(t \rightarrow Wb)_{\text{SM}}}{\mathcal{B}(t \rightarrow Wb)} \frac{\sigma_{t\text{-chan}}}{\sigma_{t\text{-chan,SM}}}, \quad (1.23)$$

where  $\Gamma(t \rightarrow Wb)_{\text{SM}}$  and  $\sigma_{t\text{-chan,SM}}$  are the partial width and cross section predicted by the SM. The precision of these kinds of measurements is most affected by the efficiency to identify a jet as originating from a  $b$  quark.

Direct top quark width measurements, on the other hand, use a likelihood-based method to compare the shape of variable distributions in data to those that are simulated under certain top quark width hypotheses. The CDF experiment performs a two-dimensional top quark width measurement using top quark pair events in the semileptonic decay channel [64, 65]. In addition to the variable used to measure the top quark decay width, a second variable, i.e. the reconstructed mass of the hadronically decaying  $W$  boson, is employed to simultaneously constrain the jet energy scale. An unbinned likelihood fit is performed and the systematic uncertainties are folded into the likelihood to obtain confidence intervals. Other than using more data, which improves the still significant statistical uncertainty, the main difference between the CDF measurements in Figure 1.11 is that the later one actively tries to constrain the jet energy scale and resolution, the latter being a dominant systematic uncertainty for both measurements, before constructing the likelihood function. The ATLAS collaboration, on the other hand, uses a one-dimensional approach, but performs a simultaneous fit of two variables [67]. Whereas the CDF measurements only used variables connected to the hadronically decaying top quark, the ATLAS experiment employs information from both top quarks. Using templates for different top quark widths, a binned likelihood is performed for each set of templates. Systematic uncertainties are taken into account by producing extra templates including the systematic variations. These are then compared to the nominal template using pseudo experiments. The dominant systematic uncertainties are the jet energy scale and resolution, and modelling uncertainties. A different analysis strategy is used by the CMS collaboration, where a likelihood ratio scan is performed in the dileptonic decay channel. The distribution of a variable assuming the SM top quark decay width is directly compared the same distribution subject to an alternative width [66]. The systematic uncertainties are treated as nuisance parameters and the CLs criterion is used to determine confidence intervals [72]. Also here, the theoretical modelling of the signal is a dominant uncertainty and presents one of the greatest challenges for a precise measurement.

Using semileptonically decaying top quark pair events that were collected by the CMS experiment, this thesis presents a direct top quark decay width measurement in one dimension. As different decay channels are used, this measurement is complementary to the CMS measurement described above. A maximum likelihood method is employed, where probability density templates are constructed using the distributions of variables that are sensitive to the top quark decay width. Further, the likelihood functions of different variables are combined in order to reduce the systematic uncertainties that affect the measurement.

# The CMS Experiment and the LHC

---

# 2

To investigate elementary particles one must think big. Present-day research makes use of colliders and large multi-layered particle detectors to examine the open questions of nature. Two beams of particles are accelerated to high energies and made to collide in the centre of a particle detector, which records the debris of the collision. The effectiveness of this technique is based on Einstein's well-known principle  $E = mc^2$ , which states that energy can be transformed into matter (mass) and vice versa. The centre-of-mass energy in the so-called interaction point of the detector is equal to the sum of the energies of the beams.

In order to accelerate particles, they must be charged and stable. In practice, electrons and charged hadrons are most commonly used. Which type of particle is most beneficial depends on the type and purpose of the accelerator. Hadrons are composite particles, so it is unclear which part of the hadron is involved in the actual collision and how much energy that composite carries. Collisions with electrons thus have a much cleaner signature. On the other hand, electrons are much more prone to energy loss when they follow a curved trajectory due to synchrotron radiation. As the amount of energy lost is inversely proportional to  $m^4$ , it is much easier to accelerate hadrons, the lightest of which are about 2000 times heavier than the electron, to high energies in circular colliders.

## 2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is located at the European Organisation for Nuclear Research (CERN) at the Franco-Swiss border near Geneva. Being 27 km in circumference, it is the largest hadron collider in the world. Protons and lead ions are accelerated to energies of 6.5 TeV and 2.56 TeV per nucleon, respectively [73]. To reach these energies, the entire accelerator complex, outlined in Figure 2.1 [73], is used. Protons are created by ionising hydrogen. They are first accelerated in a linear accelerator to get the protons up to an energy of 50 MeV. Then the circular booster accelerates them to an energy of 1.4 GeV before injecting them into the Proton Synchrotron (PS), which in its turn takes the protons to an energy of 25 GeV. Next, the protons are injected into the Super Proton Synchrotron (SPS), where they acquire an energy of 450 GeV before reaching the final stage of their journey, the LHC. Two beam pipes running in opposite directions

are filled such that the protons can collide head-on at well-defined points. In order to avoid collisions with air particles, the beam pipes are kept at an ultrahigh vacuum [32]. The acceleration of lead ions starts from vaporised lead. The ions are first accelerated by a linear collider and then injected into the Low Energy Ion Ring (LEIR), before continuing their journey in the PS, SPS and LHC.

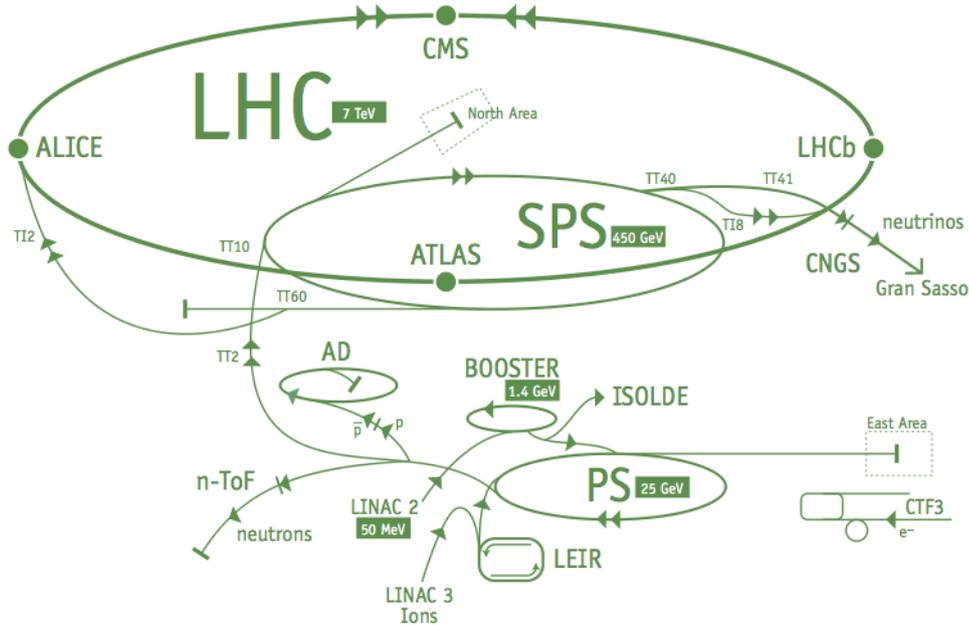


Figure 2.1: A schematic representation of the accelerator complex at CERN [73].

In fact, the LHC is not completely circular, but rather a sequence of eight straight and eight curved sections. The sectors from the middle of a curved section to the middle of the following curved section are called octants. In the straight sections, experiments can be built to record and investigate particle collisions. At the moment, four such experiments have been installed. Two general-purpose detectors, ATLAS [30] and CMS [31], look for any sign of new physics and aim to make precise measurements of the particles they observe. LHCb [74] is specifically designed to observe hadrons that contain a  $b$  quark and measure their properties, while ALICE [75] is optimised for heavy-ion collisions and tries to understand the properties of the quark-gluon plasma that is created when heavy ions collide. Apart from these, there are a couple of smaller LHC experiments. LHCf [76] and TOTEM [77] are located close to the beam line near to the ATLAS and CMS detectors, respectively, in order to measure particles originating from these experiments that are moving close to the beam line in what is called the very forward region of the detectors. TOTEM, for example, is located about 200 m from the interaction point on both sides of the CMS detector and can therefore measure particles that escape undetected from CMS, e.g. particles that traverse the CMS detector through the non-sensitive region occupied by the beam pipe. In 2013, the TOTEM and CMS collaborations decided to join efforts to pursue common goals and under the name of CT-PPS [78], TOTEM is at the moment being incorporated into the CMS experiment. It is expected that CT-PPS will improve

electroweak and QCD measurements in the very forward region when it starts taking data in 2017 [78]. The CMS experiment will be explained in more detail in Section 2.2.

In the curved sections of the LHC, superconducting magnets are installed to bend the particles when they travel through the accelerator. The magnets are made of niobium titanium (NbTi) cables and are cooled to 1.9 K with superfluid helium [32]. The majority of the bending is done by dipole magnets, which can obtain a maximum magnetic field of 8.3 T. They are supplemented by multi-pole magnets that correct for small instabilities at the edges of the dipoles, such that a stable trajectory is attained [79].

Specialised magnets are used for the injection and the dump of the proton beam. As can be seen in Figure 2.2 [80], the clockwise beam is injected into the LHC in octant 2 and the anticlockwise beam in octant 8. In octants 3 and 7 the beams are cleaned by absorbing stray particles that might damage the machine and, whenever needed, they are dumped in octant 6, where large blocks of carbon are placed to absorb the energy of the beams. Each block is cylindrical in shape, with a length of 7.7 m and a diameter of 70 cm, and is surrounded by about 900 tons of radiation shielding material [32]. In the octants where the experiments are positioned, the beam pipes cross over and the beams are focused such that high-intensity collisions are possible.

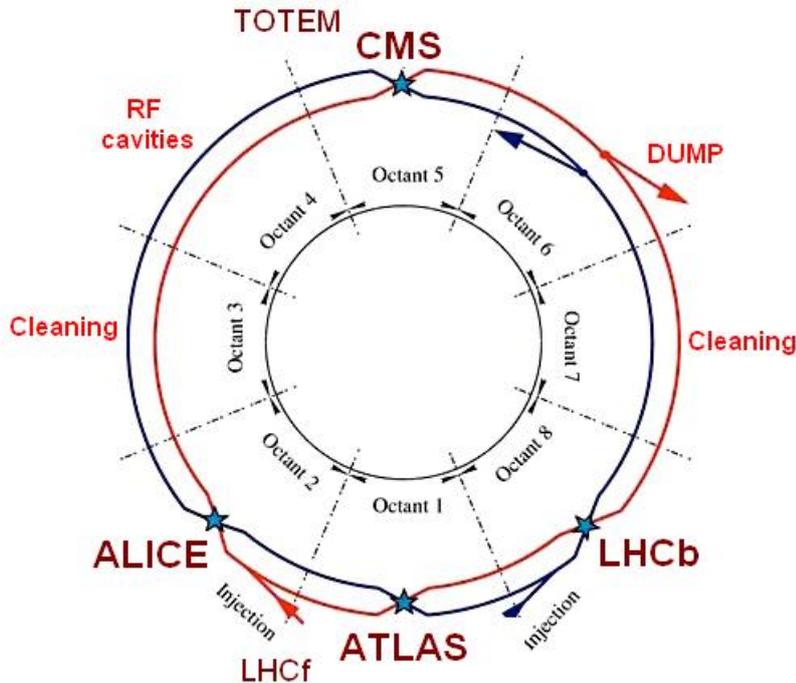


Figure 2.2: Schematic overview of the LHC ring [80]. Beam 1 (red) travels in the clockwise direction and beam 2 (blue) in the anticlockwise direction.

As the individual particles are incredibly small, it is near impossible to aim two protons directly at each other. In order to increase the probability that two protons collide, the particles are organised in bunches with a large amount of protons. Using radiofrequency (RF) cavities to accelerate the particles, bunches of about  $10^{11}$  protons

are created [73, 81]. RF cavities consist of an oscillating electric field that exchanges energy with the particles that pass through the cavity [81]. The amount of energy the particles receive depends on the time of their arrival in the cavities. Particles that are slightly faster than average will be accelerated less than particles that are slightly slower, thus bringing the energies of the particles in the bunch closer together. The frequency of the RF cavities is tuned to 400 MHz, such that protons with an energy of 6.5 TeV are not accelerated any further.

During the 2016 run of the LHC, the bunches were spaced about 7.5 m [73] or 25 ns apart. In order to allow for a safe beam evacuation, a gap of  $3 \mu\text{s}$  is kept free, ensuring that the dedicated beam dump magnets have enough time to reach their full magnetic field, which is needed to deflect the beams towards the beam dump absorbers [32]. In general, the bunches have a width of about 0.2 mm, but close to the interaction points of the experiments they are squeezed further together by quadrupole magnets until their diameter is about  $16 \mu\text{m}$  [79]. After the collisions, the beams are separated again and refocused to minimise the spread of the bunches.

When the beams cross in the interaction point, it is possible that multiple protons from one bunch collide with protons from the other bunch. This will result in several interactions or primary vertices (PVs) from whence particles measured in the detector can be traced back. This is referred to as (in-time) pileup. As there are only 25 ns between each bunch crossing, it is possible that the next collision occurs when particles of the previous bunch crossing are still travelling through the detector. This is called out-of-time pileup. An average pileup of 27 is observed for the 2016 run period [82].

The amount of collisions that occur in a certain time frame is quantified by the luminosity. It depends on the beam parameters of the LHC and is defined as

$$\mathcal{L} = \frac{N^2 n_b f}{4\pi} F', \quad (2.1)$$

where  $N$  is the number of particles per bunch,  $n_b$  the number of bunches,  $f$  the revolution frequency of the bunches and  $F'$  is a geometric factor that takes account of the size of the bunches and their crossing angle in the interaction point [32]. During the 2016 run a maximum instantaneous luminosity of  $1.4 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  was reached and the total integrated luminosity delivered by the LHC during that time period is  $40.8 \text{ fb}^{-1}$ .

## 2.2 The CMS Experiment

With its 12 500 tonnes, the Compact Muon Solenoid (CMS) is by far the heaviest of the LHC experiments. It has a cylindrical shape with a length of 21.6 m and a diameter of 14.6 m and is composed of several subdetectors in an onion-like structure, as is visualised in Figure 2.3 [31]. Each subdetector is optimised to recognise certain particle interactions.

The CMS detector consists of a central section, called the barrel, and two ‘endcaps’ that are shaped such that a maximal hermeticity for particles coming from the interaction region is attained. Its central feature is a 12.5 m long superconducting solenoid magnet

with an inner diameter of 6.3 m. When cooled to 4.7 K using liquid helium, it can support a magnetic field of 3.8 T [31]. Inside the solenoid, a silicon tracker and an electromagnetic and hadronic calorimeter are installed. The iron return yoke that closes the magnetic field lines is interleaved with different types of gaseous muon detectors. A more detailed description for each of these subdetectors can be found below.

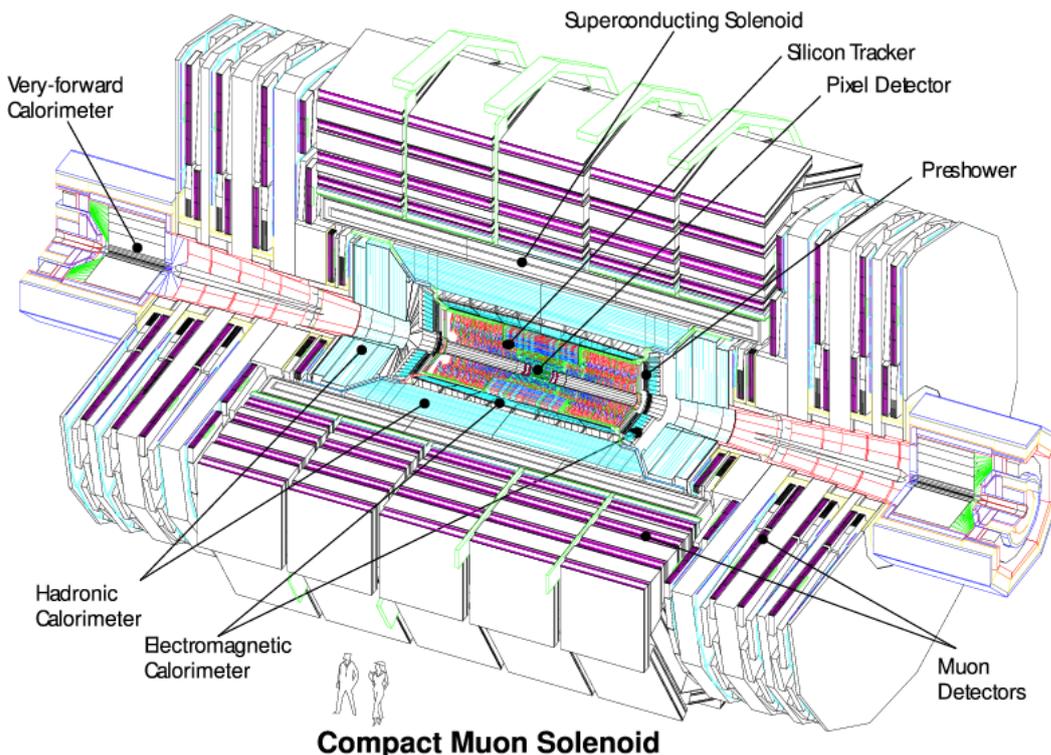


Figure 2.3: A schematic overview of the CMS experiment [31].

In order to describe the detector and the events that are recorded by it, a coordinate system is defined as follows. The  $x$  axis is directed towards the centre of the LHC, the  $y$  axis points upwards and the  $z$  axis lays along the beam line such that a right-handed coordinate system is formed. This corresponds to the direction of the beam that travels in the anticlockwise direction. The  $xy$  plane is also called the transverse plane and the position of a vector in this plane is described by the angle  $\phi$ . The angle between a vector and the positive  $z$  axis is called  $\theta$ . It is often replaced by the pseudorapidity, which is defined as  $\eta = -\ln \tan(\theta/2)$ . This means that a pseudorapidity equal to zero corresponds to a vector lying in the transverse plane, while  $\eta \rightarrow \infty$  corresponds to a vector pointing in the direction of the beam line. The angular distance between two vectors is then defined as  $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ .

As protons are composite particles, the exact energies of the components that are involved in the collision are not known. Therefore, energy conservation laws can only be applied in the transverse plane.

### 2.2.1 The silicon tracker

At the heart of the CMS experiment a silicon tracking detector can be found. It is 5.8 m long, has a diameter of 2.5 m and surrounds the interaction point and the beam pipe. Up to a radius  $r$  of about 10 cm the active volume of the tracker consists of  $320\ \mu\text{m}$  thick silicon pixel modules with a size of  $100 \times 250\ \mu\text{m}^2$ . At larger radii silicon strip modules are used. These are  $10\ \text{cm} \times 80\ \mu\text{m}$  large for radii between  $20\ \text{cm} < r < 55\ \text{cm}$  and  $25\ \text{cm} \times 180\ \mu\text{m}$  between  $55\ \text{cm} < r < 110\ \text{cm}$ . As the noise increases linearly with the strip length, the latter modules have a thickness of  $500\ \mu\text{m}$  to ensure a good signal-to-noise ratio [31].

The silicon pixels are arranged in three barrel layers and two endcap discs. The strip tracker barrel region consists of ten layers, four in the tracker inner barrel (TIB) and six in the outer barrel (TOB), and three discs (TID), which are complemented by nine endcap discs (TEC). This is visualised in Figure 2.4 [83]. Together, the 66 million silicon pixels and 9.3 million strips make up an active area of about  $200\ \text{m}^2$  that covers a pseudorapidity region up to  $|\eta| < 2.5$ . The average hit resolution is  $10\ \mu\text{m}$  in the transverse and about  $25\ \mu\text{m}$  in the longitudinal direction.

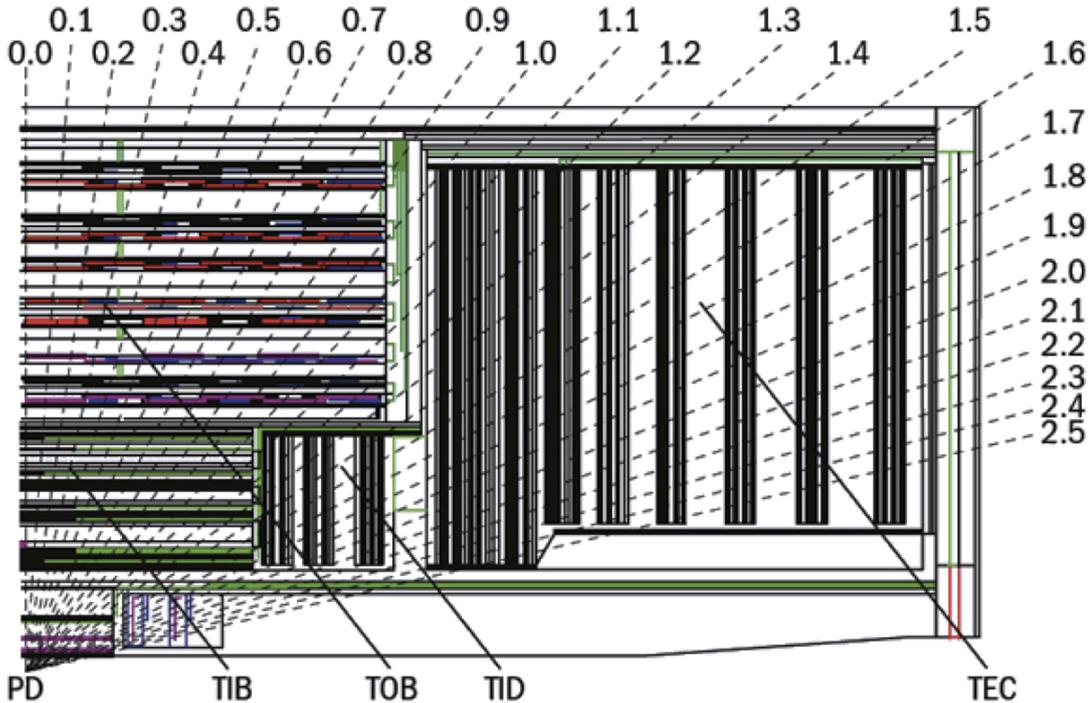


Figure 2.4: Schematic diagram of a quarter of the silicon tracking detectors [83]. The interaction point can be found in the lower left corner. It is enveloped by the pixel detector (PD), the tracker inner barrel (TIB), outer barrel (TOB), the tracker inner discs (TID), and endcap discs (TEC). On the upper and right edges the pseudorapidity  $\eta$  is indicated.

The detection principle of silicon sensors relies on the creation of electron-hole pairs when a charged particle interacts with the material. This creates a small electrical signal that is collected by applying a potential difference to the sensor. CMS uses

$p$ -on- $n$  type sensors for the strips and more radiation hard  $n$ -on- $n$  type sensors for the pixel modules [31], which see a larger flux of particles because they are closer to the interaction point.

All sensors are mounted onto a carbon fibre and graphite support structure, together with the on-detector electronics that are required to enable a fast response time. So despite the lightness of the silicon sensors themselves, particles crossing the tracker volume have to traverse a significant amount of material, as can be seen in Figure 2.5 [31], and many are thus subject to multiple scattering, Bremsstrahlung or nuclear interactions in the tracker material. To counter the heat dissipation caused by the many read-out

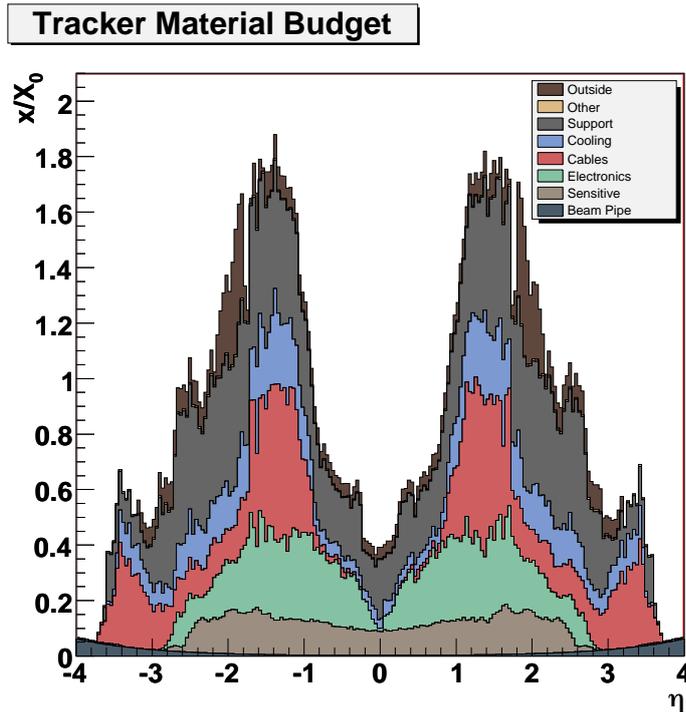


Figure 2.5: Tracker material budget as a function of the pseudorapidity [31]. The material thickness is indicated in units of radiation length.

channels during operation, the pixel detector is cooled to  $-10^{\circ}\text{C}$  and the strip tracker to  $-15^{\circ}\text{C}$  using fluorocarbon ( $\text{C}_6\text{F}_{14}$ ) gas. As the surface of the nearby electromagnetic calorimeter needs to be kept at around  $18^{\circ}\text{C}$  to ensure good performance, a thermal screen is installed around the tracker to prevent condensation of water vapour. In order to avoid humidity inside the tracker, the tracker volume is flushed with cooled nitrogen gas at a rate of about  $400\text{ m}^3/\text{h}$  [84].

## 2.2.2 The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) is designed to estimate the energy of particles such as electrons and photons. It consists of a barrel section (EB), an endcap section (EE) and the so-called preshower (ES), as is shown in Figure 2.6 [17]. The barrel and endcap subdetectors consist of homogeneous scintillating lead tungstate ( $\text{PbWO}_4$ )

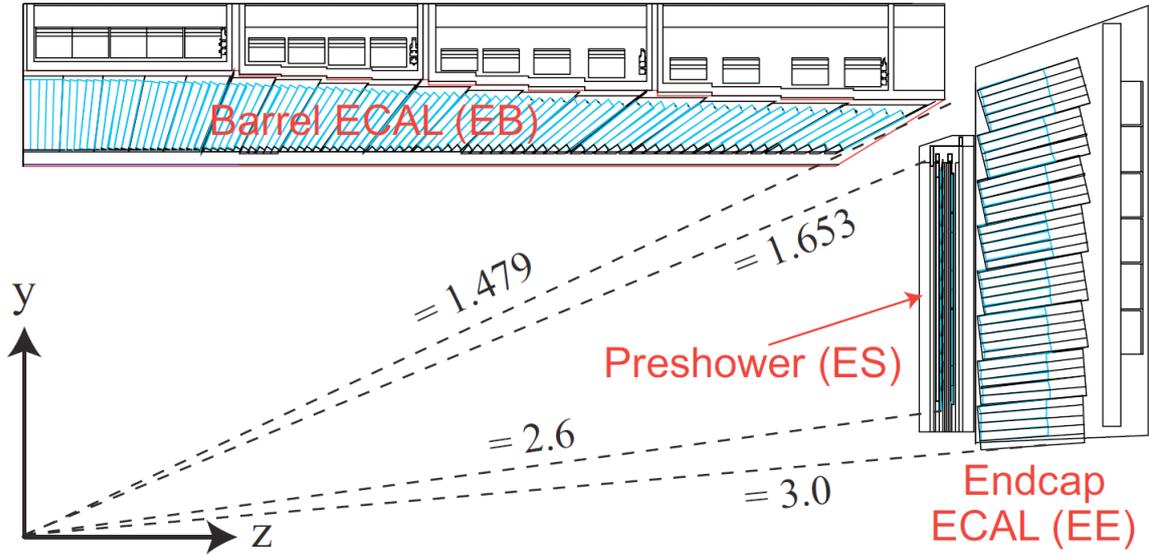


Figure 2.6: Schematic representation of a quarter of the CMS detector showing the position of electromagnetic calorimeters [17]. The interaction point can be found in the lower left corner.

crystals, which have a radiation length  $X_0$  of 0.85 cm and a Moliere radius  $R_0 = 2.19$  cm, thus resulting in a relatively compact subsystem with good granularity [31]. The size of the crystals depends on their position with respect to the interaction point. In the barrel, covering a region up to  $|\eta| < 1.479$ , the crystals are 230 mm long ( $\sim 25.8 X_0$ ), measuring on average  $2.2 \times 2.2 \text{ cm}^2$  at the front face and  $2.6 \times 2.6 \text{ cm}^2$  at the rear face. The truncated pyramidal crystals are arranged in a cylindrical pattern with an inner radius of 1.29 m. They are slightly tilted in the  $\eta$  direction, emulating the direction of flight of particles originating from the interaction point. The endcap crystals, in the range  $1.479 < |\eta| < 3.0$ , are arranged in an  $x$ - $y$  grid. They are 220 mm long ( $\sim 24.7 X_0$ ) and have a front face of  $2.86 \times 2.86 \text{ cm}^2$  and a rear face of  $3.0 \times 3.0 \text{ cm}^2$  large.

When particles traverse the crystals, they induce electromagnetic showers and the emitted photons are collected using dedicated low-light detectors that are able to operate in strong magnetic fields. In the barrel, avalanche photodiodes are used, while the collaboration opted for vacuum phototriodes in the endcaps [31]. On average about 4.5 photons are collected per MeV at a temperature of  $18^\circ\text{C}$  and the light output diminishes with about  $2.1\%/^\circ\text{C}$ . Therefore, the temperature of the EB and EE is closely monitored. A layer of insulating foam is installed between the detectors and the electronics and the temperature is stabilised with a precision of about  $0.05^\circ\text{C}$  using water at  $18^\circ\text{C}$  [31].

In the region  $1.653 < |\eta| < 2.6$  the endcap calorimeter is complemented with the preshower detector. This sampling calorimeter consists of two layers, each containing a lead radiator to initiate electromagnetic showers followed by silicon strip sensors of  $61 \times 1.9 \text{ mm}^2$  and  $320 \mu\text{m}$  thick [31]. The preshower detector has a thickness of 20 cm and adds around  $3 X_0$  to the radiation length. The ES helps with the identification of neutral pions and improves the position measurement of particles in the endcap regions

due to its high granularity.

The energy resolution in the electromagnetic calorimeter is parameterised in Equation (2.2) as the quadratic sum of a stochastic term  $S$ , a noise term  $N$  and a constant term  $C$ ,

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2. \quad (2.2)$$

The latter takes account of non-uniformities of the light collection in the crystals, energy leakage from the back of the crystals, and calibration errors, while  $N$  combines the influences of noise created by electronics, digitisation, and pileup [31]. The stochastic term is determined by event-to-event fluctuations. This includes variations in the number of photoelectrons released in the crystals per unit of energy, fluctuation of the energy deposited in the ES absorbers compared to what is measured in the silicon sensors, and variations in the lateral containment of showers in the crystals. For energies below 500 GeV, the terms are typically of the order of [31]

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.30\%)^2, \quad (2.3)$$

where  $E$  is expressed in GeV. For higher energies the energy leakage from the back of the crystals becomes significant and the constant term will therefore become the dominant factor in the resolution.

During operation, the transparency of the calorimeter crystals degrades due to ionisations caused when particles traverse the crystals. This induces a wavelength-dependent decrease in the light transmission, but the radiation damage can be estimated by doing regular calibrations [31].

### 2.2.3 The hadron calorimeter

A different approach is used to measure the energy of hadrons. The hadron calorimeter (HCAL) is a sampling calorimeter, consisting of non-magnetic absorber plates interleaved with plastic scintillators. As for the ECAL, there is a barrel (HB) and an endcap (HE) region, stretching from  $|\eta| < 1.3$  and  $1.3 < |\eta| < 3.0$ , respectively, complemented with detector units outside the solenoid (HO) and in the forward ( $3.0 < |\eta| < 5.2$ ) region (HF). The exact positions of the subdetectors are indicated on Figure 2.7 [31].

Being non-magnetic and sufficiently dense, 14 out of 16 absorber layers in the HB and the 18 absorber layers in the HE are made of brass. The HB is complemented with a stainless steel front and back plate. In between the absorbers 15 (17) layers of plastic scintillator tiles are installed for the HB (HE) with a granularity  $\Delta\eta \times \Delta\phi$  equal to  $0.087 \times 0.087$  for  $|\eta| < 1.6$  and  $\Delta\eta \times \Delta\phi \simeq 0.17 \times 0.17$  for  $|\eta| \geq 1.6$  [31].

The combined thickness of the HE and the EE is equivalent to about 10 interaction lengths  $\lambda_I$ . For the HB (including the EB) this is 7 to 11  $\lambda_I$ , depending on the traversed distance in the calorimeter, which increases with pseudorapidity. Since the thickness of the HB, which is severely restricted by the size of the solenoid, is not sufficient to contain all hadron showers, the HO is crucial to detect hadrons that did not deposit all of their energy before they reach the muon system described in Section 2.2.4. The HO

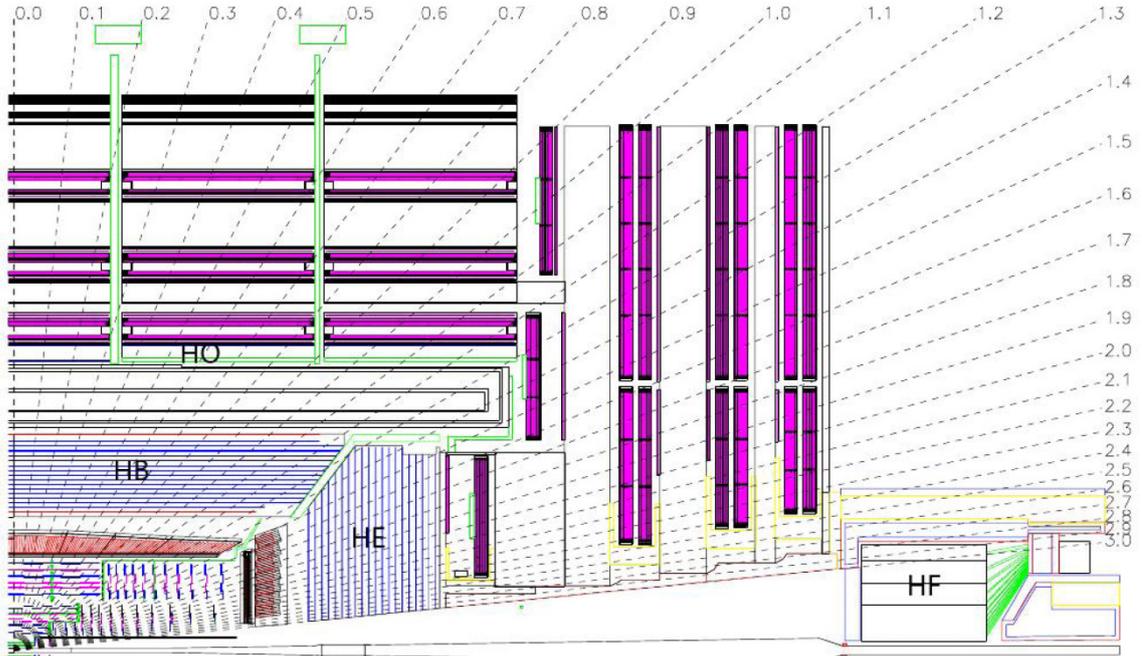


Figure 2.7: Schematic representation of a quarter of the CMS detector showing the position of the hadronic calorimeters [31], i.e. the barrel HCAL (HB), the endcap HCAL (HE), the outer HCAL (HO) and the forward HCAL (HF). The interaction point can be found in the lower left corner and the  $\eta$  range is indicated on the upper and right edges.

consists of two layers of scintillator tiles. The first one is positioned directly outside the solenoid and only covers a region up to  $|\eta| < 0.35$ . The second layer can be found on the outside of the first layer of the iron return yoke for  $|\eta| < 1.3$ . The position and size of the scintillator tiles is such that they roughly map onto the tiles in the HB. In this way, the amount of interaction lengths is increased to between 10 and  $15 \lambda_I$ .

The scintillation light produced in the HB and the HE is collected by hybrid photodiodes, which have a large gain, do not require precise temperature stabilisation and are ideal to use in strong magnetic fields [31, 85]. Silicon photomultipliers are used to read out the HO [85–87].

The HF is subject to very large particle fluxes, so radiation-hardness is a key factor in its design. It consists of quartz-fibres embedded into a steel absorber. The fibres are organised in a square grid, about 5 mm apart, parallel to the beam pipe [88]. When particles cross the fibres, Cherenkov light is created, which is read out by photomultiplier tubes [88]. These can be used because the HF is positioned outside the magnetic field of the CMS solenoid and they are shielded from the high particle fluxes by 40 cm of steel.

The hadron energy resolution is parametrised as

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + C^2, \quad (2.4)$$

where  $E$  is measured in GeV. For the barrel and endcap regions  $S = (0.847 \pm 0.016) \text{ GeV}^{1/2}$  and  $C = 0.074 \pm 0.008$ , while  $S = 1.98 \text{ GeV}^{1/2}$  and  $C = 0.09$  for the

HF [89]. As the HF primarily measures particles with high energies, its energy resolution is comparable to that of the rest of the hadron calorimeter.

## 2.2.4 The muon detectors

The muon spectrometer is located outside the CMS solenoid, interleaved in the magnet's return yoke. In order to take account of the varying magnetic and particle fluxes in different parts of the detector, multiple techniques are used to optimise the detection of muons. In the barrel, up to  $|\eta| < 1.2$ , drift tubes (DTs) are used, while cathode strip chambers (CSCs) cover the region  $0.9 < |\eta| < 2.4$ . They are complemented with resistive plate chambers (RPCs) up to  $|\eta| < 1.8$ . The position of these gaseous detectors is visualised in Figure 2.8 [90].

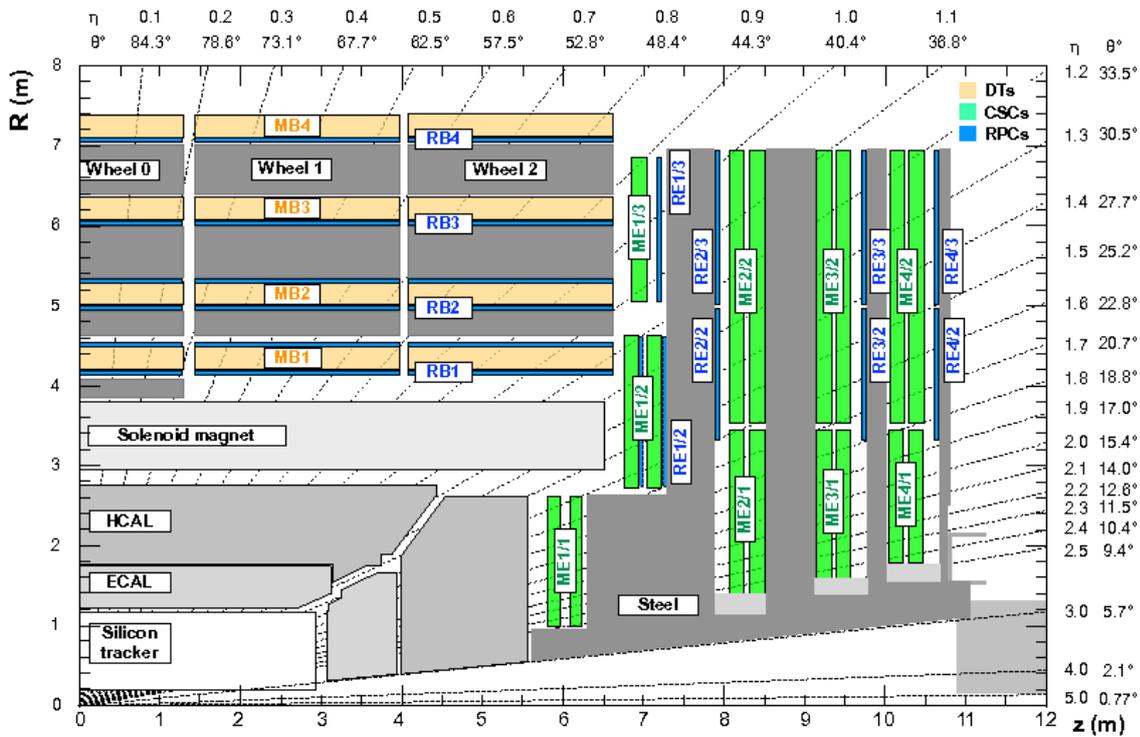


Figure 2.8: Schematic representation of a quarter of the CMS detector showing the position of the different types of muon detector [90], i.e. drift tubes (yellow), cathode strip chambers (green) and resistive plate chambers (blue). The interaction point can be found in the lower left corner and the  $\eta$  range is indicated on the upper and right edges.

As the magnetic field is uniform and the particle rate relatively low, drift tubes can be used in the barrel to observe muons [31]. A DT cell measures  $4.2 \times 1.3 \text{ cm}^2$  and is filled with an 85/15% Ar/CO<sub>2</sub> mixture. Its working principle is explained in [31]. The cells are organised such that they form layers, each shifted by half a cell width with respect to the previous one in order to avoid uninstrumented regions as much as possible. Four layers of cells together are called a superlayer (SL). As can be seen in Figure 2.8, the drift tubes are organised in four concentric stations. The first

three stations consist of three SLs, two that measure the coordinate in the  $r$ - $\phi$  plane and one that measures the  $z$  coordinate. The outermost DT station only has two SLs oriented in  $r$ - $\phi$ . One SL has a time resolution of a couple of nanoseconds [31]. The spatial resolution for a single hit is about  $250\ \mu\text{m}$  and improves to about  $100\ \mu\text{m}$  in the  $r$ - $\phi$  plane [91] for reconstructed muon segments (see Section 4.3.1).

Cathode strip chambers are used in the endcaps, because they are more radiation hard and cope better with non-uniform magnetic fields. The CSC chambers are trapezoidal in shape and they are installed in two or three concentric rings in each of the four endcap disc stations. The last station was completed during the long shutdown between LHC Run 1 and Run 2 (LS1) [87]. The size of the CSC chambers depends on their position in the CMS detector, the largest being  $3.4 \times 1.5\ \text{m}^2$ . CSCs are multiwire proportional chambers [31] containing seven panels with cathode strips that are interleaved with six layers of anode wires. The gaps between the panels are  $9.5\ \text{mm}$  wide and filled with a 40/50/10% Ar/CO<sub>2</sub>/CF<sub>4</sub> gas mixture. A two-dimensional position measurement is made by letting the cathode strips run radially outwards and the anode wires along concentric circles perpendicular to the strips, thus getting a measurement in the  $r$ - $\phi$  plane and in  $\eta$ , respectively. The width of the strips varies from  $8.4\ \text{mm}$  at the narrow end of the chamber to  $16\ \text{mm}$  at the wide end. The cathode strips are interspaced by  $0.5\ \text{mm}$  and the wires are spaced  $3.2\ \text{mm}$  apart. The spatial resolution for a single hit is estimated to be  $75 - 150\ \mu\text{m}$ . Because of the fast charge collection on the anode wires, a temporal resolution of about  $5\ \text{ns}$  is achieved [31].

Resistive plate chambers act as a redundancy in most of the muon spectrometer range. Their fast timing properties, with a resolution of  $2\ \text{ns}$ , allows to unequivocally assign a hit to a certain bunch crossing. Their spatial resolution is of the order of a centimetre [91]. The CMS detector uses double-gap bakelite RPCs in avalanche mode [31]. The  $2\ \text{mm}$  gaps between the high-resistivity plates are filled with a mixture of 96.2/3.5/0.3 C<sub>2</sub>H<sub>6</sub>F<sub>4</sub>/iso-C<sub>4</sub>H<sub>10</sub>/SF<sub>6</sub> gas. Both gaps have a common read-out system consisting of copper charge-collection strips [31]. In the barrel (endcaps) the arrangement of the RPC chambers is similar to that of the DTs (CSCs). Also the fourth endcap RPC station was installed during LS1 [87].

### 2.2.5 Trigger and data acquisition

During normal operation, the CMS experiment sees about 4 million bunch crossings per second, which leads to at least an equal amount of collisions. As the recorded data size of an average collision event is of the order of  $2\ \text{MB}$  [92], it is clear that not all of these events can be stored for further analysis. A selection is made in two stages, based on the properties of the event. The Level-1 (L1) trigger makes a first selection using hard-coded algorithms in the firmware, thus reducing the rate to  $100\ \text{kHz}$ . The events that pass the L1 trigger are reconstructed and sent to the High-Level trigger (HLT), which further reduces the rate to about  $1\ \text{kHz}$ .

The Level-1 trigger needs to decide if an event is interesting to keep in less than  $3.8\ \mu\text{s}$  [93]. Therefore, it can only use event information that is easily retrievable.

In practice, the L1 trigger will use information from the calorimeters and the muon chambers. The L1 algorithms will perform a simplified reconstruction of muons, electrons, photons, and jets. The calorimeter information is processed in so-called trigger towers, having a granularity of about  $0.087 \times 0.087$  in the barrel region. The reconstruction algorithms start as soon as a minimal amount of data is received [93]. The information from the muon subsystems is combined at an early stage and muon track finding is separated in a barrel, overlap and endcap region, in order to take account of the differences in background and magnetic field [94]. Duplicates of muons reconstructed in several regions are removed [93]. Once a provisional reconstruction is in place, a simple set of selection requirements, such as a minimum amount of transverse energy, is applied to determine if the event will be removed or kept for further analysis.

In the meantime, the complete event information is stored in a buffer. If the trigger decision does not return on time, the event is deleted. To minimise the time the electronic signals need to travel, the L1 trigger is located in the underground service cavern, close to the CMS detector itself.

When an event passes the first trigger stage, it is sent to a computer farm at the surface. There, a full event reconstruction will be made. First, the event information from all detector read-out channels, including those from the silicon tracker, will be collected and combined [92]. Then, the event is reconstructed using algorithms [95, 96] that are similar to those described in Chapter 4. Having access to the full event information, the HLT can perform complicated calculations and track down the signatures of specific event topologies. As it is software-based, the HLT algorithms and selection requirements are updated when necessary.

### **Towards a Level-1 track trigger**

As mentioned above, the L1 trigger does not use information from the tracker in order to decide which events to keep for further analysis. This is planned to change during the high luminosity phase of the LHC programme (HL-LHC), which is expected to start in 2026 [97, 98]. In order to be able to cope with the high radiation fluxes and the increased amount of pileup, the tracker will be replaced and a completely new design is envisaged. The general structure of the so-called phase-2 tracker, which is visualised in Figure 2.9, remains similar to the current one, consisting of an inner tracker, comparable to the current silicon pixel detector, and an outer tracker. Only the latter will contribute to the revised L1 trigger. The pseudorapidity range of the phase-2 tracker is extended to  $|\eta| = 4$  and its granularity is increased.

The outer tracker consists of double-layered silicon modules with a gap of 1.6 to 4.0 mm between the sensors, depending on their position. Some modules contain two layers of silicon strip sensors of about  $5 \text{ cm} \times 90 \mu\text{m}$  and others contain one layer of strips measuring around  $2.4 \text{ cm} \times 100 \mu\text{m}$  and one layer of (macro-)pixel sensors with a size of  $1.5 \text{ mm} \times 100 \mu\text{m}$ . These modules are indicated in red and blue, respectively, in Figure 2.9. Part of the modules in the inner barrel layers are tilted such that they are oriented perpendicular to the particles coming from the interaction point. This geometry increases the reconstruction efficiency at high  $|\eta|$ , while also reducing the amount of modules, and thus the amount of material, necessary to cover the sensitive

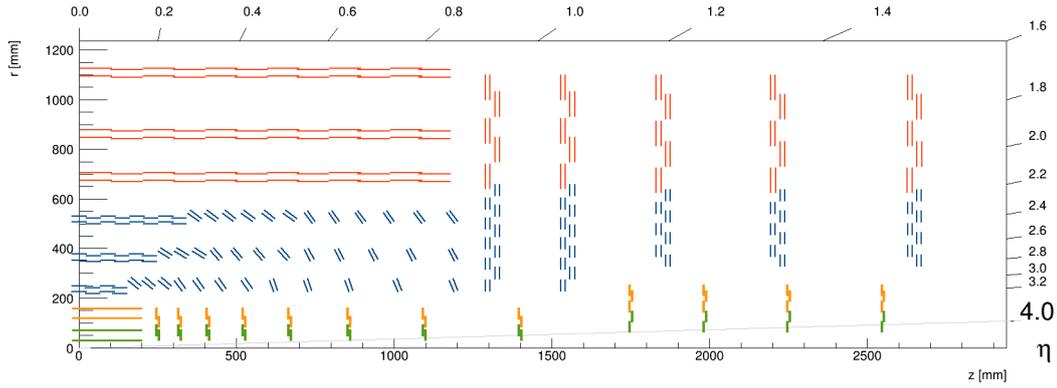


Figure 2.9: Schematic diagram of a quarter of the phase-2 tracker [99]. The inner tracker is indicated in green and orange and the outer tracker in blue and red. The interaction point can be found in the lower left corner. On the upper and right edges the pseudorapidity  $\eta$  is indicated.

volume [99–101].

The double-layered structure of the modules in the outer tracker allows to differentiate charged particles with a low and a high  $p_T$ . When a particle interacts with a module, the hits are clustered in each layer separately. Charged particles bend in a magnetic field, so the position of the outer cluster will be displaced compared to the position of the inner cluster. This displacement will be larger for low- $p_T$  than for high- $p_T$  particles, so in order to select particles with a  $p_T$  above a certain threshold, an upper limit on the displacement is installed. Since the relation between the displacement and the particle  $p_T$  depends on the distance between the two layers and the position of the module in the outer tracker, the value of this limit depends on the outer tracker region. When two clusters can be associated to each other within this ‘acceptance window’, they are collectively referred to as a stub. Figure 2.10 shows a particle whose clusters form a stub and one where the outer cluster falls outside the acceptance window. The

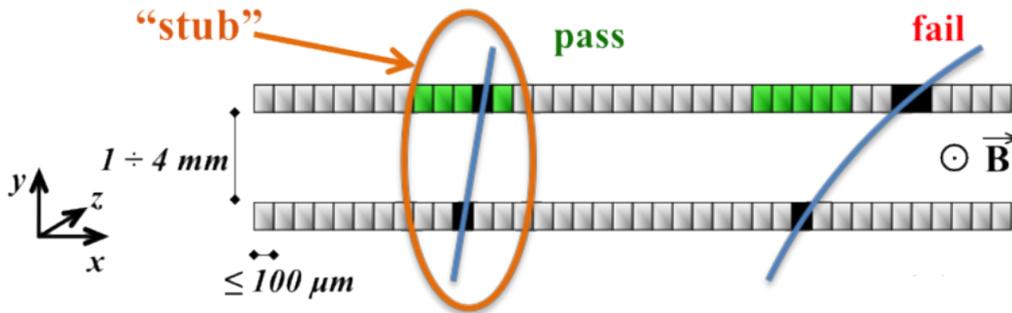


Figure 2.10: Schematic drawing of a high- $p_T$  (left) and low- $p_T$  particle (right) crossing the outer tracker [99]. The acceptance window for the displacement of the outer (upper) cluster compared to the inner (lower) cluster is indicated in green. The particle on the left is accepted as a stub, in contrast to the one on the right.

simulated distribution of the displacement per barrel layer is indicated in Figure 2.11. The boundary between the purple and the white region reflects the maximum acceptance window of each layer when the particle  $p_T$  is required to be larger than 2 GeV. Most particles in this simulation have a small displacement.

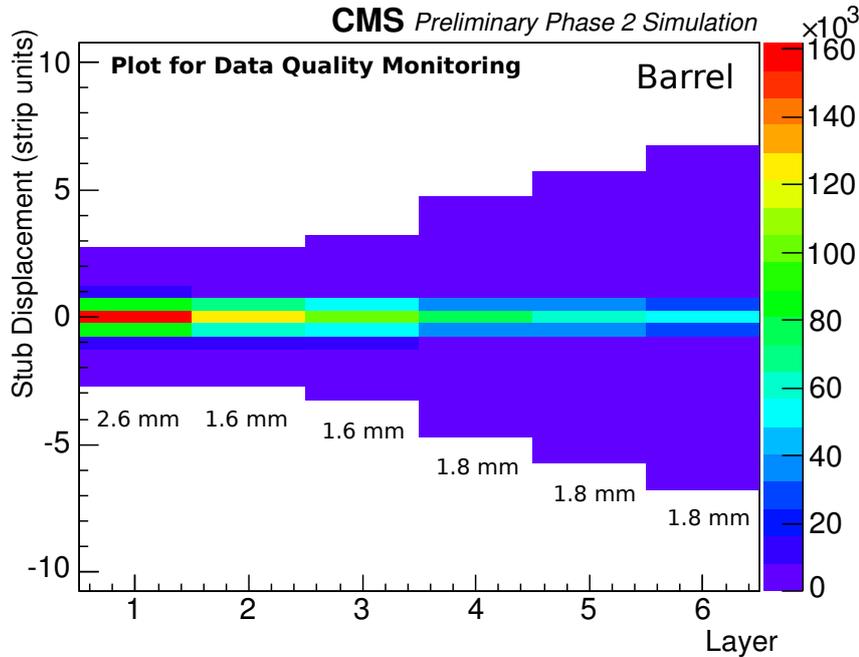


Figure 2.11: Displacement of a stub as a function of the barrel layer for simulated particles. The white regions fall outside the acceptance windows and the spacing between the sensors is indicated for each layer.

The phase-2 L1 trigger is planned to only reduce the event rate to 750 kHz during the HL-LHC programme [102]. This allows to increase the decision time to about  $12.5 \mu\text{s}$  [102]. It is expected that the reconstruction of tracks to be used by the L1 trigger takes about  $4 \mu\text{s}$  [99], so enough time remains to match these tracks to the calorimeter and muon system information. This fast reconstruction is a consequence of the stub procedure outlined above, i.e. only the tracker hits that can be related to stubs will be used for the L1 track trigger reconstruction. If the particle  $p_T$  is required to be at least 2 GeV, the number of hits is reduced by a factor 10 [99], which greatly simplifies the tracking procedure.

One of the main advantages that is expected from including the track trigger information is that it provides an increased sensitivity to the identification of isolated leptons, which is at the same time more robust for the effects of pileup [102]. The tracking information allows to better estimate the  $p_T$  of a charged particle and determine its isolation compared to the tracks of neighbouring particles. Further, due to the small sizes of the macro-pixels in the pixel-strip modules, two hits in these sensors provide enough information to estimate from which vertex the track originates [99]. This enables to only consider particles originating from a limited amount of vertices and thus reduce

the influence of the about 140–200 pileup interactions that are expected under these running conditions.

## 2.2.6 The CMS experiment during the 2016 proton run

As mentioned in Section 2.1, the LHC delivered a total integrated luminosity of  $40.8 \text{ fb}^{-1}$  during the 2016 proton run. Due to technical malfunctions, such as power supply failures and problems with the trigger system, only  $37.8 \text{ fb}^{-1}$  was recorded by the CMS detector. This still amounts to a recording efficiency of 92.5% for the 2016 proton run. Additionally, the recorded events are scrutinised by detector experts to see if they comply with basic quality standards. This further reduces the integrated luminosity that is certified to be good for analysis to  $35.9 \text{ fb}^{-1}$ , as is indicated in Figure 2.12.

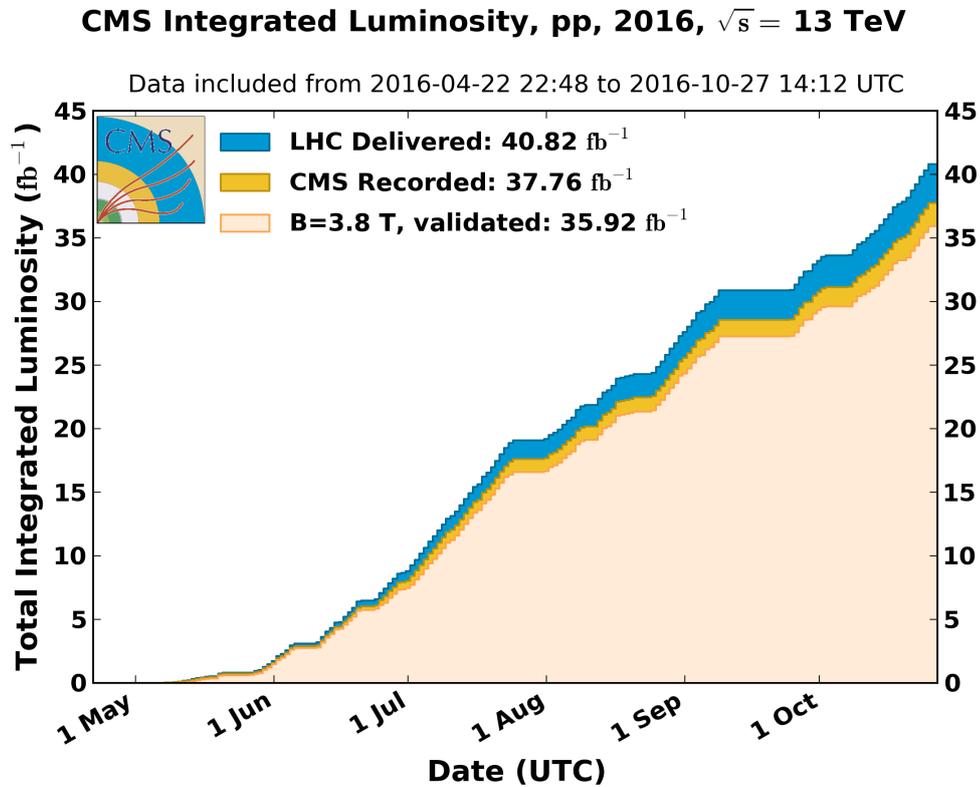


Figure 2.12: Overview of the integrated luminosity that was delivered to (blue) and recorded by (yellow) the CMS detector throughout 2016 [103]. The validated integrated luminosity that was recorded with the CMS magnet at full strength is indicated in pink.

Throughout the data-taking period, several improvements have been made to hardware and software whenever possible. These range from small changes, such as the replacement of power supplies, to major changes in the running conditions, such as the strength of the magnetic field. In order to keep track of when changes were applied, the recorded data is subdivided into so-called eras. E.g. era A contains data recorded when the CMS magnet was switched off. As this has a major influence on the reconstruction of particles, only the data recorded during eras B to H are used

in the analysis described by this thesis. Concerning the other era changes, only the change from era F to era G concerns a major update. During the previous eras, a measurement inefficiency was observed in the silicon strip tracker at high instantaneous luminosities. Fewer signal hits than expected were observed due to lingering charge depositions from previous hits causing a saturation in the read-out of the silicon sensors. From era G onwards, a reduced signal integration time increased the sensitivity of the detector immediately after recording a hit and thus recovered the otherwise missing hits. Figure 2.13 [104] shows the efficiency difference to record hits before and after the problem was resolved. The largest differences are observed in the barrel region, where some layers suffered an efficiency reduction of about 5%. The efficiency was most affected when the CMS detector was taking data at high instantaneous luminosities, which is visualised in Figure 2.13(b) for the layer that was worst affected.

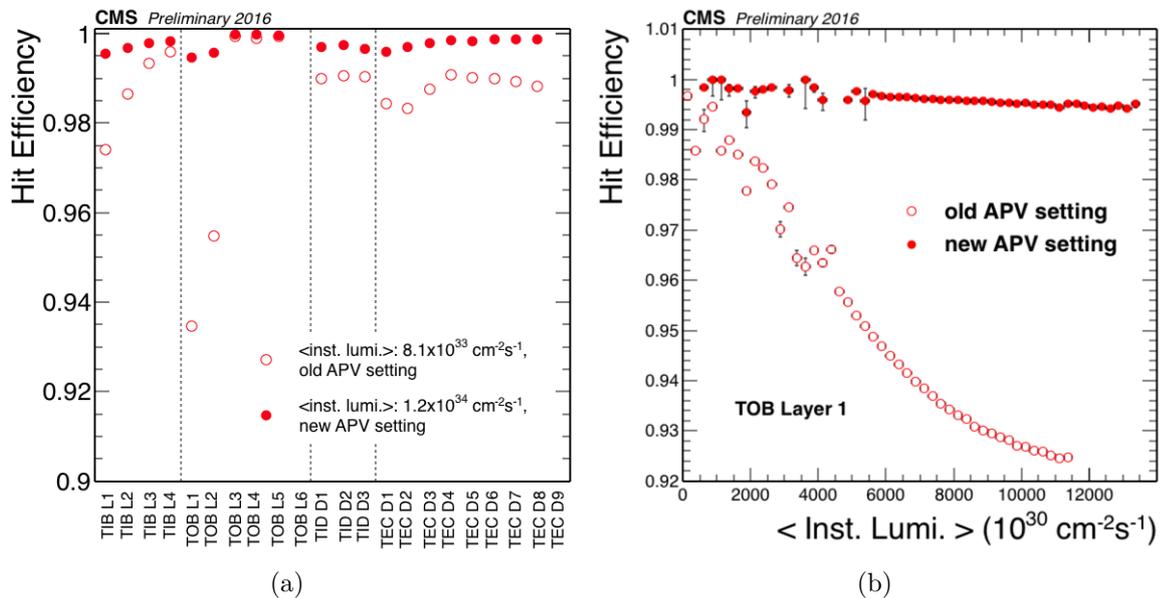


Figure 2.13: (a) Comparison of the hit efficiency during the strip tracker dynamic inefficiency problem (open circles) and after it was resolved (full circles) per layer/disc; (b) Influence of the instantaneous luminosity on the hit efficiency for TOB layer 1, which is the worst affected layer [104].

As missing hits impact the track reconstruction (see Section 4.1), the CMS collaboration developed mitigation strategies to reduce the impact of the so-called strip tracker dynamic inefficiency on the early 2016 data. As a result, the effect is expected to be negligible for this analysis.



# 3

## Event Generation and Simulation

---

When protons collide in the centre of a particle detector, only the final state of the interaction is measured. In order to make comparisons with theoretical models and SM predictions, it is necessary to understand which processes can give rise to the final state that is observed. Therefore, simulations are made of how particles are produced and how they evolve to the final state observed in the detector. Section 3.1 gives a general overview of how interactions at colliders are simulated, while Section 3.2 summarises the simulated samples relevant when studying top quarks. An alternative method for simulating samples with different top quark decay widths and masses is described in Section 3.3.

### 3.1 Simulation of Events at Proton Colliders

In order to describe particle interactions due to proton collisions, the simulation of an event is subdivided into several steps. These are outlined in Figure 3.1 and summarised below.

#### **Hard-scattering process**

As protons are composite particles, only one of their partons (quarks and gluons) will be involved in the main interaction or hard-scattering process. The two partons then interact to produce the process of interest. The differential cross section of the interaction is related to the theoretical matrix element (ME) of that process. The aspects of the hard scattering are described in more detail in Section 3.1.1.

#### **Parton shower**

The particles involved in the hard-scattering process radiate energy in the form of quarks and gluons. Additional partons from initial (ISR) and final state radiation (FSR) are included in the simulation either as part of the hard-scattering process or during the parton shower (PS). This is described in Section 3.1.2.

#### **Hadronisation and decay**

Quarks and gluons are colour charged and will hadronise to form colour-neutral particles. The models describing this so-called hadronisation are discussed in Section 3.1.3. During this step, the decay of unstable hadrons is simulated as well.

### Underlying event

The partons that are not involved in the hard-scattering process contribute to the so-called underlying event (UE), which is described in Section 3.1.4.

### Detector simulation

To simulate the interaction of particles in the detector material, the simulated events are convolved with a model of the detector. More details are given in Section 3.1.5.

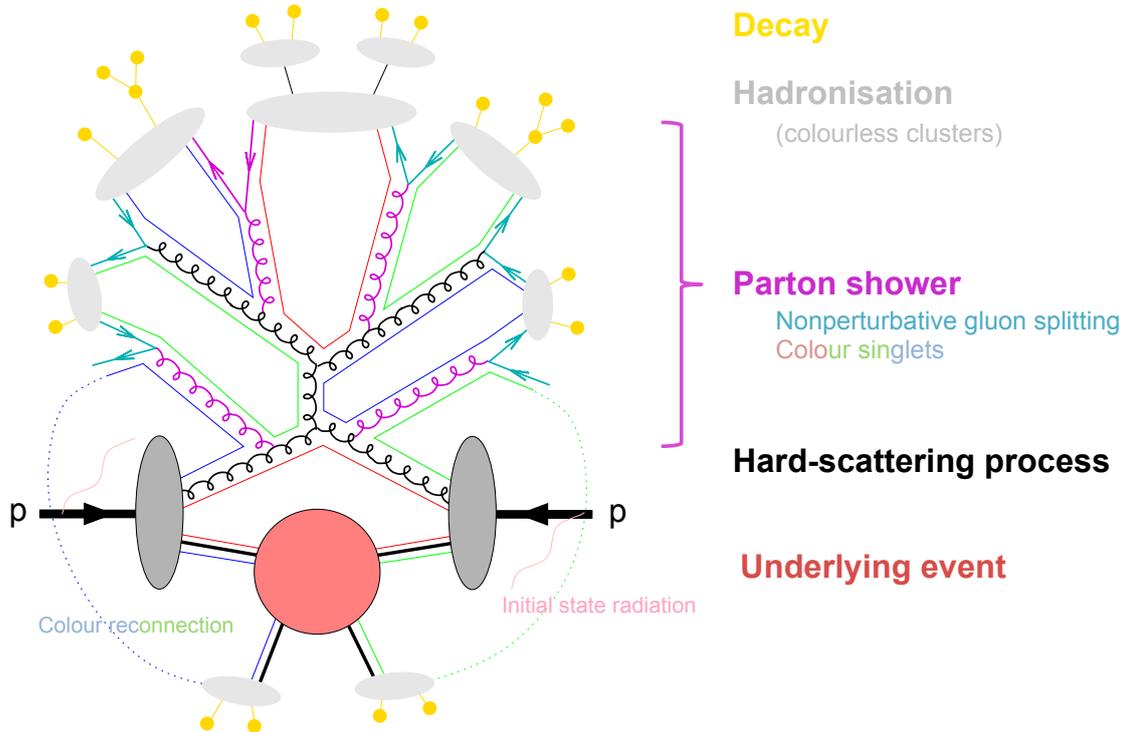


Figure 3.1: Schematic representation of a proton collision where the hard-scattering process (black) gives rise to a parton shower, which hadronises (grey) and decays (yellow). The partons that are not involved in the hard-scattering process contribute to the underlying event. Image adapted from [105].

### 3.1.1 The hard-scattering process

The partons that produce the hard interaction carry only a fraction of the momenta of the protons involved in the collision. This momentum fraction depends on the type of parton and the energy of the interaction. The way the momentum of the proton is distributed over its partons is described by the parton distribution functions (PDFs).

If two partons  $i$  and  $j$ , carrying a momentum fraction  $x_i$  and  $x_j$ , respectively, interact, the effective centre-of-mass energy of the hard-scattering process is equal to  $\sqrt{Q^2} \equiv \sqrt{\hat{s}} = \sqrt{x_i x_j s}$ . The differential cross section of a proton interaction can then

be written as a combination of the differential cross section of the parton interaction,  $d\hat{\sigma}_{ij \rightarrow X}$ , and the PDFs  $f(x_i, Q^2)$  and  $f(x_j, Q^2)$ ,

$$d\sigma_{pp \rightarrow X} = \sum_{i,j=q,\bar{q},g} \iint dx_i dx_j f(x_i, Q^2) f(x_j, Q^2) d\hat{\sigma}_{ij \rightarrow X}. \quad (3.1)$$

The scale  $Q^2$  at which Equation (3.1) is valid is called the factorisation scale, also denoted as  $\mu_F$ .

The PDFs are determined by fitting experimental data obtained from deep inelastic scattering and electroweak vector boson production at fixed target experiments. For Run 2 of the LHC the NNPDF3.0 PDF set [106] is used in the simulation. The PDFs corresponding to this set, multiplied by the momentum fraction, are visualised in Figure 3.2 for gluons, u, d and s quarks as a function of the momentum fraction. While s quarks are always virtual or sea quarks, u and d quarks can be either valence or sea quarks. As can be expected, valence quarks have a higher probability to have a fraction  $x$  of the proton momentum close to one. At an energy scale  $Q = 350$  GeV, which corresponds to about the energy needed to produce a top quark pair, it is most likely to probe a gluon from the proton, which can be seen in Figure 3.2.

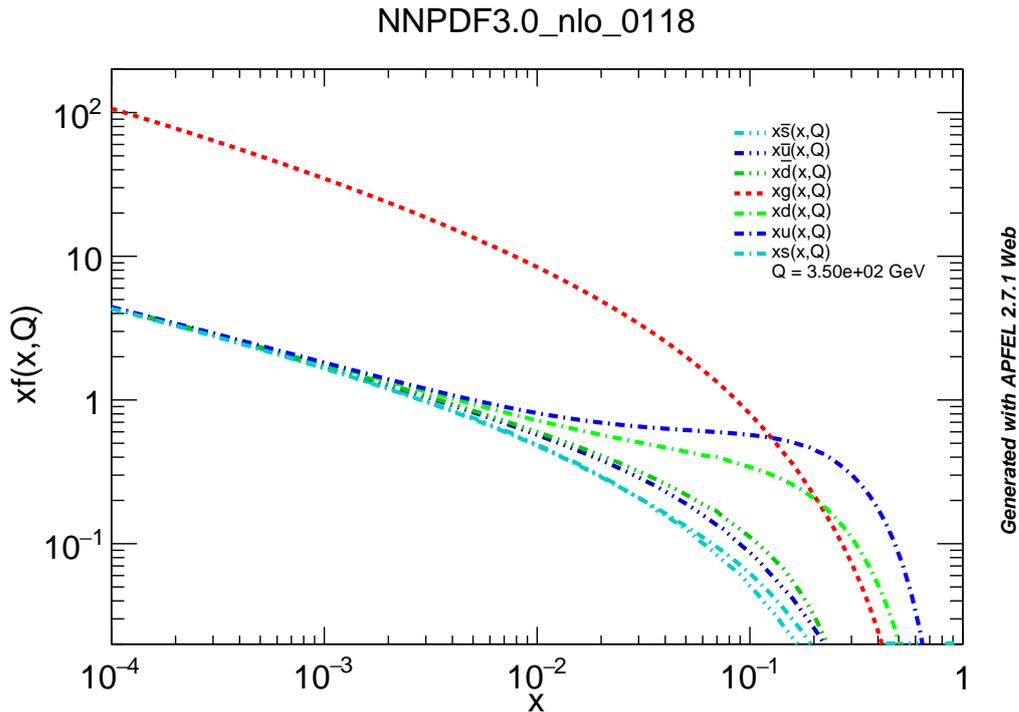


Figure 3.2: Visualisation of the parton distribution functions for the proton according to the PDF set NNPDF3.0 at an energy scale  $Q^2 = (350 \text{ GeV})^2$ . Image produced with [107].

The differential cross section of the interaction is related to the amplitude squared of the ME,  $d\hat{\sigma} \propto |\mathcal{M}|^2$ , which can be expressed as a sum of Feynman diagrams [108], and parameterises the hard-scattering process. Specialised event generators, such as MADGRAPH [109, 110] and POWHEG [111–116] calculate MEs at leading order (LO) and

next-to-leading order (NLO), respectively. When events are generated at NLO level, the hard-scattering process contains an additional real or virtual particle. The latter is visualised by a loop in the Feynman diagram, while the former can manifest itself as an additional parton.

### MadGraph/MadEvent

**MADGRAPH5**, now incorporated in **MG5\_aMC@NLO**, generates and calculates the Feynman diagrams and MEs for processes such as decays and  $2 \rightarrow n$  scatterings at LO. As the number of Feynman diagrams increases with the amount of final state particles, the number of additional partons that is supported depends on the complexity of the considered process. E.g. for top quark pair production three additional partons are supported, while this is up to five for W boson production. Using the ME and the aforementioned PDFs at LO, events are generated with **MadEvent**. It is worth noting that the differences between LO and NLO PDFs are small and diminish at increasing values of  $Q^2$  [117].

### POWHEG

The POSITIVE WEIGHT HARDEST EMISSION GENERATOR (**POWHEG**) uses NLO MEs to generate the hard-scattering process. This means that, contrary to **MADGRAPH**, also virtual emissions are taken into account. **POWHEG** generates up to one additional parton with NLO accuracy.

## 3.1.2 Parton shower

Both the initial and the final state partons of the hard-scattering process can radiate additional partons at an energy scale  $Q^2$ . A parton branches according to the mechanisms

$$\overset{(-)}{q} \rightarrow \overset{(-)}{q} g, \quad g \rightarrow q\bar{q}, \quad g \rightarrow gg,$$

and divides its energy over the two newly created partons. The probability that a parton will branch is described by the DGLAP equations [118–120]. The daughter partons can branch in their turn, thus creating a parton shower (PS). As the energy of the partons decreases with each branching, the  $Q^2$  scale at which the branching happens will also become smaller. At a scale of about 1 GeV [108] hadronisation effects will become non-negligible and a different way to model the process is needed (see Section 3.1.3).

For ISR one should also take account of the probability that the parton involved in the hard interaction originates from a parton with higher proton momentum fraction. In regions of  $x$  where the PDF falls off quickly, ISR is strongly suppressed.

Sometimes a radiated gluon reconnects with the parton it branched from, thus forming a loop. These virtual branchings are taken into account by considering the strong coupling to be a running coupling constant  $\alpha_s(Q^2)$ . The scale at which  $\alpha_s$  is evaluated is also called the renormalisation scale  $\mu_R$ .

The PS is simulated using **PYTHIA** [121] or **HERWIG++** [122, 123]. Instead of using the  $Q^2$  scale as the shower evolution parameter, **PYTHIA** simulates the PS such that the  $p_T$  of the subsequent branchings decreases, while **HERWIG++** organises the branchings according to their angle with respect to the original parton.

Since the ME and PS are simulated using different generators, a matching between the ME and the PS should be performed. When partons in the PS have a small angular separation, they will be clustered together into one object, called a jet (see also Section 4.3.3). Events will be rejected when not all partons from the ME can be matched to PS jets. If the angular separation between two partons is larger than a pre-defined threshold, they will both give rise to a jet. In this way additional jets are created compared to the final state of the ME. Another important task of the matching algorithm is to avoid double counting of processes with additional partons, since additional partons can also be produced by the ME generator. As the jets from the ME generator are better modelled, these are preferred in case there are duplicates [124].

The matching between the `MADGRAPH` ME generator and `PYTHIA` on the one hand and between `POWHEG` and `PYTHIA` or `HERWIG++` on the other hand is discussed in the following paragraphs.

### MadGraph/MadEvent

The partons from the ME are matched to jets from the PS using the MLM matching scheme [124, 125]. The partons are ordered in  $p_T$  and, starting from the parton with the highest  $p_T$ , the angular distance with respect to each of the jets is calculated. If the minimal distance is smaller than a certain threshold, the jet is matched to the parton and, in order to avoid matching the same jet multiple times, it is removed from the jet collection.

If the ME is generated with up to  $N$  additional partons, the matching procedure is performed separately for each process with  $n = 0 \dots N$  additional partons. This ensures that all possible configurations are studied. Double counting is avoided by requiring that the number of jets is equal to the amount of partons when  $n < N$ .

### POWHEG

As mentioned before, `POWHEG` generates up to one additional parton with NLO accuracy. Subsequent emissions are generated during the PS stage. By definition, the additional ME parton is the hardest emission and its  $p_T$  is considered to be an upper limit for the shower evolution scale. This means that events generated by `POWHEG` can be directly interfaced to  $p_T$ -ordered PS generators, such as `PYTHIA 8`. Extra care is needed to interface `POWHEG` to e.g. `HERWIG++`, which produces angular-ordered showers. These types of PS generators often produce large-angle soft emissions before the hardest emission in the PS, so extra steps are needed to avoid PS emissions with a higher  $p_T$  than that of the additional ME parton [111, 112, 126].

When the  $Q^2$  scale of the hard-scattering process is large, also the soft radiation can have a relatively high  $p_T$  (it is soft compared to the hard scattering). In order to compensate for an overestimation of high- $p_T$  contributions from radiation, a damping factor is introduced which takes values between 0 and 1. In `POWHEG` this takes the form,

$$\frac{h^2}{p_T^2 + h^2},$$

where  $h$ , also called `hdamp`, is usually set equal to the scale of the hard-scattering process [127, 128].

### 3.1.3 Hadronisation and decay

Phenomenological models are used to describe how the colour-charged particles originating from the PS will form the colour-neutral hadrons that are observed. This is also called fragmentation. In `PYTHIA 8`, which is used to perform the hadronisation, the Lund string model is considered [108, 129].

Two partons with opposite colour charges are connected by a string, which is visualised as a colour flux tube in Figure 3.3 [130]. The potential energy stored in the string is linearly proportional to the distance  $r$  between the partons and is defined as  $V(r) = -\kappa r$ , with  $\kappa \approx 1 \text{ GeV/fm}$  [121]. When the quark and the antiquark move apart, the potential energy increases. At a certain point sufficient energy is available to create another quark-antiquark pair,  $q'\bar{q}'$ . The newly created quarks are re-arranged into two colour singlets,  $q\bar{q}'$  and  $q'\bar{q}$ , which can each move apart and create more quark-antiquark pairs. Hence  $n$  hadrons are formed from the re-arranged quark-antiquark pairs  $q\bar{q}_1$ ,

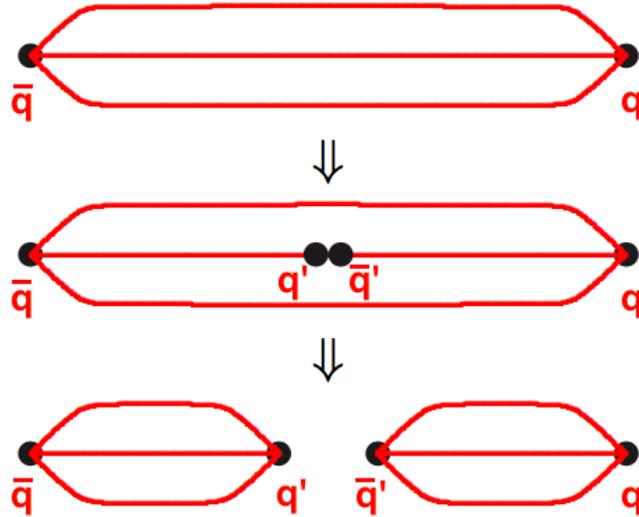


Figure 3.3: Visualisation of a Lund string between two oppositely (colour) charged partons being stretched and breaking into two quark-antiquark pairs. In this particular case the quark-antiquark pair  $q\bar{q}$  is connected through a string with a red colour charge. Image adapted from [130].

$q_1\bar{q}_2, q_2\bar{q}_3, \dots, q_{n-1}\bar{q}_n$ . The order in which the string breaks occur is irrelevant as they are causally independent [121]. The probability to end up with a certain  $n$  hadron state is thus equal to the product of the probabilities for the  $n - 1$  string breaks and the functions that fix the  $n$  hadron masses [108]. As the probability to create a quark-antiquark pair is proportional to  $1/\exp(m_q^2)$ , the production of heavy quarks is strongly suppressed.

Baryons are produced when a string break results in a diquark-antidiquark pair. The probability for this to happen is about a tenth of the probability for creating a quark-antiquark pair [108].

An alternative approach is the cluster model, as implemented in `HERWIG++`. At the end of the PS, all remaining gluons are split into quark-antiquark pairs, which favours the production of light quarks. The quarks are then clustered according to their colour

connections with other quarks. An example of the colour flow in a PS can be found in Figure 3.4 [108]. If a parton has few emissions during the PS stage, the cluster containing it can have a very large momentum. Similar to the Lund string model, it will split into two smaller clusters through an intermediate quark-antiquark pair. The clusters then decay isotropically into hadrons [122].

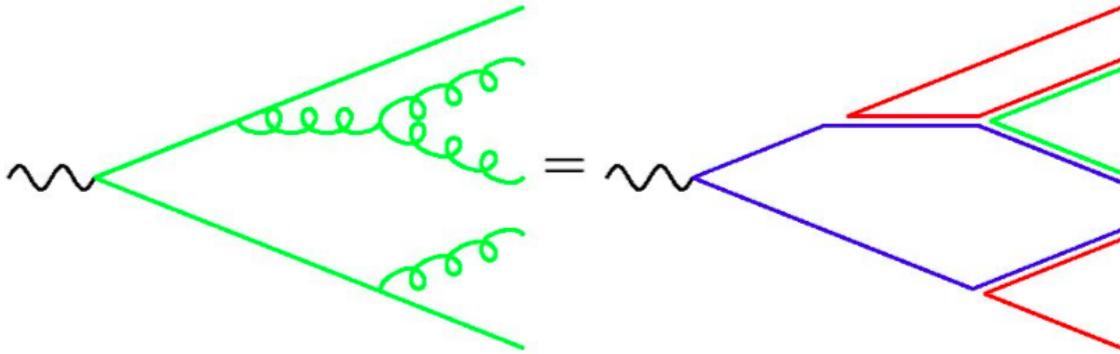


Figure 3.4: Schematic drawing of the colour flow in a parton shower [108]. If the cluster hadronisation model is applied to this example, gluons are first split into quark-antiquark pairs. Then, the colour-connected quarks form, in this case, four colourless clusters.

Many of the hadrons are unstable and decay to the final state that is observed in the detector. It is important to note that very heavy quarks, such as the top quark, have such a short lifetime that they will decay before they can hadronise. In this case, their decay products are often considered as a part of the hard-scattering process.

### 3.1.4 Underlying event

All activity that is not connected to the hard-scattering process is amalgamated in the so-called underlying event (UE). Many parameters control the modelling of the UE and these are tuned by comparisons to data. A specified set of parameters that describe the data well is called a ‘tune’. In this thesis the CUETP8M1 and CUETP8M2T4 tunes are used [131, 132]. The latter is used for the simulation of top quark pair events, as for high jet multiplicities differences between data and simulation were observed with the former tune. Both tunes are very similar and only have slight variations in a couple of parameters.

A number of the aspects impacting the UE description are discussed below.

#### Proton remnants

Up till now only the partons involved in the hard interaction have been considered. In order to simulate a complete proton collision, the activity of the proton remnants needs to be described as well. As the partons in the proton remnants carry a non-negligible amount of momentum and are colour charged, they will also undergo parton showering and hadronisation. This process is not completely independent from the hard-scattering process since all partons are colour connected.

## Multiparton interactions

Remnant partons can also interact amongst themselves before showering and hadronising. These multiparton interactions (MPIs) are mostly soft interactions, though multiple hard-scattering processes are also possible [121]. MPIs are very common in proton collisions and are enhanced by requesting a hard-scattering process, which gives a bias towards head-on proton collisions [108]. Since the proton can be seen as a collection of partons moving in the same direction, the probability for MPIs increases when the overlap area of the collision is large. In fact, PYTHIA 8 considers all proton collisions to be MPIs, parameterised by the number of parton interactions [121]. When modelling the MPIs, momentum and colour conservation is taken into account.

## Colour reconnection

In the Lund model all colour-charged particles are connected by strings. When there are many partons, e.g. by the presence of many MPIs, the colour flows can be mixed. When a string connecting two quarks is stretched, it can be represented by an infinite amount of colour-charged soft gluons. When two strings overlap, they can interact and the colour encompassed in the strings is redistributed over the system. This can be modelled in different ways. Below, some of the models implemented in PYTHIA 8 are described [133, 134]. So as to be able to describe the variations depending on resonance decays, it is assumed that at least one top quark is produced in the hard-scattering process.

### MPI-based model

In order to minimise the total string length  $\lambda$ , partons from lower- $p_T$  MPIs are added to the strings from higher- $p_T$  MPIs. Starting from the lowest- $p_T$  system, the reconnection probability with higher- $p_T$  systems is calculated and this is repeated for all systems in order of increasing  $p_T$  until the optimal  $\lambda$  is achieved. Then, the merging is performed starting from the highest- $p_T$  system. As low- $p_T$  systems have a larger spatial area, the probability that it overlaps with another system is much higher than for high- $p_T$  systems. Therefore, low- $p_T$  systems are easily merged with others, while it is much harder to merge two high- $p_T$  systems.

In the default version of this model it is assumed that colour reconnection occurs before the top quark decays. The decay products are thus not involved in the reconnection. Another option is to assume early resonance decays (ERD), such that also the top quark decay products are included in the colour reconnection.

### QCD-based model

Whereas the MPI-based model only considers a colour structure where the number of colours  $N_c = 3$ , the QCD-based model allows extensions using the SU(3) colour algebra [135]. A set of nine additional colour indices are defined that are not related to the nominal indices, often referred to as ‘red’, ‘blue’ and ‘green’. A quark gets a colour index, an antiquark an anticolour index and a gluon gets one of each, where the latter index has a different value than the former index [136]. Two partons can be connected by a string when their colour indices match. In

this way, more reconnections are possible compared to the MPI-based model. Similarly, a reconnection is performed when this reduces the total string length.

### Gluon-move model

A gluon on the string connecting a quark-antiquark pair can be moved to another quark-antiquark pair when the move decreases the total string length. This is repeated as long as  $\lambda$  can be minimised and may involve moving the same gluon multiple times. It is also possible to introduce a ‘flip’ such that a quark can be reconnected to the antiquark of a different quark-antiquark system and the other way around.

## 3.1.5 Simulation of the CMS detector

In order to give a more realistic description of the simulated events, such that they have a closer resemblance to the recorded data, they are passed through a simulation of the CMS detector, which is made using the **GEANT4** toolkit [137, 138]. First, all of the detector elements are simulated in great detail. For e.g. each tracker pixel or calorimeter cell logical blocks describing their size and material are built and these are assigned a physical position with respect to the interaction point. This is also done for the support structures, cooling pipes, etc. that are present in the detector volume. In this way a model of the detector is created that shows both the regions that are sensitive to detect particles and those that are not (due to the presence of non-sensitive material or gaps between modules). A detailed description of the magnetic field is added as well, since this will have a large influence on how the particles move through the detector.

Then, the particles created during the various simulation steps described in the previous sections are transported through the detector. For each detection layer, the interaction of particles with the material is simulated, taking account of the detection efficiency of each sensor and possible unwanted interactions, such as multiple scattering and Bremsstrahlung. This information is passed on to the next layer until the entire detector volume has been covered. The result is a collection of simulated detector hits and their electronic response, much like the data would produce, which has the advantage that both data and simulation can be reconstructed using the same algorithms (see Chapter 4).

Also pileup is added to the event description at this point. The hits produced by the additional proton interactions mix with those of the main collision and complicate the event reconstruction.

The detector convolution is by far the most time-intensive step of the simulation.

## 3.2 Simulated Events for Top Quark Studies

The simulation sequence described above can be applied to any hard-scattering process. This thesis will focus on top quark pair production, or  $t\bar{t}$ , in the semileptonic decay channel. Events are generated with **POWHEG** at NLO using the UE tune **CUETP8M2T4** and the PS is performed with **PYTHIA 8**. The top quark pair production cross section, which is calculated at NNLO, is equal to  $832_{-46}^{+40}$  pb at 13 TeV [139].

Apart from  $t\bar{t}$  events, also other hard-scattering processes that produce the same final state have to be taken into account. In this case, background events consist of singly produced top quarks, also referred to as ‘single top’ (ST), and leptonically decaying vector bosons produced with additional jets. Table 3.1 gives an overview of the simulated samples, the generators used for their production, the amount of events generated for each sample, and their production cross sections at 13 TeV. The UE tune used in the generation of the background events is CUETP8M1 and also for these processes showering and hadronisation is performed with PYTHIA 8. Single top events are generated with POWHEG and the cross section is  $217_{-8}^{+9}$  pb for t channel production and  $72 \pm 4$  pb for single top quark production in association with a W boson, both calculated at NLO [140–142]. In practice, samples are generated separately for singly produced top and antitop quarks and the respective cross sections for these processes are mentioned in Table 3.1. W + jets and Drell–Yan (DY) + jets events are generated with MADGRAPH5 at LO with up to four additional partons in the ME calculation. As can be seen in Table 3.1, the cross sections become smaller when the processes are generated with more additional partons.

The number of events for a certain production process is determined as the product of the production cross section and the collected integrated luminosity,  $N = \sigma \cdot L$ . In order to minimise statistical fluctuations in the simulation, the number of simulated events should be larger than the amount that is measured. Hence, the equivalent luminosity, defined as  $L_{eq} = N'/\sigma$ , where  $N'$  is the number of events in the simulated sample, should be larger or equal to the integrated luminosity of the data sample. The higher  $L_{eq}$ , i.e.  $N'$ , the smaller the statistical fluctuations in the simulation when comparing to data. Table 3.1 shows that the top quark samples have an equivalent luminosity that is at least five times higher than the integrated luminosity of the data. By contrast, the integrated luminosities of the W + 1jet and the W + 2jets samples are smaller than the integrated luminosity of the data. However, these samples will not be used in the analysis, since the number of selected events becomes negligible after applying the selection requirements proposed in Chapter 5.

Systematic uncertainties arise due to uncertainties in the modelling of the simulation, e.g. variations in the amount of ISR/FSR that is expected or in how the ME is matched to the PS. In order to take account of these uncertainties, additional samples have been produced where several parameters related to certain modelling aspects are varied with respect to those in the nominal  $t\bar{t}$  sample. The size of each variation is chosen such that it describes the modelling uncertainty when the simulation is compared to data. These supplementary samples are summarised in Table 3.2. In addition, other systematic effects are implemented on an event-weight basis. This is e.g. the case for systematic effects related to experimental uncertainties such as the uncertainty on the integrated luminosity of the data sample. A full treatment of the systematic uncertainties, applied to the analysis described in this thesis, can be found in Section 6.3.

Table 3.1: Overview of the simulated samples at  $\sqrt{s} = 13$  TeV. Showering and hadronisation is performed with PYTHIA 8.

Dataset	Generator	$\sigma$ (pb)	# events	$L_{eq}$ (fb $^{-1}$ )
$t\bar{t}$ + jets $m_t = 172.5$ GeV	POWHEG tune CUETP8M2T4	831.76	154.7 M	185.9
Single top	POWHEG tune CUETP8M1			
t channel t		136.02	65.6 M	193.9
t channel $\bar{t}$		80.95	20.0 M	193.4
tW channel t		35.85	7.0 M	481.9
tW channel $\bar{t}$		35.85	6.9 M	468.3
DY + jets	MADGRAPH tune CUETP8M1			
DY + 1 jet		1016.0	61.4 M	60.4
DY + 2 jets		331.4	20.0 M	60.3
DY + 3 jets		96.4	5.9 M	60.7
DY + 4 jets		51.4	4.2 M	81.7
W + jets	MADGRAPH tune CUETP8M1			
W + 1 jet		9493.0	45.4 M	4.8
W + 2 jets		3120.0	60.2 M	19.3
W + 3 jets		942.3	55.2 M	58.5
W + 4 jets		524.2	27.9 M	53.3

Table 3.2: Overview of the simulated samples that are used to estimate some of the systematic uncertainties due to modelling.

Dataset	Generator	$\sigma$ (pb)	# events	$L_{eq}$ (fb <sup>-1</sup> )
$t\bar{t}$ + jets	POWHEG + PYTHIA 8 tune CUETP8M2T4	831.76		
<i>Shower scales</i>				
ISR up			59.0 M	71.0
ISR down			58.8 M	70.7
FSR up			59.1 M	71.1
FSR down			59.2 M	71.2
<i>ME-PS matching</i>				
hdamp up			58.7 M	70.6
hdamp down			58.1 M	69.9
<i>Underlying event</i>				
tune up			58.8 M	70.7
tune down			58.3 M	70.1
<i>Colour reconnection</i>				
MPI (ERD)			59.9 M	72.0
QCD-based (ERD)			59.6 M	71.7
Gluon move			59.0 M	71.0
Gluon move (ERD)			56.2 M	67.5

### 3.3 Reweighting Procedure for Simulated Samples

The  $t\bar{t}$  events used in this thesis are generated with POWHEG using  $M_{t,\text{gen}} = 172.5$  GeV and  $\Gamma_{t,\text{gen}} = 1.31$  GeV as the top quark mass and decay width, respectively, but it is possible to generate samples with different combinations of  $M_{t,\text{gen}}$  and  $\Gamma_{t,\text{gen}}$ . The CMS collaboration produced a number of simulated samples with non-nominal masses and/or widths, which are summarised in Table 3.3. The samples that were generated with a different decay width all have  $M_{t,\text{gen}} = 172.5$  GeV. When an alternative value was used for the top quark mass, the decay width was adapted accordingly (see Section 1.2.3).

Table 3.3: Overview of the simulated samples that are used to validate the reweighting procedure.

Dataset	Generator	$\sigma$ (pb)	# events	$L_{eq}$ ( $\text{fb}^{-1}$ )
$t\bar{t}$ + jets	POWHEG + PYTHIA 8 tune CUETP8M2T4	831.76		
<i>Top quark width</i>				
	$\Gamma_{t,\text{gen}} = 0.2 \times 1.31$ GeV		19.9 M	23.9
	$\Gamma_{t,\text{gen}} = 0.5 \times 1.31$ GeV		19.9 M	23.9
	$\Gamma_{t,\text{gen}} = 0.8 \times 1.31$ GeV		18.5 M	22.3
	$\Gamma_{t,\text{gen}} = 2 \times 1.31$ GeV		13.5 M	16.2
	$\Gamma_{t,\text{gen}} = 4 \times 1.31$ GeV		19.6 M	23.6
	$\Gamma_{t,\text{gen}} = 8 \times 1.31$ GeV		19.5 M	23.5
<i>Top quark mass</i>				
	$M_{t,\text{gen}} = 169.5$ GeV ( $\Gamma_{t,\text{gen}} = 1.23$ GeV)		58.5 M	70.4
	$M_{t,\text{gen}} = 171.5$ GeV ( $\Gamma_{t,\text{gen}} = 1.28$ GeV)		19.6 M	23.5
	$M_{t,\text{gen}} = 173.5$ GeV ( $\Gamma_{t,\text{gen}} = 1.34$ GeV)		19.3 M	23.2
	$M_{t,\text{gen}} = 175.5$ GeV ( $\Gamma_{t,\text{gen}} = 1.39$ GeV)		59.3 M	71.2

#### 3.3.1 Reweighting of the simulated top quark width

Table 3.3 contains a couple of  $t\bar{t}$  samples that have been generated with a different value for the top quark decay width. The equivalent integrated luminosity of these samples, however, is about half the size of the integrated luminosity of the data sample. Therefore, the top quark decay width of the nominal  $t\bar{t}$  sample will be reweighted and the generated samples are used to validate the procedure. Reweighting the nominal sample has the additional advantage that it is possible to ‘create’ a sample with nearly any desired value of the top quark decay width.

In order to reweight the events to obtain a sample with a different value of the top quark decay width, the distribution of the top quark mass is studied in more detail. The mass of the generated top quark  $m_t$  follows a Breit–Wigner distribution,

$$\text{BW}(m_t, s) = \frac{1}{\pi} \frac{s\Gamma_{t,\text{gen}}/2}{(m_t - M_{t,\text{gen}})^2 + (s\Gamma_{t,\text{gen}}/2)^2}, \quad (3.2)$$

where  $M_{t,\text{gen}}$  and  $\Gamma_{t,\text{gen}}$  represent the top quark mass and decay width used in the generation and  $s$  is a scale factor introduced for the reweighting. Events are generated with  $s = 1$  and the top quark decay width can be reweighted by varying the parameter  $s$ . An event weight is constructed by dividing the result of Eq. (3.2) evaluated for a certain value of  $s$  by the nominal value for  $s = 1$ . This is done for both the top quark and the antitop quark present in the  $t\bar{t}$  event,

$$\text{weight}(m_t, m_{\bar{t}}, s) = \frac{\text{BW}(m_t, s)}{\text{BW}(m_t, 1)} \cdot \frac{\text{BW}(m_{\bar{t}}, s)}{\text{BW}(m_{\bar{t}}, 1)}. \quad (3.3)$$

Figure 3.5 shows the top quark mass distribution for the nominal  $t\bar{t}$  sample, where the event weights have been calculated for several values of  $s$ . The expected behaviour is observed. When the event weights are constructed with  $s < 1$ , the top quark mass distribution is narrower and has shorter tails, which corresponds to a smaller top quark width. The opposite occurs when  $s > 1$ .

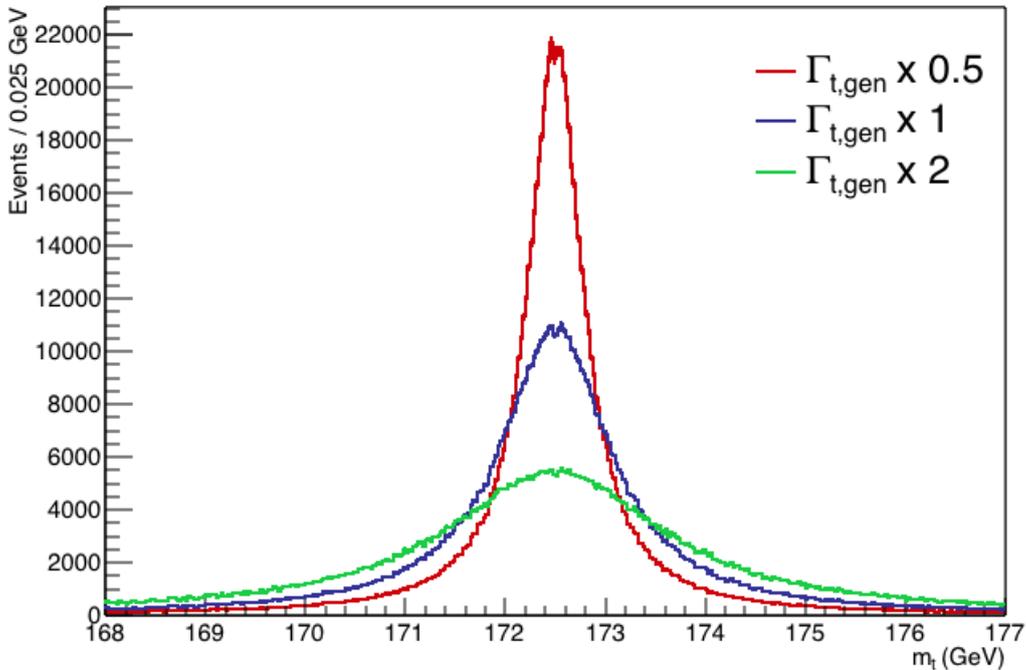


Figure 3.5: Distribution of the top quark mass for top quark decay widths reweighted with  $s = 0.5$  (red),  $s = 1$  (blue), and  $s = 2$  (green).

In Figure 3.6 the reweighted distributions are compared to the top quark mass distributions obtained using the samples that were generated with top quark widths

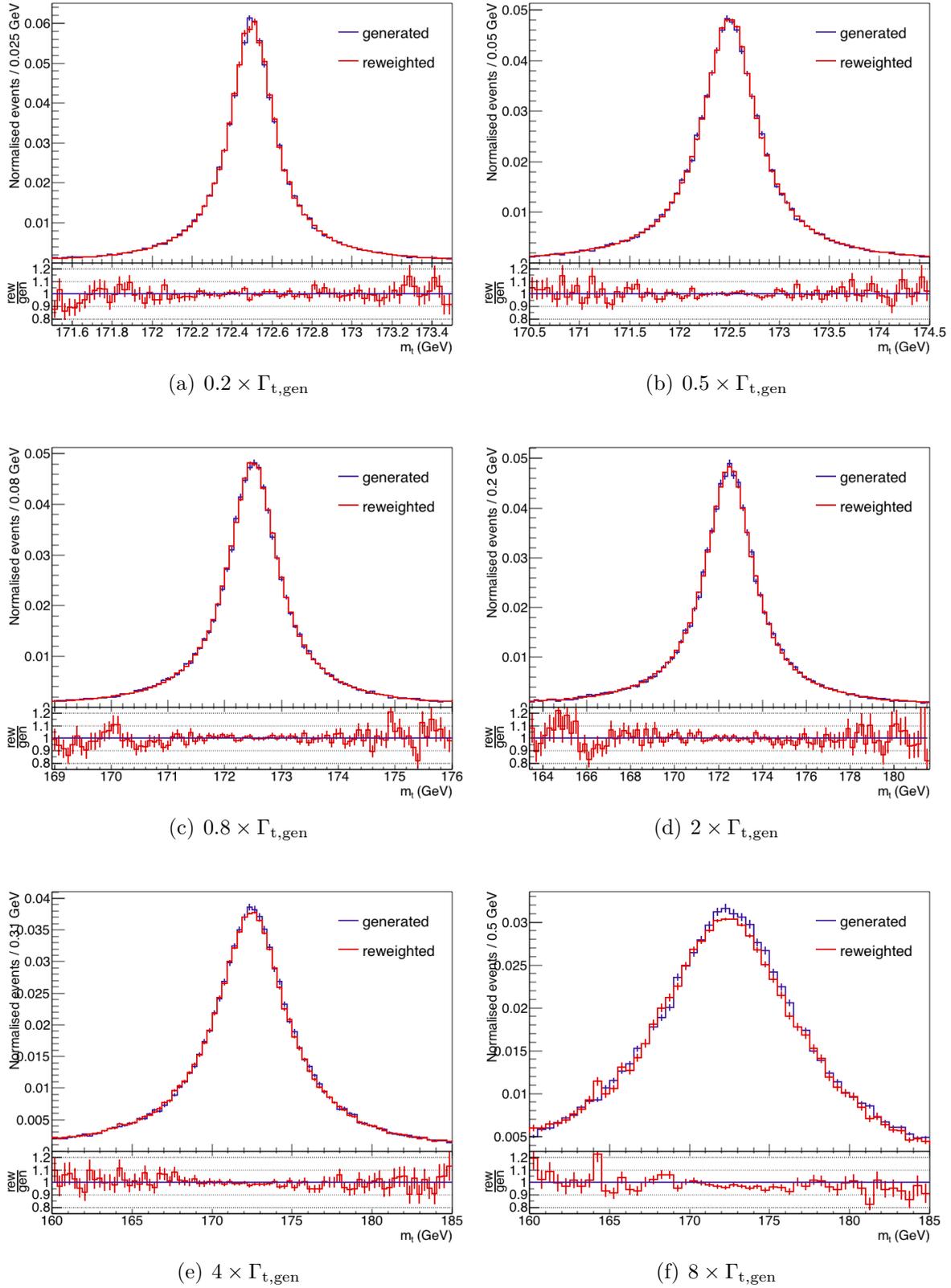


Figure 3.6: Comparison of the normalised distributions of the top quark mass for samples generated with a different top quark decay width (blue) and the reweighted samples (red) with a top quark width of (a)  $0.2 \times \Gamma_{t,gen}$ , (b)  $0.5 \times \Gamma_{t,gen}$ , (c)  $0.8 \times \Gamma_{t,gen}$ , (d)  $2 \times \Gamma_{t,gen}$ , (e)  $4 \times \Gamma_{t,gen}$ , and (f)  $8 \times \Gamma_{t,gen}$ .

different from the SM one. For top quark decay widths up to  $4 \times \Gamma_{t,\text{gen}}$  ( $s = 4$ ) the two distributions agree within the statistical uncertainty of the simulated samples. For  $8 \times \Gamma_{t,\text{gen}}$  ( $s = 8$ ) some deviations can be seen in the peak of the distribution, as well as instabilities in the tails. When the width of the distribution increases, the limited amount of events in the tails of the nominal distribution get larger event weights, resulting in statistical fluctuations. Therefore, the top quark decay width will only be reweighted up to  $s = 6$ .

### 3.3.2 Reweighting of the simulated top quark mass

Events can also be reweighted to change the value of the top quark mass. This is done by replacing  $M_{t,\text{gen}}$  in Eq. (3.2) by the desired  $M_t$  value. Contrary to reweighting the top quark width, where  $s$  can take nearly any value, the top quark mass reweighting only gives sensible results when the difference between  $M_t$  and  $M_{t,\text{gen}}$  is limited. As the position of the top quark mass distribution changes during the reweighting, the events in the peak of the distribution are scaled down, while the restricted amount of events in one of the tails get an event weight that is larger than one. If  $|M_t - M_{t,\text{gen}}|$  is large, statistical fluctuations become significant and the new distribution will be distorted. This is visualised in Figure 3.7, where the nominal  $t\bar{t}$  sample, in the centre, is reweighted with a mass difference of  $\pm 1$  GeV and  $\pm 3$  GeV, respectively. Whereas the uncertainties on the reweighted distributions obtained by shifting  $M_{t,\text{gen}} = 172.5$  GeV by 1 GeV are moderate, they are non-negligible when the shift is 3 GeV. Furthermore, it can be seen that the statistical fluctuations of the reweighted distributions are larger in the tails that are farthest away from the nominal distribution.

The reweighted top quark mass distributions cannot be directly compared to the generated samples in Table 3.3, as these also have a different generated width. Therefore, the widths of the generated samples are first reweighted such that they match the decay width of the nominal distribution, i.e.  $\Gamma_t = 1.31$  GeV. Figure 3.8 shows the normalised top quark mass distributions of the samples generated with a different  $M_t$  before and after reweighting the top quark decay width. The distributions where the mass is reweighted are compared to those of the generated samples in Figure 3.9. This comparison confirms the ideas expressed before, i.e. it is safe to reweight the top quark mass of the nominal sample as long as  $M_t$  and  $M_{t,\text{gen}}$  differ by a small amount.

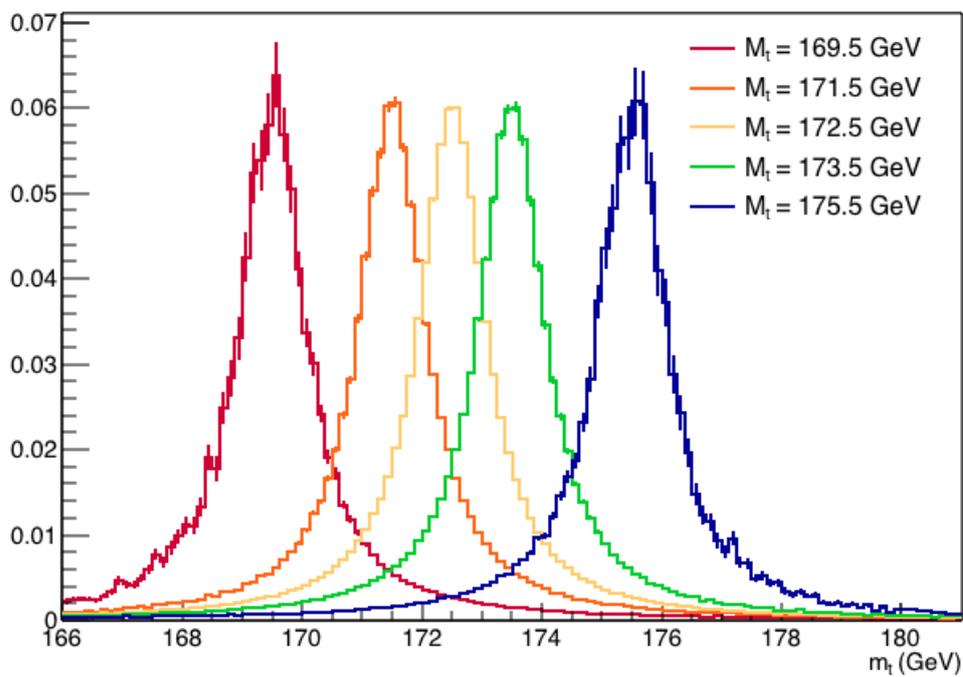
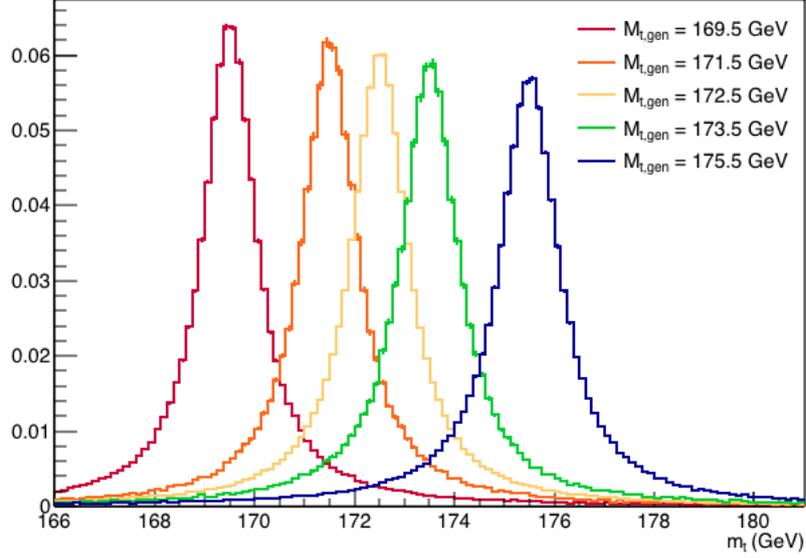
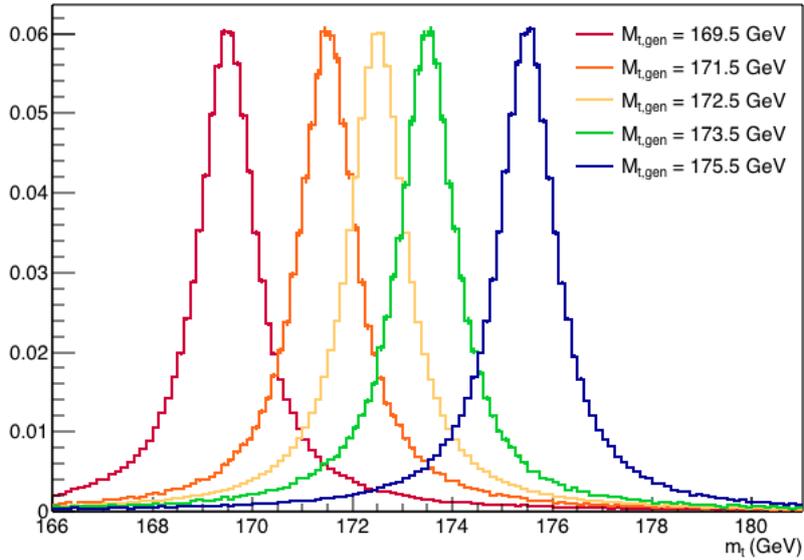


Figure 3.7: Comparison of the normalised distributions of the top quark mass when the nominal  $t\bar{t}$  sample, generated with  $M_{t,\text{gen}} = 172.5$  GeV (yellow), is reweighted to a top quark mass  $M_t = 169.5$  GeV (red),  $M_t = 171.5$  GeV (orange),  $M_t = 173.5$  GeV (green), and  $M_t = 175.5$  GeV (blue).



(a)



(b)

Figure 3.8: Normalised distributions of the top quark mass for samples generated with a top quark mass  $M_{t,\text{gen}} = 169.5 \text{ GeV}$  (red),  $M_{t,\text{gen}} = 171.5 \text{ GeV}$  (orange),  $M_{t,\text{gen}} = 172.5 \text{ GeV}$  (yellow),  $M_{t,\text{gen}} = 173.5 \text{ GeV}$  (green), and  $M_{t,\text{gen}} = 175.5 \text{ GeV}$  (blue). The widths of the distributions are (a) equal to those indicated in Table 3.3 or (b) reweighted to  $\Gamma_t = 1.31 \text{ GeV}$ .

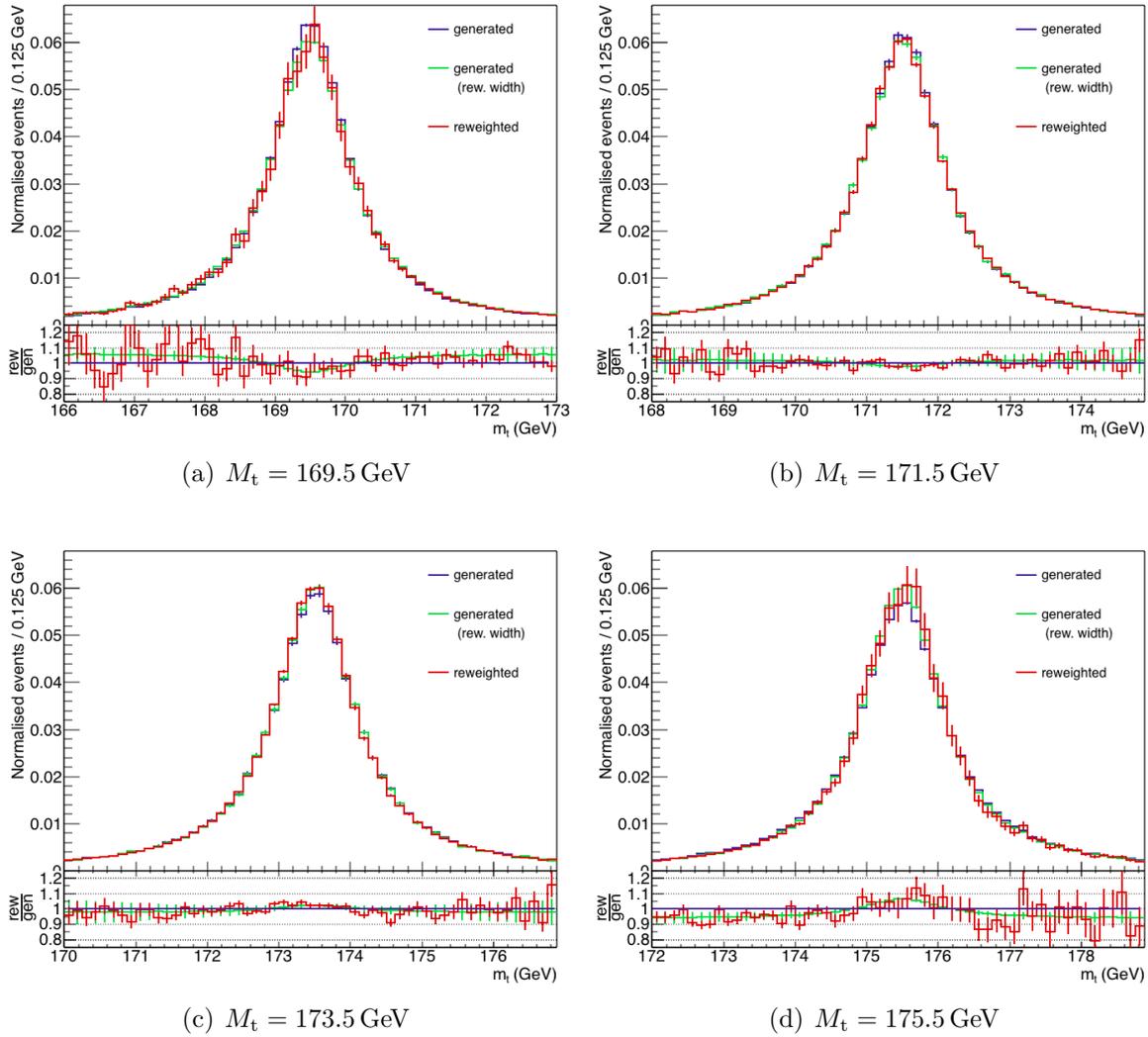


Figure 3.9: Comparison of the normalised distributions of the top quark mass for samples generated with a different top quark mass (blue) and the reweighted samples (red) with a top quark mass of (a)  $M_t = 169.5 \text{ GeV}$ , (b)  $M_t = 171.5 \text{ GeV}$ , (c)  $M_t = 173.5 \text{ GeV}$ , and (d)  $M_t = 175.5 \text{ GeV}$ . As a comparison, the generated samples have also been reweighted such that they all have a top quark decay width of 1.31 GeV (green).



# Reconstruction of Proton Collisions

---

# 4

Each part of the CMS detector is specifically designed to identify and reconstruct certain types of particles. Because of its layered structure and high granularity, the tracker can efficiently record the trajectories of charged particles that cross the detector volume (see Section 4.1). Its design also allows to reconstruct the vertices where the particles originated from. Displaced vertices enable to identify particles with a longer time-of-flight, such as b quarks (see Section 4.3.4) and tau leptons. The electromagnetic calorimeter detects electromagnetic showers initiated by electrons and photons. By clustering close-by cells with an energy deposit above a certain threshold, the direction and energy of these particles can be calculated (see Section 4.2). Likewise, energy deposits of closely interspaced hadrons in the hadronic calorimeter are clustered into jets. As all energy from electrons, photons, and hadrons should be absorbed in the calorimeters, only muons and neutrinos will progress into the muon chambers. Since neutrinos interact only very rarely, any trajectory observed in the outer parts of the CMS detector is attributed to muons.

On itself, this procedure allows for decent particle identification and reconstruction, but a reduction of the misidentification of particles and an improved reconstruction efficiency can be acquired by connecting the observed signals in the different subdetectors. This is achieved by a Particle Flow (PF) algorithm [143]. This algorithm exploits maximally the spatial resolution of the subdetectors to reconstruct individual particles and provide a full event description.

Figure 4.1 presents a visualisation of the interactions that occur when different kinds of particles flow through the CMS detector from the interaction point on the left to the outer layers on the right. Charged particles, indicated by full lines, will leave a track in the silicon tracker. Due to the magnetic field of the solenoid, they will follow a helical trajectory. The momentum of a charged particle is determined by the curvature of the track and its charge by the direction of the bend. Muons will travel almost unimpaired through the calorimeters and the magnet material and will also leave a track in the muon chambers. Note that the direction of the bend reverses outside the solenoid due to the closing of the magnetic field lines. Electrons will lose all their energy in the electromagnetic calorimeter, while charged hadrons will deposit the majority of their energy in the hadronic calorimeter, although small energy losses in the ECAL can be expected. Neutral particles will not interact with the tracker material and will dump their energy in the calorimeters; photons in the electromagnetic calorimeter and neutral

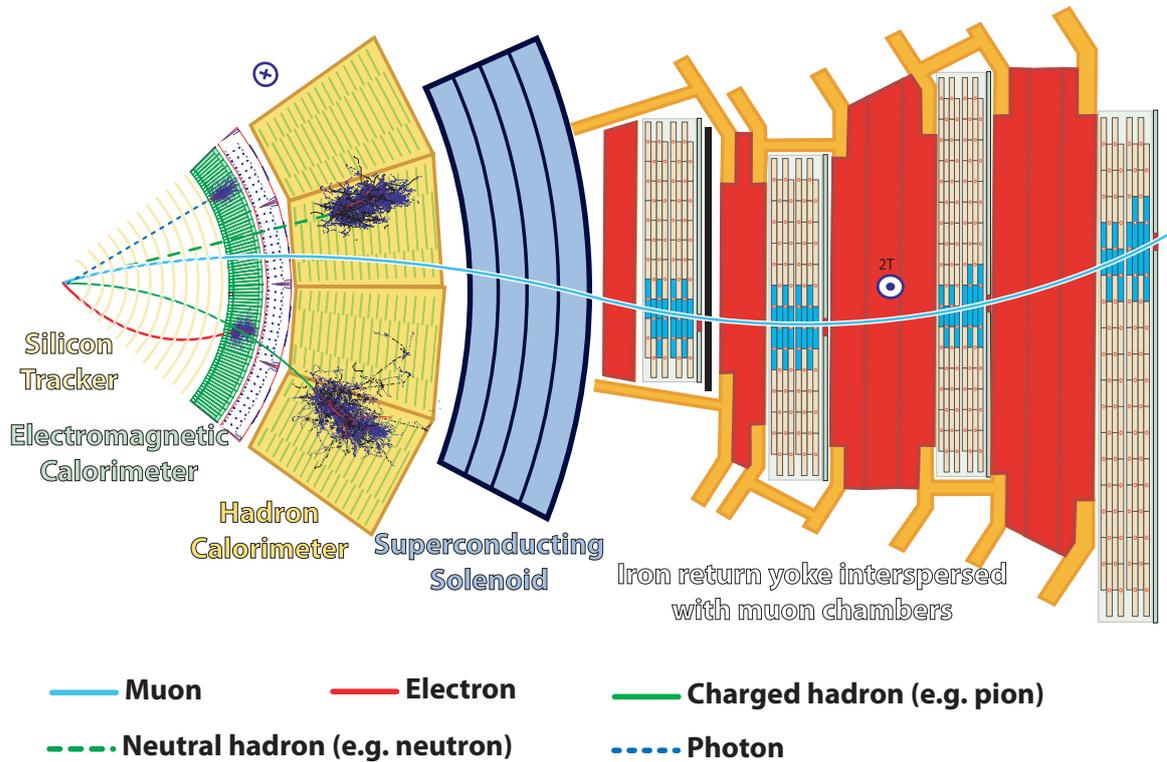


Figure 4.1: A slice of the CMS detector with a visualisation of the interactions of different types of particles when they move through the detector [144].

hadrons mainly in the hadronic calorimeter.

Due to the hermeticity of the CMS detector and the conservation of energy in the transverse plane, the presence of a neutrino can be inferred from missing transverse energy (see Section 4.3.5).

## 4.1 Track Reconstruction

In order to reconstruct a track, the hits in the different layers of the silicon tracker need to be combined. As mentioned in Section 2.2, the tracker consists of three pixel and ten strip layers in the barrel and two pixel and twelve strip layers in the endcap. Due to this complicated geometry, a clever way of predicting in which layers the next hit can be found is needed. Therefore a simplified model is made both of the tracker and of the trajectory of the particles navigating through it [145]. Since the magnet is a large solenoid that extends beyond the span of the tracker, the magnetic field in the tracker can be considered to be uniform [31] and the reconstruction comes down to determining the parameters of a helical trajectory. Because the tracker is very dense and the particles bend a non-negligible amount during their course, multiple scattering, pair-production, and the emission of Bremsstrahlung photons happen frequently. In order to reduce the complexity of the track reconstruction, the material making up the support structure, cooling, and read-out cables is attributed to the sensitive area on the layers. This allows to treat deviations from a simple helical trajectory due to interactions with the tracker

material or photon emissions as uncertainties on the track parameters [146].

The full procedure for the reconstruction of tracks is presented in [147]. A brief summary is given below.

The first step to reconstruct a track is the formation of a seed consisting of a couple of hits. This seed track is then propagated through the tracker using pattern recognition, based on a combinatorial Kalman filter method [147, 148]. Using a first estimate of the track parameters from the seed, the rough position of the hit in the next tracker layer is inferred. When a compatible hit is found, the track parameters are recalculated and the next hit is sought. When several hits are compatible, multiple track candidates are created. Also a track candidate with a missing hit in one of the tracker layers due to measurement inefficiencies is propagated. The candidates are grown in parallel and are propagated to the next layer until the final layer or a stopping condition is reached. At each layer the track parameters of the candidates are known with a better precision and, based on the uncertainty determined by the Kalman filter, the number of track candidates gets truncated to avoid an exponential increase.

Nevertheless, one seed can give rise to several tracks and also one track can be made by several seeds. In order to avoid the double counting of tracks, the number of shared hits between tracks is compared to the total number of hits in those tracks. If more than half of its hits are shared, the track with the least number of hits is discarded, whereas the track with the highest  $\chi^2$  value is discarded when the tracks are equally long. When all ambiguities are resolved, the track is refitted with a Kalman filter and a smoothing function. The first stage, starting from the seed and moving to the outermost hit, is similar to the procedure described above, scaling the covariance matrix with a large factor to avoid any bias. Then, the track is refitted outside-in, using the final results for the track parameters from the previous step as a starting point. At each stage, the updated parameters from the second filter are combined with the parameter values obtained from the first filter. This greatly improves the estimates of the parameters at the surfaces associated with the first and last hit of the trajectory [147].

Considering the multitude of particles that produce hits in the sensitive area of the tracker, there are an enormous amount of possibilities to recombine the recorded hits to form tracks. In order to minimise the number of fake tracks reconstructed from unrelated hits, an iterative tracking procedure is applied.

During each iterative step, the hits that are associated to certain types of reconstructed tracks are masked and new seeds are formed from the remaining unassociated hits in the hit collection. The order of the steps is such that the complexity of the track reconstruction increases with decreasing amount of hits.

During the first couple of iterations, a seed needs to consist of three hits in the pixel detector and the extrapolation towards the centre of the detector needs to be within a certain distance from the interaction point. The efficiency for reconstructing tracks with these kinds of seeds is around 80% and removes about 40% (20%) of hits from the collections recorded by the pixel (strip) detector [143]. Next, seeds with one or two hits in the pixel detector are considered in order to cover for detector inefficiencies and short-lived particles. Then, seeds with no pixel hits are used to reconstruct tracks that

are displaced compared to the interaction point (e.g. originating from the decay of short-lived particles). At this point the reduction of the number of available hits by the previous iterations has become important to reduce the fake rate. Finally, high- $p_T$  tracks and muon tracks, where information from the muon detectors can be used for seeding, are reconstructed.

Primary vertices (PVs) are reconstructed using pixel tracks with a  $p_T > 1$  GeV that have a hit in each layer. As the vertices are expected to lay along the beam line, the transverse distance of the tracks to the  $z$  axis, also called the impact parameter (IP), must be smaller than 1 mm. These tracks are ordered according to their extrapolated position on the  $z$  axis,  $z_{IP}$ , and an average vertex position is calculated for tracks that lay within a certain separating distance  $z_{sep}$  from each other [145, 149]. If there are tracks that are not compatible with this average position, they are re-used to find another primary vertex candidate. This is repeated until it is not possible anymore to make vertices with at least two tracks [149]. The primary vertex that has the largest  $p_T^2$  sum of its associated charged particles and jets is chosen to be the signal primary vertex of the event. The PV reconstruction efficiency is above 95% for most types of events, but deteriorates for events with few charged-particle tracks [145].

## 4.2 Calorimetry

The clustering of calorimeter deposits is performed separately in each ECAL and HCAL subdetector, since these have different compositions and granularities.

First, seeds are identified, i.e. calorimeter cells with an energy above a certain threshold that constitute a local maximum compared to their neighbouring cells. When these neighbours, which have at least one corner in common with the seed, have an energy above a (usually lower) threshold, they are aggregated into a topological cluster together with the seed. The exact values of the thresholds for each subdetector can be found in Table 2 of [143] and are at least twice as high as the electronic noise.

As neutral particles do not generate tracks, their presence can only be inferred from calorimeter clusters. Often there is an overlap with the clusters from charged particles and, because the calorimeter energy of the latter particles can be compared to the energy sum calculated from the corresponding tracks, neutral particle clusters are usually seen as an excess of energy. In order not to overestimate neutral particle deposits, a calibration of the calorimeter response to photons and neutral hadrons is performed [143].

## 4.3 Reconstruction and Identification of Particles and Jets

The PF algorithm combines the observed signals in the subdetectors to reconstruct physical objects, such as muons, electrons, photons, and jets. This involves linking tracks in the inner tracker to calorimeter clusters or tracks in the muon chambers and

calorimeter clusters in one subdetector to those in another subdetector. In order to quantify the quality of the link between two PF elements, a distance between these elements is defined in the  $(\eta, \phi)$  plane. The speed of the algorithm is made independent of the event complexity by only investigating ‘nearest neighbours’ for a possible link [143]. A further optimisation is achieved by first reconstructing objects that are easily identified and masking the corresponding PF elements in the collection. In practice, this means that muons will be reconstructed first, followed by electrons, isolated photons, and, in a later stage, hadrons and non-isolated photons. When all objects have been identified, a post-processing step is carried out to reduce the effect of possibly misreconstructed objects, as explained in Section 4.3.5. Objects reconstructed using the PF algorithm are referred to as ‘PF objects’.

A track is extrapolated from the last hit in the inner tracker to the calorimeter. When the extrapolated track falls within the cluster area, a link between the track and the calorimeter cluster is established. At the intersection points between the track and each tracker layer, a tangent to the track is extrapolated to the ECAL to take account of Bremsstrahlung photons. If clusters can be linked in this way, their energies are added to the energy of the main cluster. The tracks of electron–positron pairs originating from photon conversions are linked with a dedicated conversion finder [150].

Similarly, calorimeter clusters are linked together by extrapolating the cluster position in the more granular calorimeter and investigating if it falls within the envelope of the cluster in the less granular subdetector. If multiple links are possible, the one with the smallest distance is preferred.

How to establish a link between an inner tracker track and a track in the muon chambers is explained below.

### 4.3.1 Muon reconstruction

Muons will in general give rise to a track in the tracker, called a tracker-muon track, and a track in the muon chambers (‘stand-alone muon track’). Depending on the properties of the muon, these can be combined into a ‘global muon track’.

The first step in reconstructing a stand-alone muon track is to combine the hits in the muon chambers into muon segments or stubs. This is done separately for each type of muon subdetector. The segments are then matched to generate seeds for track fitting [143]. The amount of material complicates the track fitting, as multiple scattering happens more often than not. A track in the muon chambers should at least combine two muon segments [143] and is extrapolated to the centre of the CMS detector to check its compatibility with one of the proton collision vertices.

To construct tracker-muon tracks, all tracks with a  $p_T > 0.5$  GeV and a total momentum larger than 2.5 GeV are considered as potential muon candidates. They are extrapolated to the muon system using the magnitude and direction of the magnetic field, while taking the average expected energy loss and possible multiple scattering in the detector material into consideration [143]. A tracker-muon track is accepted as such if there is a match with at least one muon segment.

A combination of a tracker-muon track and a stand-alone muon track is possible by comparing the properties of both tracks when they are propagated onto a common

surface. All hits are then refitted using a Kalman filter and the different tracks are merged into one global muon track.

For muons with a  $p_T > 200$  GeV a global muon track will have an improved resolution compared to the tracker-muon track. In other cases, the tracker-muon track will be more suited to determine the particle properties [143]. While stand-alone muon tracks are excellent to identify muons, their resolution is generally worse than the other cases. In addition, they are more likely to originate from cosmic muons compared to other muon tracks.

An overall muon reconstruction efficiency of more than 95% is attained, while the misidentification rate is lower than 1% [151]. Depending on the need for a high efficiency or a high purity, further requirements can be imposed on an analysis level. When selecting muons with a ‘loose working point’, a PF muon is required to be a global muon or a tracker muon. A PF muon with a ‘tight ID’ always has to be a global muon. On top of that, a tight muon needs to have at least one hit in the pixel detector, contain hits in at least five tracker layers, and have a track  $\chi^2$  smaller than 10 times the number of degrees of freedom. Further, the transverse impact parameter with respect to the primary vertex needs to be smaller than 2 mm and the longitudinal one smaller than 5 mm. The combination of these requirements gives a good discrimination against hadronic punch-through, cosmic muons and muons created in-flight, while ensuring a good  $p_T$  measurement [152]. The efficiency to identify muons with loose and tight IDs is visualised in Figure 4.2 [152]. The drops in the distributions correspond to partially uninstrumented regions in the muon chambers. As expected, a higher efficiency is attained for muons with a loose ID. Efficiency differences between data and simulation are corrected by applying scale factors to the simulation. This is explained in more

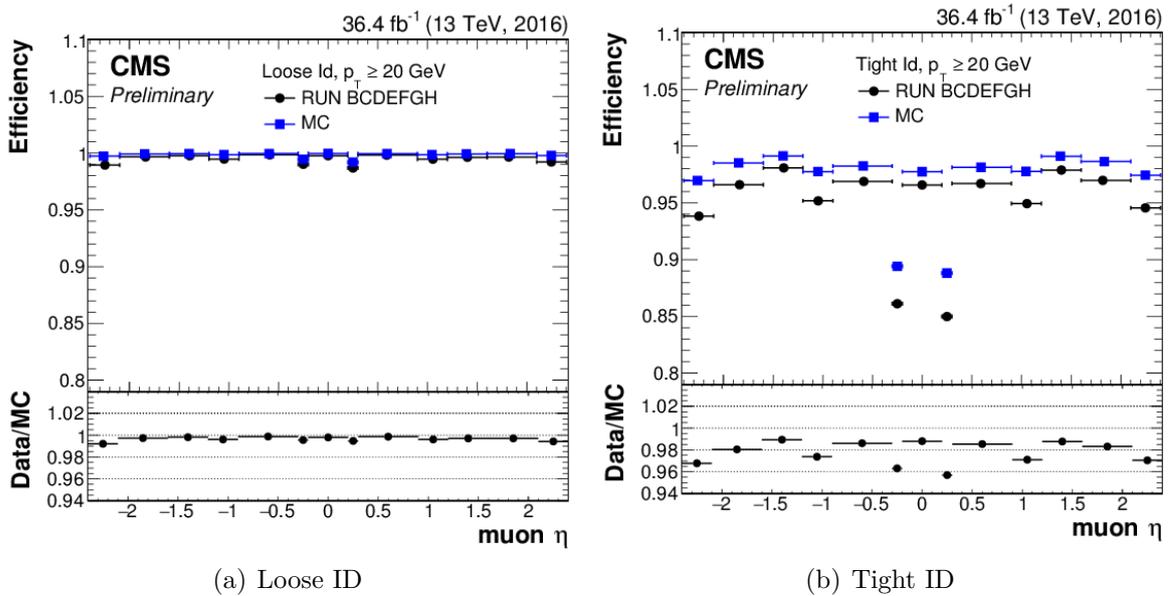


Figure 4.2: Efficiency to identify a muon with a (a) loose and (b) tight ID in data (black) and simulation (blue) as a function of the muon  $\eta$  [152].

detail in Chapter 5.

A relative isolation variable is introduced to differentiate muons produced at the primary vertex from muons appearing in jets. The latter are surrounded by a multitude of particles, so the distribution of the transverse momentum in the direct environment of the muon will be different. Therefore, a cone with an opening angle of 0.4 is defined around the muon and the momentum contributions of its surrounding particles are added per particle type, as in Equation (4.1). The momentum inside a small veto cone, as illustrated in Figure 4.3 [153], is attributed to the muon itself. Whereas charged particles originating from pileup (PU) interactions can be traced back to a different primary vertex than the muon, contributions from neutral particles are more difficult to discern. In the vicinity of a muon, it is expected that the contribution from charged PU particles is twice as large as that of neutral particles originating from PU interactions. Therefore, a correction is applied to the momentum contribution of the neutral particles,

$$\text{relIso} = \frac{\sum_{\text{charged}} p_T + \max\left(0, \sum_{\text{neutral}} p_T + \sum_{\gamma} p_T - \Delta\beta \sum_{\text{charged PU}} p_T\right)}{p_{T,\mu}}, \quad (4.1)$$

where  $\Delta\beta = 0.5$ . When the PU correction is larger than the sum of the neutral particle momenta, it is assumed that there are no neutral particles present in the isolation cone.

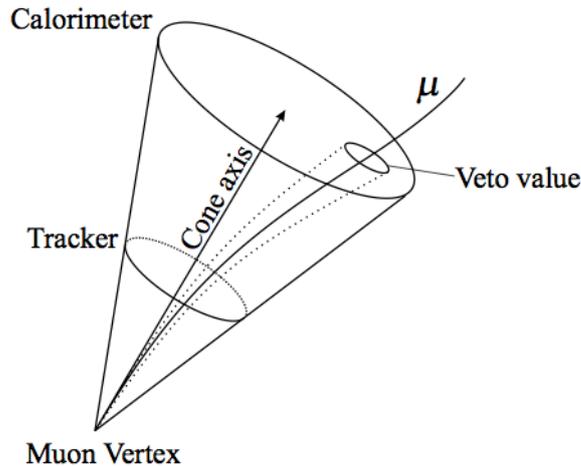


Figure 4.3: Illustration of the muon isolation cone [153]

For a loose muon ID, the relative isolation is required to be smaller than 0.25, while it has to be smaller than 0.15 for tight muons. The efficiency for muons with a tight ID and a tight relative isolation is presented in Figure 4.4 as a function of  $p_T$  and  $\eta$ . Whereas the  $\eta$  distribution is relatively flat, a drop-off is observed for muons with a low  $p_T$ . Once again, scale factors are applied to the simulation to account for efficiency differences between data and simulation.

The distribution of the relative isolation values themselves is visualised in Figure 4.5 for muons with a tight ID. Only events containing exactly one muon and four jets are considered (see Chapter 5). Since the muon is also required to have a tight relative isolation, the distribution stops abruptly at 0.15. In addition, the distribution of one of the muon ID variables, i.e. the number of hits in the pixel detector, is visualised in the

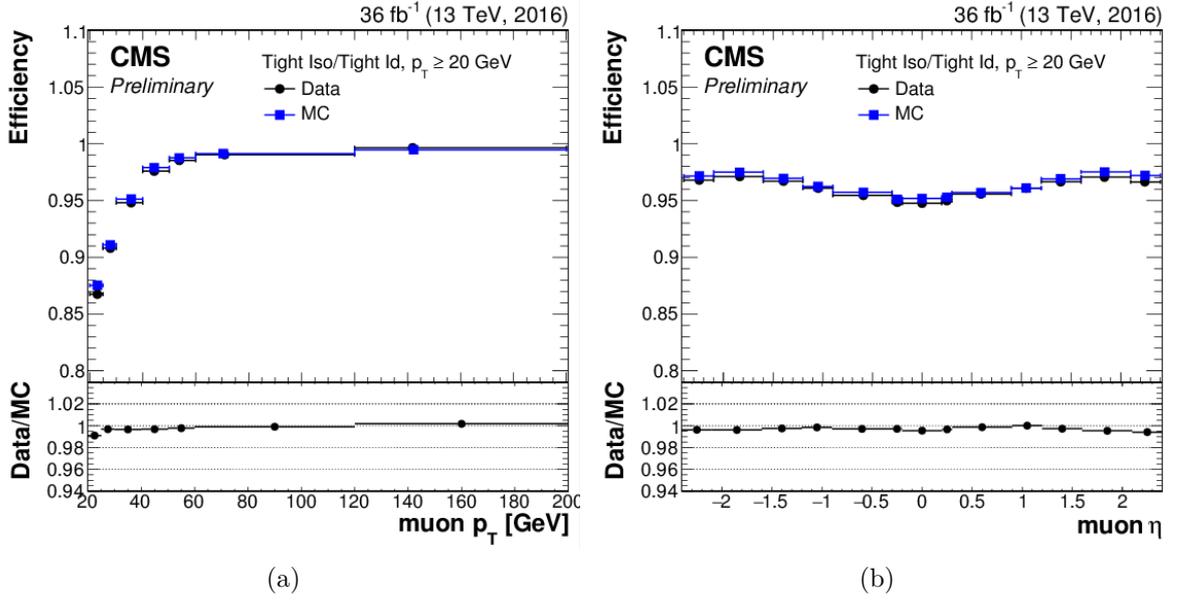


Figure 4.4: Efficiency to identify a muon with a tight ID and a tight relative isolation in data (black) and simulation (blue) as a function of the muon (a)  $p_T$  and (b)  $\eta$  [152].

same figure. As the pixel detector has three layers in the barrel and two endcap discs, the peak around three valid hits is to be expected.

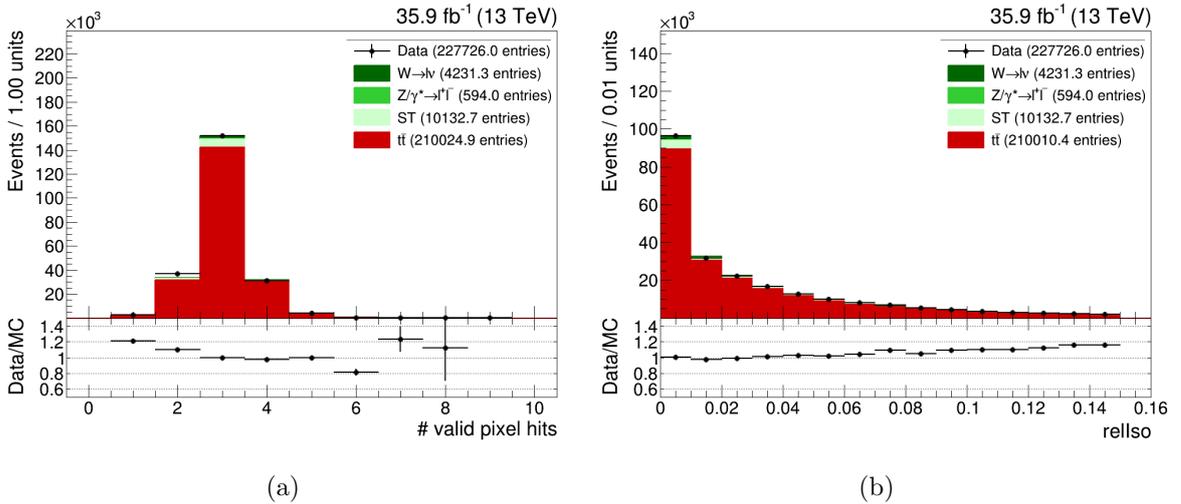


Figure 4.5: Distribution of the (a) number of valid hits in the pixel detector and (b) relative isolation for isolated muons with a tight ID and tight relative isolation, considering only events with one isolated muon and four jets (see Chapter 5).

### 4.3.2 Electron reconstruction

There are different ways to link an electron track to its calorimeter cluster. For isolated electrons with high  $p_T$  an ECAL-driven approach is used. Starting from clusters with a transverse energy of at least 4 GeV, the position of the pixel seed hits is inferred from the cluster barycentre and energy, which gives an indication of the bend of the electron trajectory. A so-called supercluster is formed when the energy radiated in the tracker material is added to the energy of the cluster.

When the electron is not isolated or has low momentum, a tracker-driven approach can be more appropriate [143]. Tracks with a  $p_T > 2$  GeV are used as seeds and propagated to the calorimeter surface. Depending on the angular position, this can be the surface of the ECAL, preshower, or HCAL. It is then investigated if the tracks can be connected to a calorimeter cluster. Usually, the amount of energy radiated in the tracker is non-negligible, so a one-to-one comparison of the energy of the cluster and the track is not possible. In order to take account of the energy loss along the trajectory, the electron track is re-fitted with a Gaussian-sum filter (GSF) [154], from which the track parameters are estimated. The GSF track is associated to the supercluster using a multivariate analysis (MVA) discriminator that combines a number of track and cluster variables, such as the number of hits in the track, the ratio  $\chi_{GSF}^2/\chi_{Kalman}^2$  and the difference between the position of the cluster and the extrapolated track [155]. Ambiguities arising from several GSF tracks being compatible with the same supercluster are resolved by choosing the track with the least amount of missing hits or, if the number of missing hits is the same, the track for which the cluster-to-track momentum ratio is closest to one.

Finally, the electron candidates reconstructed with both strategies are merged. In case of duplicates, a preference is given to ECAL-seeded electrons, because tracker-seeded candidates arise more often from track segments corresponding to photon conversions. An overall reconstruction efficiency of 93% is achieved for electrons from Z boson decay that have a  $p_T > 15$  GeV [155].

As for muons, further identification criteria are applied on an analysis level. These are based on the value of various quantities determined during the reconstruction process, such as the number of missing track hits or the difference between the cluster energy and the track momentum. Depending on the required purity, cuts on these quantities can be chosen to be stricter or looser. Due to the differences in granularity and energy thresholds, the recommended cuts are different in the barrel and the endcap region of the ECAL. A complete list of the identification variables and their respective cuts can be found in [156]. The tight working point has an average efficiency of 70%, while the efficiency is around 90% for a loose working point and 95% for an electron veto working point. The relative isolation of an electron is determined with Equation (4.1), using an effective area correction instead of a  $\Delta\beta$  correction and an opening angle of 0.3 for the cone around the electron direction. These changes provide a better pileup treatment. A smaller cone reduces the probability of other particles overlapping with the electron, while the effective area approach ensures that tracks of the electron emissions are not counted as part of the pileup contribution. The areas are determined from data as a function of  $|\eta|$  and can be found in [157].

### 4.3.3 Jet reconstruction

Hadrons are reconstructed by linking HCAL clusters to inner tracker tracks. If such a link is made, a charged hadron is found. The energy can be calculated in two different ways, namely the sum of the momenta of the tracks associated to the cluster or the sum of the energies of the associated ECAL and HCAL clusters. The maximum of these will be chosen to be the ‘true’ energy of the charged hadron [143]. However, if the sum of the track momenta is significantly larger than the calorimetric energy, a search for fake tracks and muons abiding looser criteria than the ones above is initiated. Global muons that have a momentum estimate with a relative precision of 25% or better are identified as muons [143] and their track is removed from the collection. In this way, the particle flow algorithm allows to recover more muons than a standard calorimeter-driven approach. The remaining tracks are considered to be charged hadrons. When the calorimetric energy and the momentum inferred from the tracks correspond, the measurements in the calorimeters and the tracker are refitted together in order to obtain a superior momentum resolution, even at high  $p_T$ , where the resolution of the track parameters is usually worse. When the calorimetric energy is larger than the combined track momentum, the difference is attributed to neutral particles, as well as clusters that cannot be linked to tracks. Within the tracker acceptance, clusters in the ECAL are reconstructed as photons and clusters in the HCAL as neutral hadrons. Beyond the tracker acceptance the difference between charged and neutral hadrons is not so clear. Whenever a link can be established between an ECAL and an HCAL cluster, the energy deposits will be reconstructed as a hadron. When the ECAL cluster cannot be linked, it is reconstructed as a photon.

#### Jet clustering

Particles are clustered into jets using the anti-kT algorithm [158] with a cone size of 0.4. A minimal-distance principle is used to decide whether particles are grouped into so-called protojets or if the protojet is considered to be a ‘completed’ jet. Two distances are defined as follows,

$$d_{ij} = \min(p_{T,i}^{-2}, p_{T,j}^{-2}) \frac{\Delta_{ij}^2}{R^2}, \quad (4.2)$$

$$d_{iB} = p_{T,i}^{-2}, \quad (4.3)$$

where  $R$  is a cone radius parameter and  $\Delta_{ij}$  an angular distance defined as

$$\Delta_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2},$$

with  $y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}$  the rapidity.

The distance  $d_{ij}$  can be seen as the distance between two entities, i.e. particles or protojets, whereas  $d_{iB}$  is the distance between an entity and the beam. When  $d_{ij}$  is smaller than  $d_{iB}$ , the entities will be clustered together to form a (larger) protojet. When the distance of this protojet to the beam is smaller than its distance to any other entity, the protojet is declared a jet and is masked in the set of entities. An

advantage of this approach is that the jet reconstruction is not influenced by soft radiation, as would be the case when the clustering is started from the highest  $p_T$  entities [159]. In practice, soft particles cluster around hard particles and if the jets are separated by more than  $2R$  their shape is perfectly conical [158]. If two jets are within a distance  $R < \Delta_{ij} < 2R$  their shape depends on the ratio of their momenta. When one jet has a much larger  $p_T$  than the other, the overlap area between the two jets will be attributed to that jet. If their momenta are comparable, the overlap area will be shared evenly. Jets within  $R$  of each other will be reconstructed as a single jet.

Considering only events with one isolated muon and four jets (see Chapter 5), the distribution of the minimum and the maximum angular distance  $\Delta R$  between two jets is plotted in Figure 4.6. Since the cone radius parameter in the anti-kT algorithm is equal to 0.4, the  $\Delta R$  distributions have a lower limit at this value. Whereas the distribution of the minimum  $\Delta R$  peaks around 1, the maximum  $\Delta R$  is mostly larger than or equal to  $\pi$ , which implies that the jets are positioned almost back-to-back.

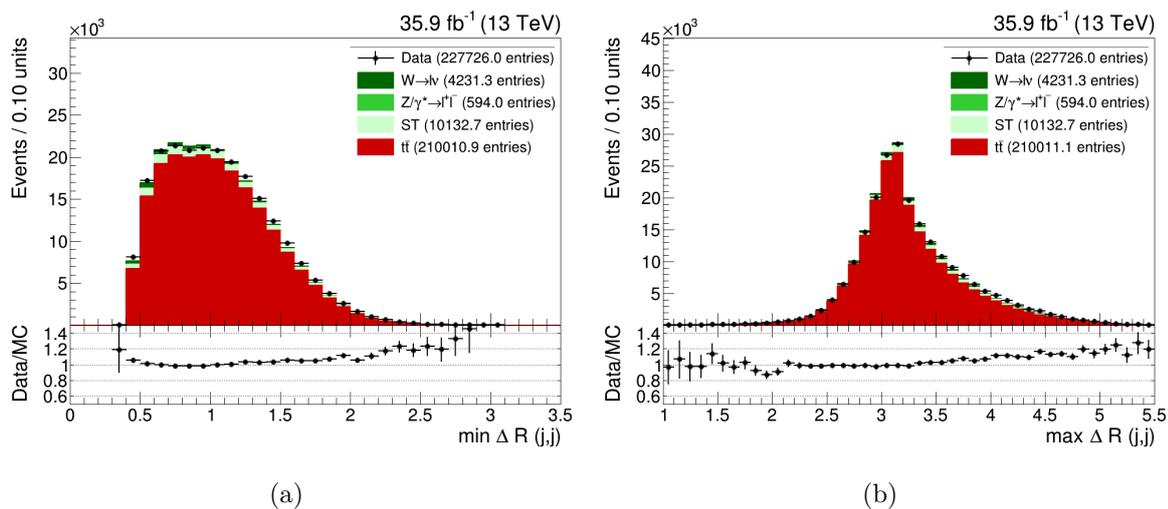


Figure 4.6: Distribution of the (a) minimum and (b) maximum  $\Delta R$  between two jets for events with one isolated muon and four jets (see Chapter 5).

### Jet energy corrections

Due to the non-linear response of the calorimeters, it necessary to calibrate the deposited energy in order to get a consistent energy measurement that is independent of the jet energy scale (JES) and its location in the CMS detector. In addition, corrections need to be applied to take account of pileup and the UE activity. This is implemented using a factorised approach where the output of one step serves as the input for the next [160, 161]. During each step the corrections are implemented by scaling the jet four-momentum. A flowchart of the jet energy corrections that are applied to data and simulation is presented in Figure 4.7 [160]. First, an energy offset correction is determined for data and simulation separately. Then, simulated events are used to estimate the detector response as a function of the jet  $p_T$  and  $\eta$  and additional corrections

are applied to data to take care of residual differences. These are explained in more detail below. Besides, optional corrections are available, such as corrections taking account of the energy scale differences between jet flavours, but these are not applied in this analysis.

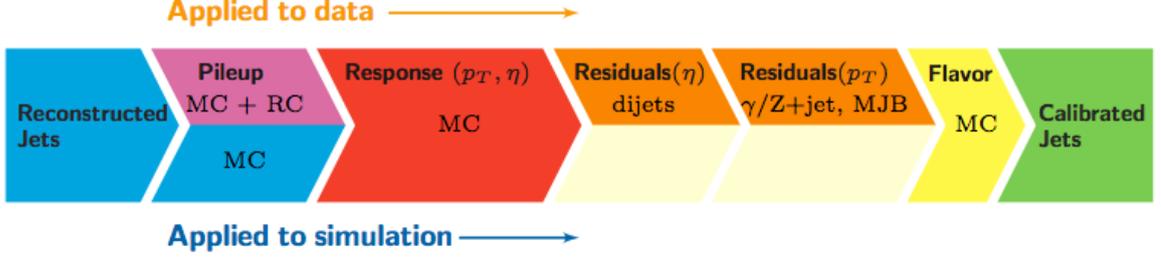


Figure 4.7: Overview of the different stages for jet energy scale corrections [160]. The flavour-dependent corrections are optional.

### Offset corrections (L1)

The activity in the underlying event and energy contributions from pileup (PU) form a pedestal whereupon the jets can be found. This results in an overestimation of the jet energy, so a correction factor as a function of the jet pseudorapidity  $\eta$  is calculated. First, the PU contributions are minimised by rejecting charged hadron tracks that can be unequivocally attributed to a PV originating from PU. This is done before the jets are clustered and is called charged hadron subtraction (CHS) [160]. Then, the hybrid jet area method is used to make an estimate of the average energy density of the event, which is then multiplied by the jet area  $A_j$ . Considering the parameters  $\beta_0$ ,  $\beta_1$  and  $\beta_2$ , each having a dependency on  $\eta$ , the corrected jet  $p_T$  can be calculated as

$$p_{T,L1} = p_{T,uncorr} - A_j [\beta_0(\eta) + \rho \beta_1(\eta) (1 + \beta_2(\eta) \log p_{T,uncorr})] , \quad (4.4)$$

where  $\rho$  is the per-event offset energy density [160]. The  $\beta_i$  parameters are determined using simulation. Events are simulated with and without PU contributions and the corresponding reconstructed jets are matched. The observed jet  $p_T$  difference gives an estimate of the offset density. Corrections are determined as a function of the jet  $p_T$  and  $\eta$ , and the amount of PU (number of PVs).

Small differences in the offset energy density are observed for data compared to simulation, as can be seen in Figure 4.8 for reconstructed events with one muon and four jets (see Chapter 5). Corrections to account for the observed differences are determined by applying the random cone (RC) method to data events that do not contain a hard-scattering process [160]. If electronic noise is considered to be negligible, the energy density in the events can be solely attributed to PU. The RC method clusters particles into jets using cones that have random positions in the  $(\eta, \phi)$  plane. This way the entire plane is filled with jets, whereof the average  $p_T$  indicates the average offset due to PU.

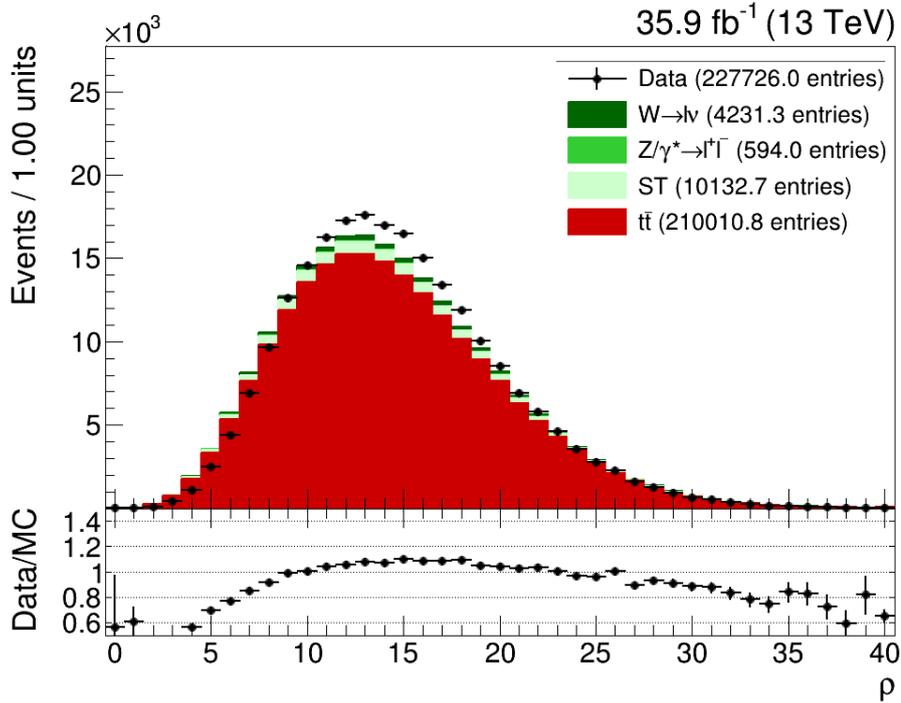


Figure 4.8: Distribution of the offset energy density for events with one muon and four jets (see Chapter 5).

### Detector response corrections (L2L3)

Jets that are observed in different locations in the CMS detector will experience a non-identical detector response. Depending on the jet  $\eta$ , the composition of the subdetectors and the amount of traversed detector material changes. In addition, a non-negligible amount of hadrons in low- $p_T$  jets are not observed because they have an energy below the calorimeter thresholds. Using simulated events, the jet  $p_T$  before and after the detector convolution step is compared, thus creating detector response scale factors as a function of the jet  $p_T$  and  $\eta$ . The events are simulated such that each  $(p_T, \eta)$  region contains a statistically significant amount of events.

### Residual corrections for data (L2L3Res)

In general, the energy response of the CMS detector is lower for data compared to simulation. In addition to this absolute energy scale offset, residual response dependencies on  $\eta$  and  $p_T$  remain after applying the corrections described above. These are corrected by investigating the transverse momentum imbalance of the events. The detector response is estimated by comparing the  $p_T$  contained in the event to the  $p_T$  of a reference object. In order to correct for the differences in  $\eta$ , dijet events are used, where at least one jet has  $|\eta| < 1.3$ . The response is thus calibrated with respect to the barrel region of the detector. The remaining energy scale corrections are calculated solely in this region, as they can easily be extended to the rest of detector.

The absolute energy scale is calibrated using  $Z + \text{jets}$  events, where the  $Z$  boson decays into a muon pair. For the  $p_T$  dependency of the response a global fit is performed using  $DY + \text{jets}$  and multijet events.

In addition to JES corrections, also variations in the jet energy resolution (JER) have to be resolved. As the JER is worse in data than in simulation, the resolution of simulated jets is smeared such that it resembles the one observed in data. Scale factors are derived as a function of  $p_T$ ,  $\eta$ , and the number of PVs using a similar technique as is applied for the JES. The major difference is that the width of the response distribution is considered instead of the mean.

### Jet identification

A typical jet in a multijet event consists mainly of charged hadrons, supplemented by neutral hadrons and photons [143]. Within the tracker acceptance region, the jet identification criteria reflect this composition by requiring at least two jet constituents, whereof one is charged, and a charged hadron fraction larger than zero. Since hadrons can also deposit small amounts of energy in the electromagnetic calorimeter, the charged electromagnetic fraction is allowed to be quite large, but it has to be smaller than 99%. The requirements on the neutral energy fractions depend on the working point of the identification. The conditions for the loose and tight working point are summarised in Table 4.1 [162].

Table 4.1: Identification criteria for jets within  $|\eta| < 2.4$  for the loose and tight working point [162].

	loose	tight
Charged hadron fraction	$> 0$	$> 0$
Charged EM fraction	$< 99\%$	$< 99\%$
Neutral hadron fraction	$< 99\%$	$< 90\%$
Neutral EM fraction	$< 99\%$	$< 90\%$
Number of constituents	$> 1$	$> 1$
Number of charged constituents	$> 0$	$> 0$

### 4.3.4 Identification of jets originating from $b$ quarks

Due to the hadronisation process, jets originating from  $b$  quarks will contain so-called  $b$  hadrons. Because of their high mass, these will decay in-flight after traversing a measurable distance. Jets originating from  $b$  quarks can therefore be recognised by the presence of a secondary vertex (SV) and the larger impact parameter of their decay products with respect to the PV. There are several algorithms that can discriminate or ‘tag’ jets originating from  $b$  quarks, based on the properties of charged particles in

jets [163]. The most simple algorithms use just one observable, but also a combination of variables can be made to increase the discriminating power. The algorithm output is a single value that reflects how likely the jet is to originate from a b quark. As always, a choice has to be made between a high efficiency and a high purity. Several working points corresponding to a certain misidentification rate are defined. These are implemented by requiring the b-tag discriminator value to be above a certain threshold. In this analysis the Combined Secondary Vertex v2 (CSVv2) algorithm will be used. Its working points can be found in Table 4.2 [164] and are usually indicated by appending a letter to the algorithm name, e.g. CSVv2M for the medium working point (WP).

Table 4.2: Overview of the discriminator values associated to the working points of the Combined Secondary Vertex v2 (CSVv2) algorithm, as well as the efficiencies to tag a jet originating from a b quark and the misidentification rate, i.e. the relative amount of tagged jets that do not originate from b quarks. The misidentification rate is estimated for jets with a  $p_T$  of around 80 GeV [164].

	WP	misID rate	$\epsilon_b$
Loose (L)	0.5426	10%	81%
Medium (M)	0.8484	1%	63%
Tight (T)	0.9535	0.1%	41%

The CSVv2 algorithm combines information from the impact parameter of tracks with kinematic variables associated with a secondary vertex [163]. Jets are divided into three categories, based on whether a secondary vertex is found or not. Secondary vertex candidates are formed when tracks that have a  $p_T > 0.8$  GeV and a longitudinal impact parameter  $z_{IP} < 0.3$  cm are clustered around seeds with an IP  $> 50 \mu\text{m}$  and an IP significance (the IP value divided by its uncertainty) larger than 1.2. They must have fewer than 79% of their associated tracks in common with the primary vertex and the flight direction of the SV should be within  $\Delta R < 0.3$  of the axis of the jet it is associated with [164]. When no secondary vertex is reconstructed, the jet may enter the pseudo vertex category when at least two tracks with an IP significance larger than 2 are found. For jets in this category some SV-based quantities are computed without performing an actual vertex fit [163]. When no SV or pseudo vertex can be identified, only information related to displaced tracks is used.

The distribution of the CSVv2 discriminator values is presented in Figure 4.9 for  $t\bar{t}$  events with one isolated muon and at least four jets. A distinction is made between jets originating from b, c and udsg partons. The CSVv2 algorithm discriminates very well between jets originating from b and udsg partons. The differentiation between jets originating from b and c quarks is less clear-cut due to the presence of c hadrons, which can also produce SVs. Though in general, these can be found closer to the PV than those produced by b hadrons.

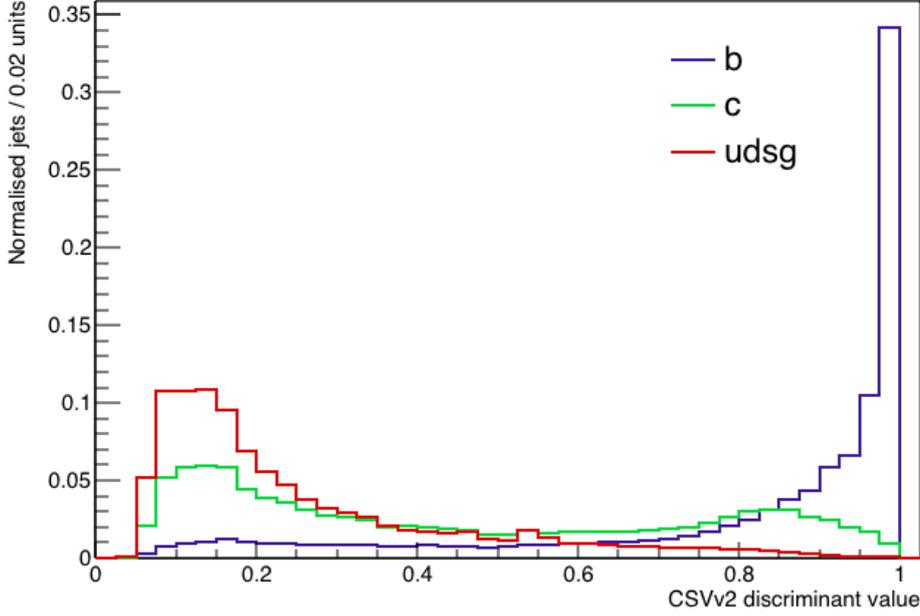


Figure 4.9: Normalised distributions of the CSVv2 discriminator values for jets originating from b, c and light-flavour partons. Only jets with a  $p_T > 30$  GeV and  $|\eta| < 2.4$  that derive from  $t\bar{t}$  events with one isolated muon and at least four jets are considered (see Chapter 5).

### 4.3.5 Missing transverse energy reconstruction

Conservation of energy dictates that the total energy and momentum of the event in the transverse plane should be equal to zero. This is easily checked by vectorially adding the transverse momenta of all reconstructed particles and jets. The missing transverse momentum is defined as the opposite of the resulting vector, if it is non-zero. The missing transverse energy  $E_T^{\text{miss}}$  is the modulus of this vector.

If energy corrections are applied to the jets, this also influences the missing transverse energy. Therefore, so-called type-I corrections are applied [160]. For each jet  $i$  that has a  $p_T > 10$  GeV, the missing transverse momentum is corrected as

$$\vec{p}_{T,\text{typeI}}^{\text{miss}} = \vec{p}_{T,\text{uncorr}}^{\text{miss}} + \sum_i \vec{p}_{T,\text{uncorr}}^i - \sum_i \vec{p}_{T,\text{corr}}^i - \sum_i \vec{\Theta}^i, \quad (4.5)$$

where  $\vec{p}_{T,\text{corr}}^i$  is the fully corrected jet  $p_T$  and  $\vec{\Theta}^i$  is the average offset due to PU.

Since most events have a negligible amount of missing transverse energy, the energy balance of the event is re-evaluated in a post-processing step when the  $E_T^{\text{miss}}$  is very large. Especially the misreconstruction of muons can have a large effect on the missing transverse energy [143]. A hadron can, for example, punch through to the first muon station, such that it is reconstructed as a muon. The calorimetric cluster of the hadron will then give rise to an additional neutral hadron, which induces a missing transverse energy component in the direction opposite to the reconstructed muon. This scenario is

---

investigated in the post-processing step and if it reduces the missing transverse energy, the reconstructed muon is changed into a reconstructed hadron and the extra neutral hadron is removed from the collection of reconstructed objects. Similarly, a muon can be reconstructed as a charged hadron when its angular position overlaps with the calorimetric cluster of an energetic neutral hadron, giving a missing transverse energy component in the direction of the muon. Also cosmic muons that cross the detector at the same time as a bunch crossing contribute to the missing transverse energy. The post-processing step will investigate a number of such scenarios to redress the energy balance of the event when appropriate.



# 5

## Event Selection & Reconstruction

---

In order to analyse the process of interest, signal events need to be discriminated from background processes. A dedicated event selection is employed, where, in addition to selecting the correct event topology, the signal-to-background ratio is enhanced by rejecting as many background events as possible. This is done by exploiting the kinematic properties of the events, which are different for signal and background. The selected events are then reconstructed and, if appropriate, additional requirements to improve the background reduction can be imposed.

This analysis will focus on pair-produced top quarks that have a semileptonic decay. More specifically, the lepton is required to be a muon. The event selection and reconstruction strategies are outlined below.

### 5.1 Trigger and Basic Event Requirements

Most proton collisions produce soft interactions. As described in Section 2.2.5, a first event selection focusing on hard scattering processes is made using the CMS trigger system. So-called trigger paths or triggers contain sets of requirements that are used to investigate if events are interesting to store for further analysis. An event is called triggered when it is accepted by one of the trigger paths. In that case, the decision for each trigger path is saved as an event variable. This enables to make a pre-selection based on the objects that are observed in the event. For this analysis, the trigger paths `HLT_IsoMu24` and `HLT_IsoTkMu24` were used in a logical ‘OR’ structure. This maximises the amount of events since not all trigger paths were tested during the entire data-taking period. These trigger paths require that events have an isolated muon, reconstructed at HLT level, that has a  $p_T > 24$  GeV and, when applied to the data set recorded and validated by CMS, a number of events corresponding to an integrated luminosity of  $35.8 \text{ fb}^{-1}$  is selected. Also simulated events are subjected to the triggers in order to get a consistent pre-selection. Scale factors are calculated to correct for the different selection efficiency of the triggers applied to data compared to simulation. These are produced centrally by the CMS collaboration as a function of  $p_T$  and  $|\eta|$  [165] and are applied to simulated events only. The scale factor values for the trigger paths above range between 0.9 and 1.0, so in general the triggers have a higher efficiency to recognise muons in simulated events than in data.

Further, event filters are applied so as to reject events that are affected by detector noise. To ensure that the event was created by a proton collision, the PV is required to be positioned within 24 cm of the nominal interaction point along the  $z$  axis and at most 2 cm away from the  $z$  axis in the transverse direction.

## 5.2 Event Selection

The topology of semileptonically decaying  $t\bar{t}$  events is visualised in Figure 5.1. Selected

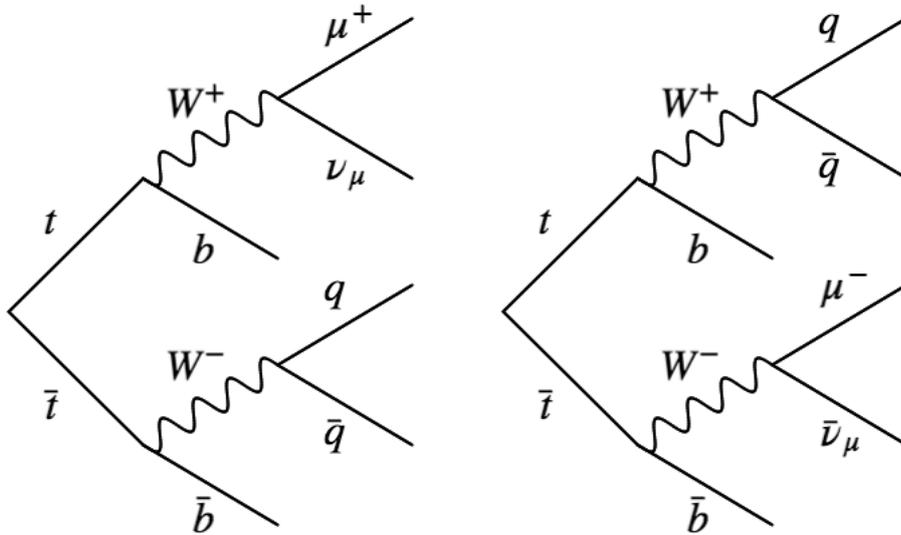


Figure 5.1: Schematic drawing of the semileptonic decay of pair-produced top quarks, where the lepton is a muon. Both the cases where the lepton originates from the decay of a top (left) or antitop quark (right) are visualised.

events need to have exactly one muon and exactly four jets, all satisfying the requirements of the tight ID. In addition, the selected muon needs to have a transverse momentum  $p_T > 26$  GeV and a pseudorapidity  $|\eta| < 2.4$ . To ensure that the events are semileptonic, they are rejected if any additional muons (electrons) with a loose (veto) ID and a  $p_T > 10$  GeV (15 GeV) are found within  $|\eta| < 2.5$ . After applying jet energy corrections, jets are required to have a  $p_T$  between 30 GeV and 250 GeV and  $|\eta| < 2.4$ . They are expected to be separated by  $\Delta R > 0.4$  (0.3) from any muon (electron) with a loose (veto) ID. At least two jets should be b tagged according to the CSVv2 algorithm at medium working point. These conditions are summarised in Table 5.1.

As can be seen in Figure 5.2, the requirement on the number of b tagged jets is excellent in reducing the amount of selected background events that do not involve top quarks, such as Drell–Yan and W boson production with additional jets. The data surplus in the first couple of bins is due to multijet events. These processes are not included in the simulated samples, since the amount of events that remains after the full event selection is negligible. The effect of each selection requirement on the

Table 5.1: Summary of the requirements for leptons and jets in order for an event to be selected.

	1 muon	veto muons	veto electrons	4 jets
$p_T >$	26	10	15	30
$p_T <$				250
$ \eta  <$	2.4	2.5	2.5	2.4
WP	Tight	Loose	Veto	Tight
b tags				$\geq 2$ CSV <sub>v2M</sub>

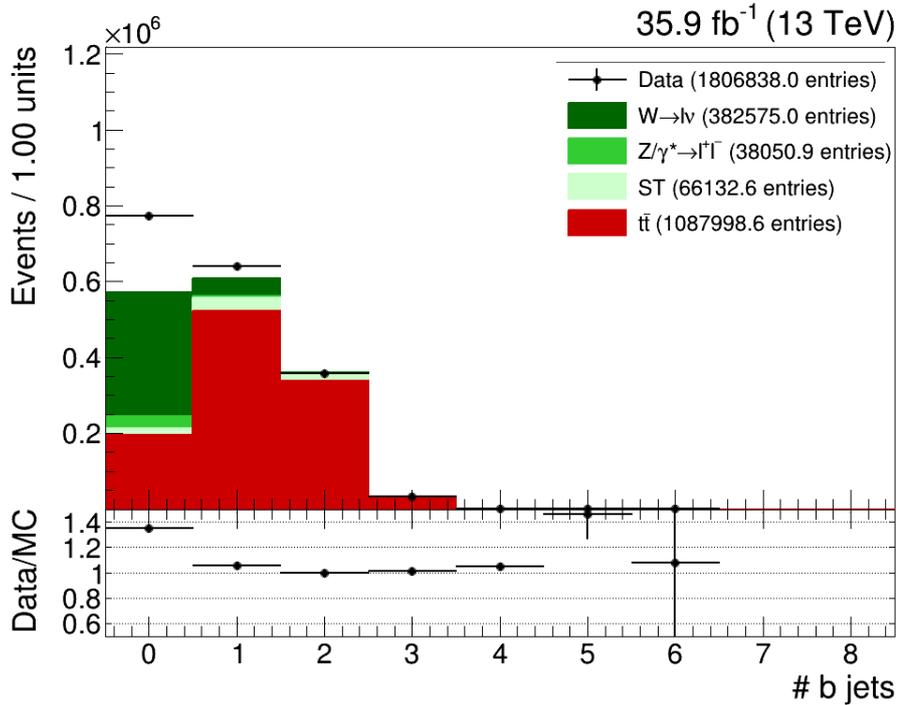


Figure 5.2: Distribution of the number of b tagged jets for events with exactly one muon and at least four jets. The overshoot of the data in the first two bins can be attributed to multijet events.

Table 5.2: Effect of each selection requirement on the number of selected events. ( $35.9 \text{ fb}^{-1}$  of int. lumi.)

	$t\bar{t}$	single top	DY + jets	W + jets	Total exp.	Observed
Preselected	29795042	10340645	53560577	52532495	146228759	785921686
Triggered	$4024706 \pm 2006$	$886677 \pm 942$	$12963479 \pm 3600$	$8637614 \pm 2939$	$26512476 \pm 5149$	469696507
PV & filters	$3978086 \pm 1995$	$877269 \pm 937$	$12835272 \pm 3583$	$8557896 \pm 2925$	$26248523 \pm 5123$	467076864
1 muon	$2847559 \pm 1687$	$701236 \pm 837$	$4132677 \pm 2033$	$7341341 \pm 2709$	$15022813 \pm 3876$	299539969
$\geq 4$ jets	$1141011 \pm 1068$	$71856 \pm 268$	$41349 \pm 203$	$423988 \pm 651$	$1678204 \pm 1295$	1898415
= 4 jets	$675725 \pm 822$	$47909 \pm 219$	$30618 \pm 175$	$321685 \pm 567$	$1075936 \pm 1037$	1273006
$\geq 1$ b jet	538280 $\pm$ 734	34984 $\pm$ 187	4580 $\pm$ 68	42620 $\pm$ 206	620463 $\pm$ 788	658074
$\geq 2$ b jets	212584 $\pm$ 461	10602 $\pm$ 103	604 $\pm$ 25	4621 $\pm$ 68	228410 $\pm$ 478	231275
Jet $p_T < 250$ GeV	210011 $\pm$ 458	10133 $\pm$ 101	594 $\pm$ 24	4231 $\pm$ 65	224969 $\pm$ 474	227726
Other	54059 $\pm$ 233	1311 $\pm$ 36	52 $\pm$ 7	240 $\pm$ 16	55662 $\pm$ 236	535553

number of selected events is presented in Table 5.2 for each data set separately. The requirements contained in the final step will be explained in more detail in Section 5.3. The amount of simulated events is corrected for differences in the selection efficiency between simulated events and data. To this purpose event scale factors are used.

### Lepton identification and isolation

Similar to the muon trigger scale factors mentioned above, also scale factors for muon identification and isolation as a function of  $p_T$  and  $\eta$  are provided by the CMS collaboration [165]. As the observed efficiency differences between data and simulation are very small, the scale factors have a value very close to one in all regions.

### b quark identification

The probability for an event to contain  $i$  b tagged jets and  $j$  jets that are not b tagged, also called ‘light’ or ‘light-flavour’ jets, is

$$P(\text{SIM}) = \prod_{i=\text{tagged}} \varepsilon_i \prod_{j=\text{not tagged}} (1 - \varepsilon_j), \quad (5.1)$$

for simulated events and

$$P(\text{DATA}) = \prod_{i=\text{tagged}} \text{SF}_i \varepsilon_i \prod_{j=\text{not tagged}} (1 - \text{SF}_j \varepsilon_j), \quad (5.2)$$

for data events, where  $\varepsilon_i$  and  $\varepsilon_j$  are the efficiencies for jets to be b tagged in the simulation and the scale factors  $\text{SF}_i$  and  $\text{SF}_j$  take account of efficiency differences between data and simulation. The latter are determined centrally by the CMS collaboration [164]. The efficiencies are determined for simulated  $t\bar{t}$  events containing exactly one isolated muon and at least four jets that comply with the requirements defined above. The b tag efficiency, or misidentification probability, is determined separately for jets originating from b, c and light-flavour quarks and these are plotted as a function of  $p_T$  and  $|\eta|$  in Figure 5.3.

The efficiency to tag a jet originating from a b quark is about 60 – 70% over the entire  $p_T$  range and decreases at higher  $|\eta|$ . A jet originating from a c quark (light-flavour parton) is tagged in about 14% (3%) of cases.

Event weights are determined as

$$w = \frac{P(\text{DATA})}{P(\text{SIM})}, \quad (5.3)$$

and applied to simulated events. The majority of the weights are smaller than one, indicating that the number of events containing b jets is overestimated in simulation when this b-tagging algorithm is used.

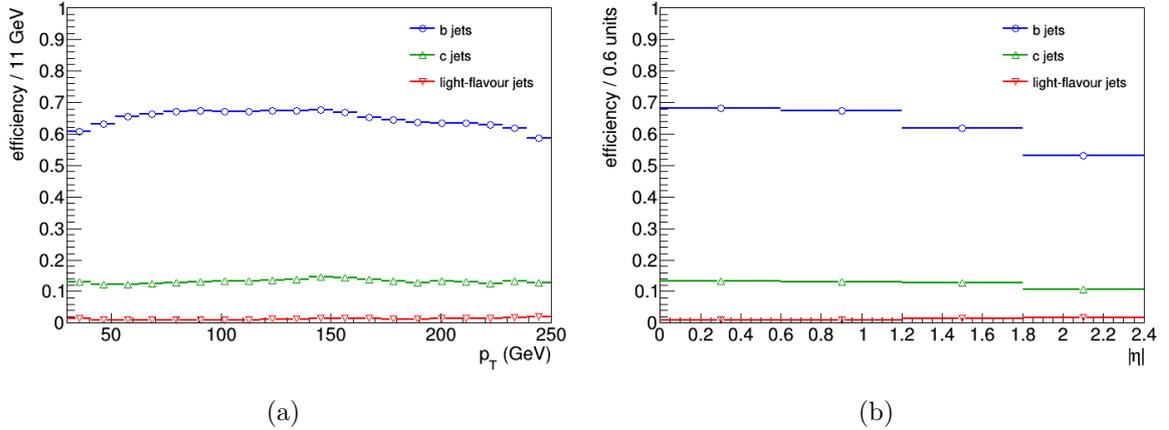


Figure 5.3: Efficiency to tag a jet as a b jet for jets originating from b quarks (blue), c quarks (green), and light-flavour partons (red) as a function of (a)  $p_T$  and (b)  $|\eta|$ .

### Pileup reweighting

The expected amount of pileup interactions is simulated during the detector simulation step explained in Section 3.1.5. This often happens before the data are recorded. Since the actual PU contribution can only be measured in data, the distribution of the number of PVs in the simulation is reweighted to take care of discrepancies. Figure 5.4 shows the distribution of the number of PVs for data and simulation before and after reweighting. Due to the strip tracker dynamic inefficiency, fewer tracks and thus fewer PVs were reconstructed during the early 2016 data. Therefore, the data distribution of the number of PVs peaks at a slightly lower value than expected for this luminosity. As a consequence, there is still a discrepancy between data and simulation in the distribution of the number of PVs, despite the clear improvement after the pileup is reweighted.

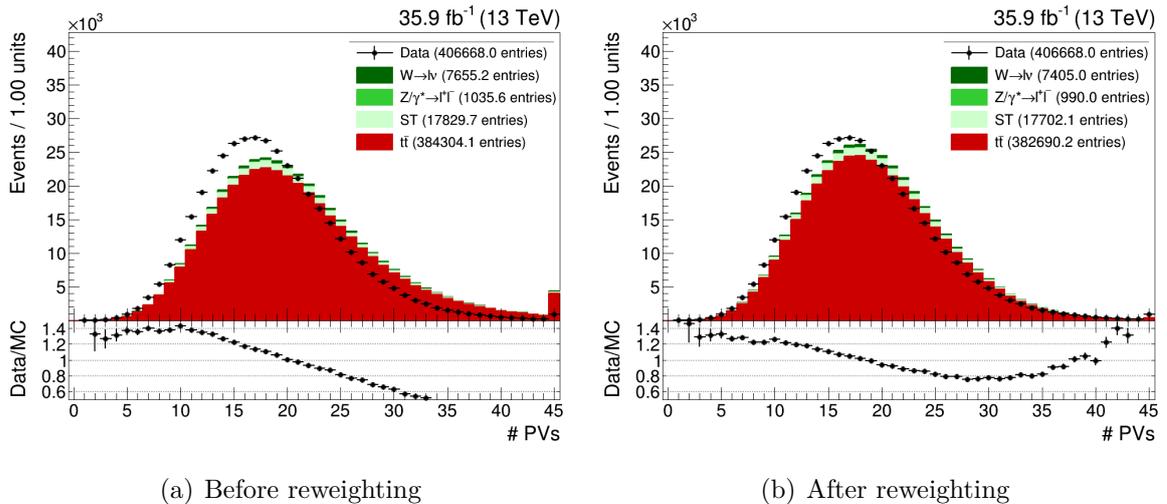


Figure 5.4: Distribution of the number of PVs (a) before and (b) after pileup reweighting. The last bin contains the overflow.

A comparison between data and simulation after applying these scale factors is presented in Figure 5.5 for the distributions of the transverse momentum and pseudorapidity of the muon and the least energetic jet in the event. A good agreement is observed.

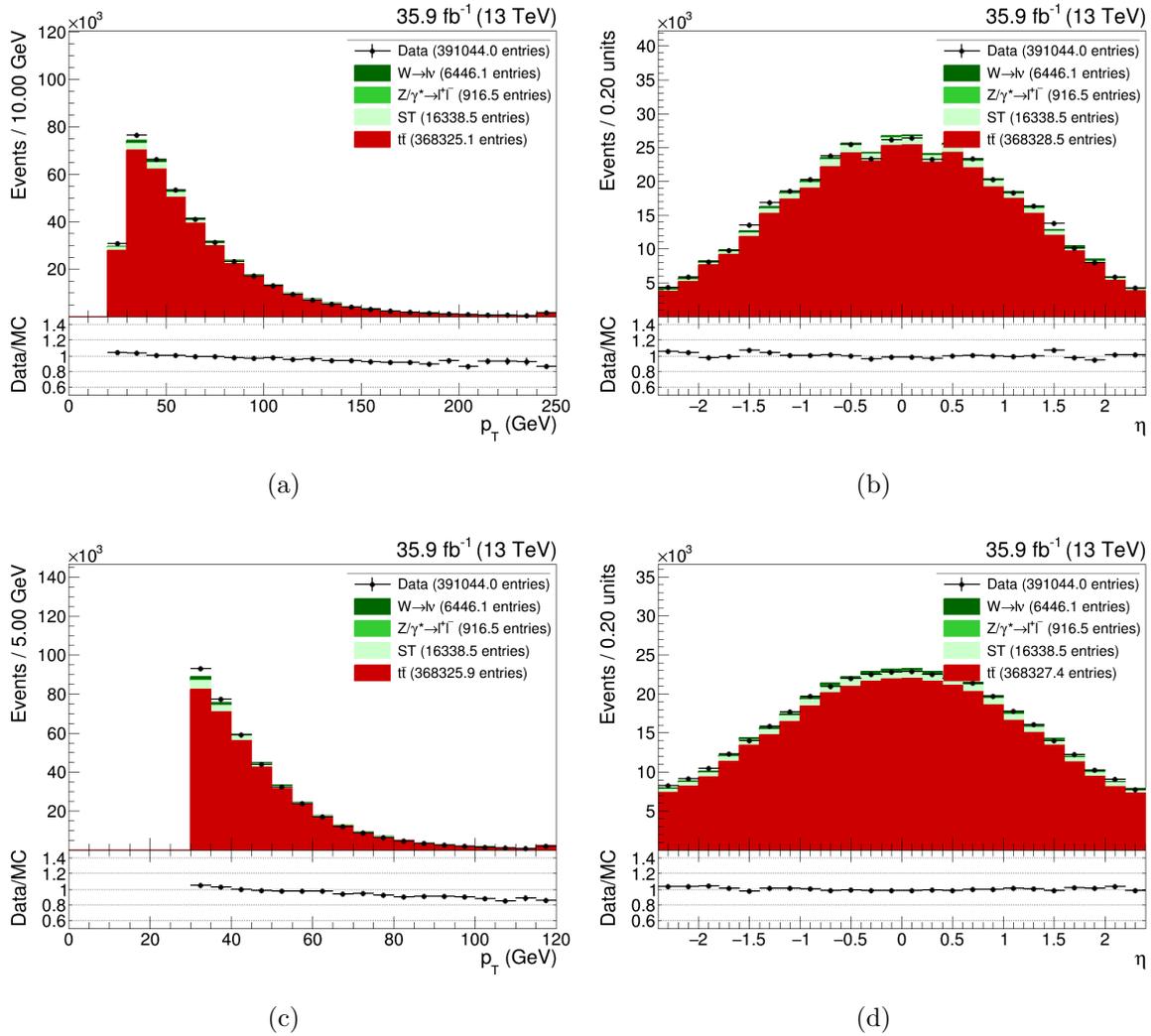


Figure 5.5: Distribution of the (a) transverse momentum and (b) pseudorapidity of the muon and the (c)  $p_T$  and (d)  $\eta$  of the least energetic jet. The last  $p_T$  bin contains the overflow.

## 5.3 Event Reconstruction

After the event selection the sample has a purity of 93% for pair-produced top quarks, which are now reconstructed. As can be seen in Figure 5.1, the top quark decays into a b quark and a W boson, which in its turn can decay into a pair of quarks or a charged lepton and a neutrino. A semileptonically decaying top quark pair can thus be divided into a hadronically decaying top quark system, represented as  $bq\bar{q}'$ , and a leptonically decaying one,  $bl\nu$ .

Section 5.3.1 describes how events are reconstructed using generator information, while Section 5.3.2 explains the reconstruction procedure for both simulated events and data, where no generated event information is available. A comparison between both methods is made for simulated events, such that the reconstruction efficiency can be estimated.

### 5.3.1 Reconstruction of generated top quark events

In addition to information related to the reconstructed leptons, jets and missing transverse energy, simulated events also contain information that enables to reconstruct how the event was generated. The ‘generator particles’ are produced by `PYTHIA 8` and, besides their four-momentum and charge, they also contain identification (ID) markers, indicating their own type and that of the particles they originated from (‘mothers’ and ‘grannies’) according to the Particle Data Group conventions [8], and so-called status codes. During the PS and hadronisation stages of the simulation, `PYTHIA 8` keeps track of what happens to each individual particle. E.g. when a particle radiates a gluon or a photon, the energy of the particle decreases. In order to store the properties of the particle before and after the radiation, `PYTHIA 8` creates a copy of the original particle which is then attributed the altered momentum. The two copies can be distinguished by their status codes, which reflect if the particle has radiated or not.

Status codes thus give information on how particles are produced and have evolved during the simulation process. This implies that also particles originating from beam remnants, MPIs or ISR/FSR can be identified using status codes. As there are many possible status codes, only those relevant for this thesis will be described below. The complete list can be found in [166].

Particles involved in the hard-scattering process have status codes between 20 and 29. For top quark production and decay only status 22 and 23, representing an intermediate hard particle such as a top quark and the outgoing partons from its decay, respectively, are relevant. Final-state particles have status 1. In practice, only the muon in this event topology will have a copy with status 1, as all partons hadronise and form jets. Whenever a choice is possible, a preference for muons with status 23 will be given, since at this stage no (photon) radiation has taken place yet and the original muon momentum and direction is preserved.

Similarly, the generated top quarks used in the reweighting procedure in Section 3.3 have status 22. In addition, generated top quarks with status 62 are used for determining the systematic uncertainty due to top- $p_T$  reweighting (see Section 6.3.1). These represent

top quarks after radiation, so with an altered  $p_T$  with respect to the originally generated top quark, but before their decay.

The hard-scattering process of the generated event can be reconstructed using the IDs of the generator particles. Each event should contain two oppositely charged top quarks, of which one decays hadronically and the other leptonically. It is important to note that, although the reconstructed events are semileptonic, this is not necessarily the case at generator level. E.g. when one of the leptons in a dilepton event is subject to photon radiation, it might be perceived as a non-isolated lepton in a jet. Events like this are easily spotted when their generator information is used. When they are encountered in this analysis, they will not be matched, but proceed directly to the event reconstruction described in Section 5.3.2.

### 5.3.1.1 Jet-parton matching

The jet-parton matching is based on a minimum-distance algorithm, where the maximum angular distance  $\Delta R$  between the parton and the selected jet has to be smaller than 0.3. As described above, only one copy of each generated parton will be considered in the matching procedure.

First, all partons are ordered in decreasing  $p_T$  and, starting from the hardest parton, the angular distance with respect to each jet is calculated. When the minimum  $\Delta R < 0.3$ , the jet is matched to the parton and removed from the jet collection to avoid matching the same jet to several partons. If  $\Delta R > 0.3$  for all jets, the parton remains unmatched. The procedure is repeated for all partons.

In about 37% of the  $t\bar{t}$  events it is possible to match all four partons from the semileptonic decay to jets. This increases to 40% if only the partons originating from the hadronic top quark decay are to be matched. This low matching efficiency has multiple reasons. Firstly, all selected jets have to abide the jet requirements explained earlier. If e.g. the jet  $p_T$  is too low or its  $|\eta|$  is too large, the jet will not be selected and the parton cannot be matched. Secondly, the direction of the jet can change if it is subject to FSR. If the change in direction is large, the angular distance between the jet and the parton can become larger than 0.3 and the parton will not be matched.

### 5.3.1.2 Resolution functions

The reconstructed and calibrated jets provide an estimate of the momentum and direction of the partons from which they originate. By studying the difference between the kinematic properties of a quark and those of its matched jet, the resolution of the momentum and direction is determined. The resolution functions are determined separately for jets originating from b and light-flavour quarks. To this purpose, only  $t\bar{t}$  events where at least the three partons originating from the hadronically decaying top quark can be matched to jets are considered. In order to increase the precision of the resolution functions, matched events with more than four jets are included as well.

For each quark and its matched jet the difference in  $E_T$ ,  $\theta$  and  $\phi$  is calculated and plotted as a function of the jet  $E_T$ . Three pseudorapidity regions are considered, i.e.  $|\eta| < 1.3$ ,  $1.3 < |\eta| < 1.5$  and  $|\eta| > 1.5$ , which roughly correspond to the calorimeter

barrel (B), overlap (O) and endcap (E) regions, respectively. The normalised distribution of the  $E_T$  difference  $\Delta E_T = E_T(\text{parton}) - E_T(\text{jet})$  in the barrel region is visualised in Figure 5.6 for b and light-flavour jets separately. Several jet  $E_T$  regions are considered. In both cases, the width of the distribution becomes larger when the transverse jet energy

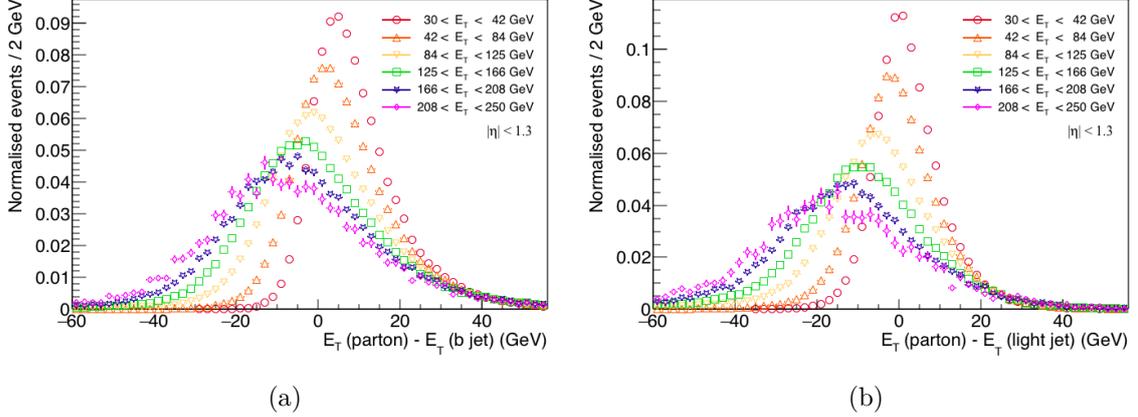


Figure 5.6: Normalised distribution of the difference between the jet and parton  $E_T$  in the barrel region as a function of the jet  $E_T$  for (a) b jets and (b) light-flavour jets.

increases. Further, a higher number of b jets have  $E_T(\text{parton}) > E_T(\text{jet})$  compared to light-flavour jets, since b hadrons also have leptonic decay modes involving neutrinos. This reflects itself in a reconstructed jet energy that is lower than expected.

In each jet  $E_T$  region the  $\Delta E_T$  distribution can be parameterised by two Gaussian functions,

$$R(E_T) = \frac{1}{2\pi\sqrt{a_2^2 + a_3 a_5^2}} \left[ \exp\left(-\frac{(\Delta E_T - a_1)^2}{2a_2^2}\right) + a_3 \exp\left(-\frac{(\Delta E_T - a_4)^2}{2a_5^2}\right) \right], \quad (5.4)$$

where  $a_1$  and  $a_2$  are, resp., the mean and the standard deviation of the Gaussian function that describes the top of the distribution and  $a_4$  and  $a_5$  the mean and the standard deviation of a broader Gaussian describing the tails of the distribution. The parameter  $a_3$  takes account of the relative contribution of the Gaussian functions.

The  $\Delta E_T$  distribution is fitted using the function in Eq. (5.4) and, in order to get a fluent parameter description as a function of  $E_T$ , the parameter values obtained in each jet  $E_T$  region are then fitted with a linear function. Inserting the  $E_T$  dependence of the parameters back into Eq. (5.4), a function as the one visualised in Figure 5.7 is obtained. One can see that the distribution is fairly narrow at small  $E_T$  and becomes broader when  $E_T$  increases. The width of the distribution corresponds to the  $\Delta E_T$  resolution and is described well by parameter  $a_2$ . The linear function describing its behaviour as a function of  $E_T$  is called the resolution function and is shown in Figure 5.8(a) for light-flavour jets in the three  $|\eta|$  regions. There is no strong dependence of the resolution on the  $|\eta|$  region.

Also the distributions of  $\Delta\theta$  and  $\Delta\phi$  can be described by Eq. (5.4) and the resolution functions for these parameters are determined correspondingly. Note that the angular resolution improves at higher energies, since high-energy particles are bent less in the

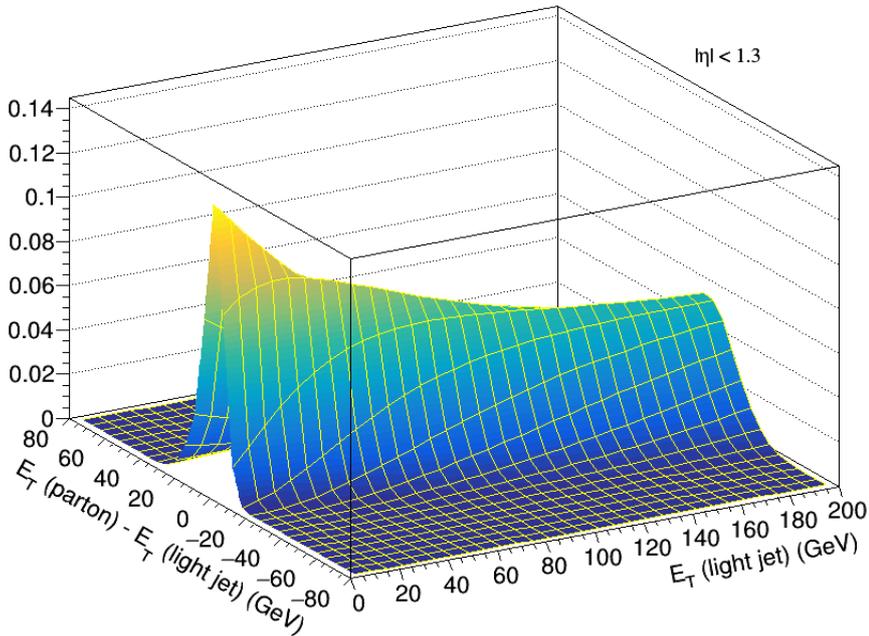


Figure 5.7: Normalised distribution of the difference between the parton  $E_T$  and that of its corresponding light jet as a function of the jet  $E_T$  in the barrel region.

magnetic field of the CMS detector and the direction of the jet is therefore better reconstructed. This effect is not linear, so a term proportional to  $\sqrt{E_T}$  is added to fit the resolution functions. These are visualised for light-flavour jets in Figures 5.8(b) and 5.8(c), respectively. For  $\theta$ , a dependence of the resolution on the  $|\eta|$  region is observed. This is an effect of the used matching criterion. A certain angle  $\Delta\theta$  describes a larger  $\Delta\eta$  in the endcap than in the barrel region. Therefore, a requirement on  $\Delta R$ , which is related to  $\Delta\eta$ , results in a better precision of  $\Delta\theta$  in the endcap region.

### 5.3.2 Reconstruction of the top quark pair system

Contrary to simulated events, it is not possible to get information about the generator particles for data. Other methods are thus needed to reconstruct the event to perform a measurement of the top quark width. If an event has exactly four jets, two of these should be originating from the b quarks. If more than two jets in the event are b tagged, only those with the highest CSVv2M discriminant values will be considered as b jet candidates. For each of these jets, the radial distance  $\Delta R$  with respect to the lepton is calculated. The b jet candidate with the smallest  $\Delta R$  is considered to be the jet associated with the decay  $t \rightarrow bW \rightarrow b\mu\nu$ . As the neutrino is not detected, the four-momentum of the leptonically decaying top quark is approximated by summing the four-momenta of the b jet candidate and the lepton. The four-momenta of the remaining b jet candidate and the two light-flavour jets are summed to reconstruct the four-momentum of the hadronically decaying top quark ( $t \rightarrow bW \rightarrow bq\bar{q}'$ ).

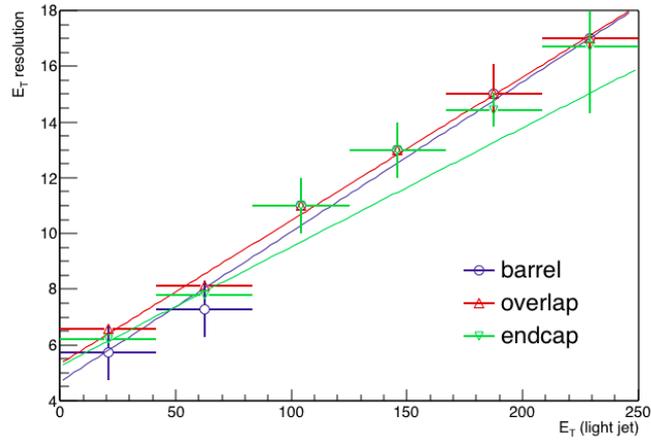
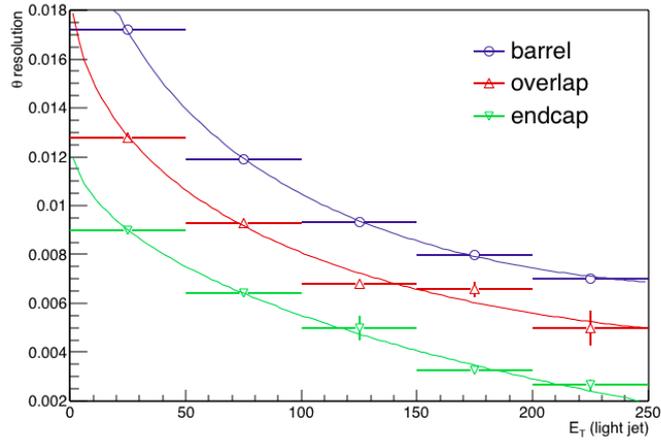
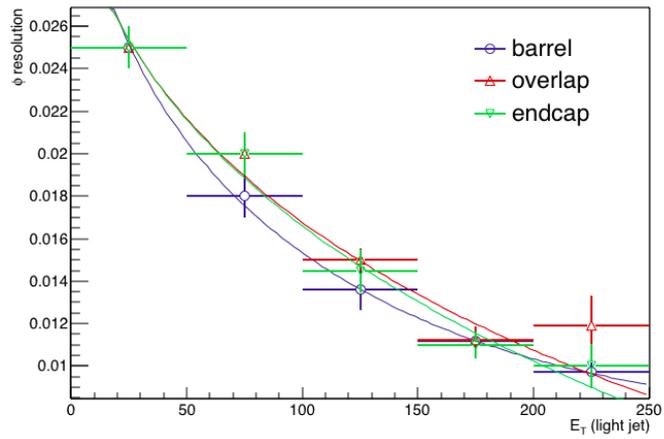
(a)  $E_T$ (b)  $\theta$ (c)  $\phi$ 

Figure 5.8: Resolution functions for (a)  $E_T$ , (b)  $\theta$ , and (c)  $\phi$  as a function of the jet  $E_T$  for light-flavour jets.

### 5.3.2.1 Categorisation of events

The reconstruction efficiency for the hadronically decaying top quark is estimated by comparing the 3-jet combination used for the reconstruction to the collection of jets that are matched to the partons originating from the  $t \rightarrow b\bar{q}q'$  decay. Simulated events can thus be subdivided into different categories, based on how the reconstruction of the event compares to the generator information.

#### Correctly matched (CM)

The 3-jet combination that reconstructs the hadronically decaying top quark is equal to the jet combination that is matched to the  $b\bar{q}q'$  partons from the top quark decay. Note that any permutation of the jets leads to a correct top quark mass reconstruction.

#### Wrongly matched (WM)

At least one jet in the 3-jet combination that reconstructs the hadronically decaying top quark is not matched to one of the  $b\bar{q}q'$  partons from the top quark decay.

#### Unmatched (UM)

During the matching procedure at least one of the  $b\bar{q}q'$  partons from the top quark decay is not matched to any jet. Hence a correct match cannot be found.

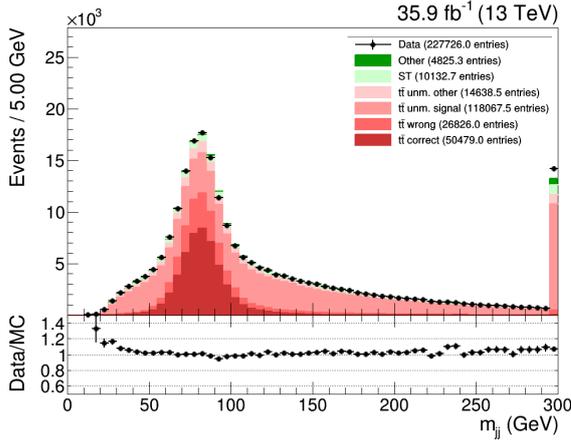
By definition, non- $t\bar{t}$  events are categorised as UM. Simulated  $t\bar{t}$  events, on the other hand, can fall into each of these categories and these will be treated separately in the following sections. A further distinction is made between unmatched  $t\bar{t}$  events with a semileptonic decay and other decay modes of the top quark pair on generator level.

Considering only the selected  $t\bar{t}$  events that decay semileptonically, about 26% is correctly matched, 14% is wrongly matched and 60% is unmatched. This means that about 65% of the events that can be matched is correctly reconstructed. The invariant mass distributions of the reconstructed W boson,  $m_{jj}$ , and top quark,  $m_{bjj}$ , are plotted in Figure 5.9. Whereas the CM events are centred around the expected mass values, the WM and UM events are distributed over the entire mass range. This is more pronounced for UM events due the large amount of events in this category. As a consequence, the distributions of the masses have large tails. If only events with a reconstructed top quark mass between 100 GeV and 245 GeV are considered, the relative fraction of CM events increases to 38%. This requirement reduces the absolute number of CM events by merely 1%, while the decrease in WM and UM events is equal to 29% and 47%, respectively. The resolution of the reconstructed mass of the top quark is further improved using the precisely measured W boson mass as a constraint in a kinematic fit (KF).

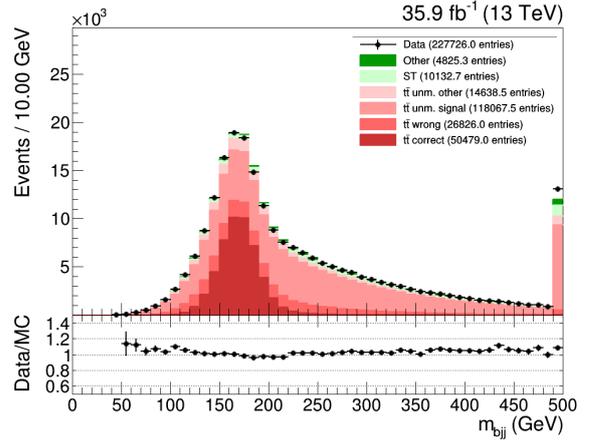
### 5.3.2.2 Kinematic fit

The kinematic fit procedure is described in detail in [167]. A short overview is given below.

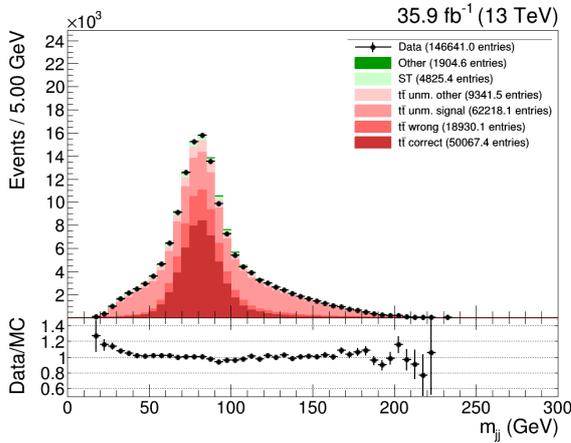
A general event topology with  $n$  particles can be described by a number of constraints, such as energy and momentum conservation. If there are  $m$  constraints and  $\bar{y}_i$  is the



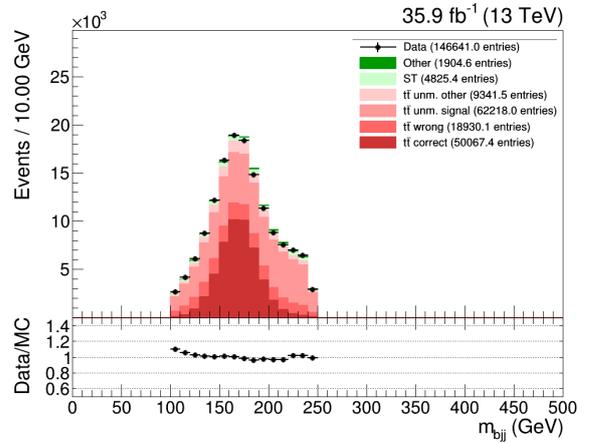
(a) W boson mass before req.



(b) t quark mass before req.



(c) W boson mass after req.



(d) t quark mass after req.

Figure 5.9: Distribution of the invariant mass of the reconstructed W boson (left) and top quark (right), before (top) and after (bottom) requiring  $100 < m_t < 245$  GeV. The last bin contains the overflow.

true four-momentum of a measured particle  $i$ , the system is described by

$$\begin{aligned} f_1(\bar{y}_1, \bar{y}_2, \dots, \bar{y}_n) &= 0, \\ &\vdots \\ f_m(\bar{y}_1, \bar{y}_2, \dots, \bar{y}_n) &= 0. \end{aligned} \tag{5.5}$$

The measured particle four-momenta  $\vec{y}$  are often not equal to  $\bar{y}$  and thus not necessarily solve the set of equations above. Therefore, they are varied within their uncertainties and the equations are solved for  $\vec{y}' = \vec{y} + \Delta\vec{y}$ . In practice, the  $\Delta\vec{y}$  corrections are determined by minimising the weighted sum

$$S(\vec{y}) = \Delta\vec{y}^t V^{-1} \Delta\vec{y}, \tag{5.6}$$

where  $V$  is the covariance matrix of the measured four-momenta. So-called Lagrange multipliers  $\vec{\lambda}$  are introduced to minimise  $S(\vec{y})$  under the conditions in Eq. (5.5),

$$L(\vec{y}, \vec{\lambda}) = S(\vec{y}) + 2 \sum_{k=1}^m \lambda_k f_k(\vec{y}). \tag{5.7}$$

This is a least-squares minimisation problem that can be solved analytically if the constraints depend linearly on the measured four-momenta. If this is not the case, an iterative procedure is applied to solve Eq. (5.7), while linearising the functions in Eq. (5.5). This is the case when e.g. a mass constraint is imposed. If the KF fails to converge within 30 iterations, the event is rejected. When it does converge, the quality of the fit is described by  $\chi^2/\text{NDF}$ , where the number of degrees of freedom NDF is defined as the difference between the number of constraints and the amount of unmeasured quantities. If there are no unmeasured quantities, the number of degrees of freedom is equal to the number of constraints.

In this analysis, only one constraint is set, namely the invariant mass of the two jets corresponding to the  $W \rightarrow q\bar{q}'$  decay is required to be equal to the  $W$  boson mass of 80.385 GeV. The jets are parameterised using the  $(E_T, \theta, \phi)$  coordinates, which are related to the jet energy and momentum as

$$E = \frac{E_T}{\sin \theta}, \quad \vec{p} = \begin{pmatrix} E_T \cos \phi \\ E_T \sin \phi \\ E_T \cot \theta \end{pmatrix}. \tag{5.8}$$

The covariance matrix for the two light-flavour jets is set up using the resolution functions determined in Section 5.3.1. As  $E_T$ ,  $\theta$  and  $\phi$  are uncorrelated, this is a diagonal matrix, where the variance of each coordinate is determined by evaluating the resolution functions in the jet  $p_T$  and  $\eta$ .

Events that fulfil the constraint without altering the jet four-momenta too much will have a small KF  $\chi^2$  value. It is therefore expected that correctly matched events generally have smaller  $\chi^2$  values than events that are not well-reconstructed. This is confirmed in Figure 5.10, which shows the  $\chi^2$  value for all event categories separately. In order to increase the relative amount of CM events, a cut is imposed on the  $\chi^2$  value of

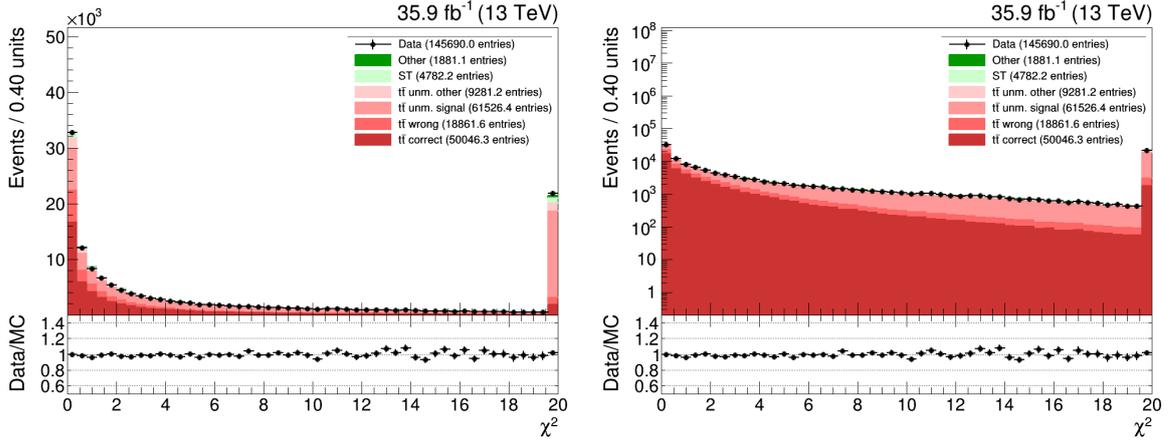


Figure 5.10: Distribution of the KF  $\chi^2$  value in linear and logarithmic scale. The last bin contains the overflow.

the KF. Balancing a good background rejection and a decrease in the absolute number of CM events, an event is required to have a  $\chi^2 < 15$  in order to be selected for this analysis. This cut reduces the number of selected CM, WM and UM events by 5%, 10% and 33%, respectively. As a consequence, the fraction of CM events increases to about 45% of all semileptonically decaying  $t\bar{t}$  events. In addition, the distribution of the reconstructed top quark mass becomes narrower and more peaked around the expected top quark mass value, as can be seen in Figure 5.11.

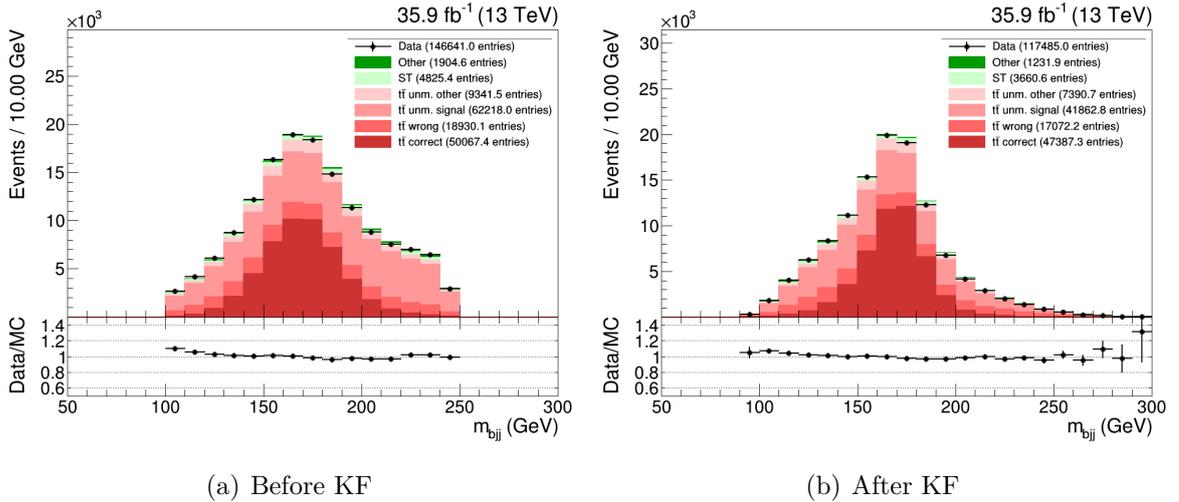


Figure 5.11: Distribution of the reconstructed top quark mass (a) before and (b) after performing a KF. The last bin contains the overflow.

### 5.3.2.3 Further selection requirements

Up till now the fraction of CM events has been increased by removing events belonging to other categories from the collection of selected events. At this point, the reconstruction of the top quark is revisited in order to increase the reconstruction efficiency. The hadronic top quark mass  $m_{b_{jj}}$  is equal to the invariant mass of the combination of the light-flavour jets and the b jet candidate that is farthest away from the lepton. An alternative top quark mass  $m_{b'_{jj}}$  is devised by calculating the invariant mass of the light-flavour jets and the b jet candidate that is closest to the lepton. The distribution of this variable is plotted as a function of  $m_{b_{jj}}$  in Figure 5.12 for CM and WM events. Whereas the top quark mass is in general smaller than the alternative top quark mass for CM events, the opposite is observed for WM events. This seems to suggest that some of the WM events could become CM events if the b jet candidates are swapped in the reconstruction.

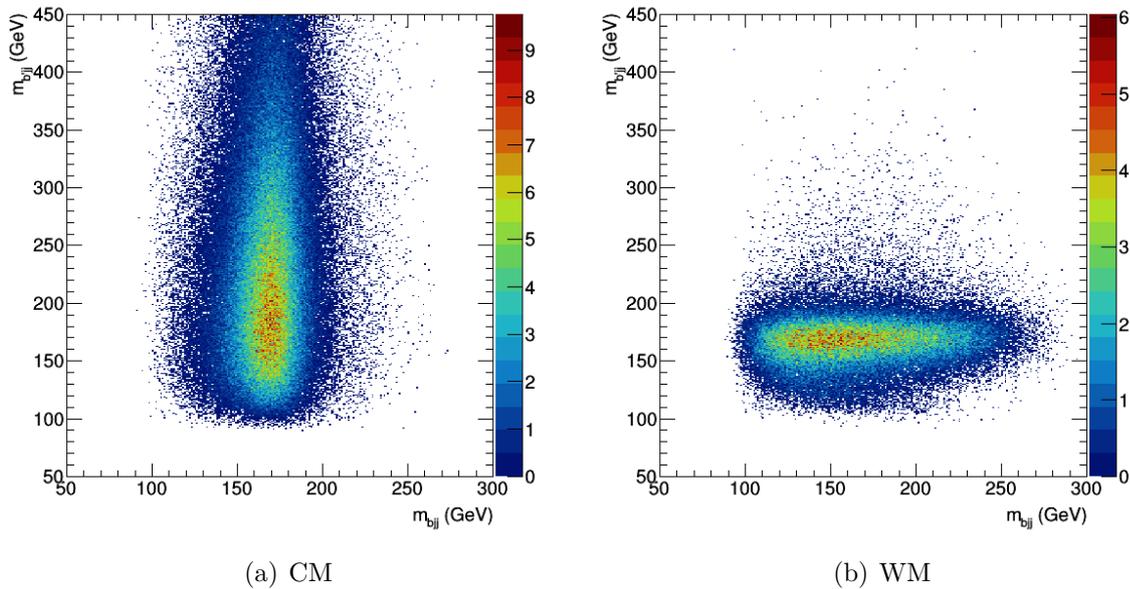


Figure 5.12: Distribution of the alternative top quark mass as a function of the top quark mass for (a) CM and (b) WM events. The colour scale indicates the number of events in each bin normalised to the recorded integrated luminosity.

Using only CM events, the average top quark mass is calculated and a value of  $\langle m_{b_{jj}} \rangle_{CM} = 169.3 \text{ GeV}$  is obtained. Then, the differences between the reconstructed masses and the average top quark mass are determined and when

$$\left| m_{b_{jj}} - \langle m_{b_{jj}} \rangle_{CM} \right| - \left| m_{b'_{jj}} - \langle m_{b_{jj}} \rangle_{CM} \right| > 8 \text{ GeV}, \quad (5.9)$$

the b jet candidates are swapped, i.e. the hadronic top quark is reconstructed by adding the four-momenta of the light-flavour jets and the b jet candidate that is closest to the lepton. This reduces the number of events in the WM category by 37%, such that they constitute only one tenth of the total number of selected events. The amount of CM events, on the other hand, is enhanced by 13%. The effect of this swap on the

distribution of  $m_{b'_{jj}}$  as a function of  $m_{b_{jj}}$  is visualised in Figure 5.13 and the number of selected events can be found in Table 5.3.

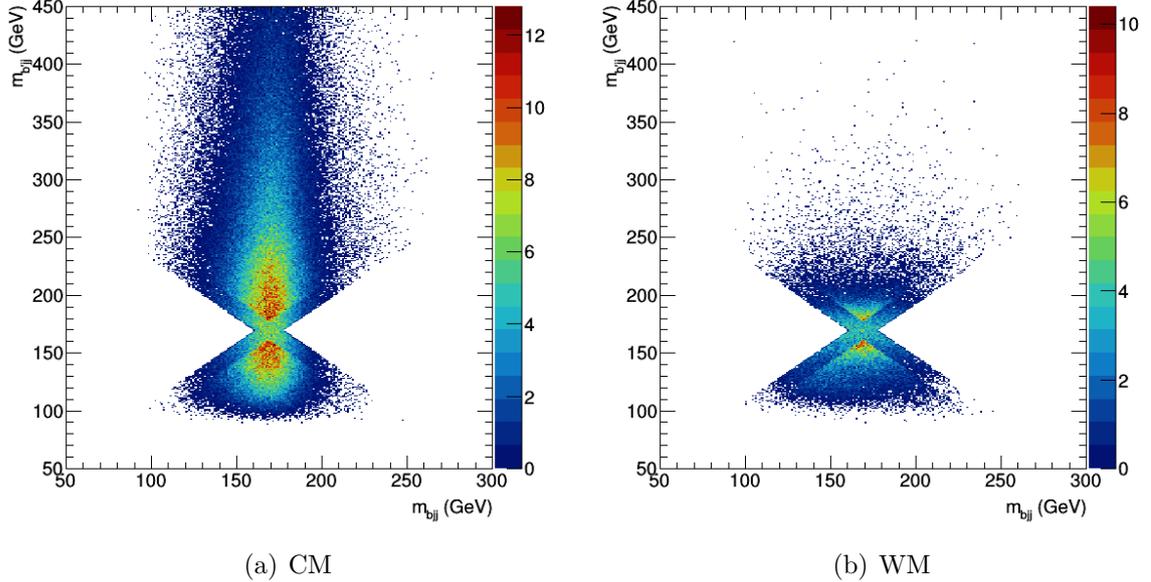


Figure 5.13: Distribution of the alternative top quark mass as a function of the top quark mass for (a) CM and (b) WM events after applying the requirement in Eq. (5.9). The colour scale indicates the number of events in each bin normalised to the recorded integrated luminosity.

As can be seen in Figure 5.11, the kinematic fit has altered the jet four-momenta in such a way that the reconstructed top quark mass extends beyond the previously defined range. Therefore, the requirement  $100 < m_{b_{jj}} < 245$  GeV is applied again. Figure 5.14 shows the distribution of the reconstructed  $m_{lb}$  mass and the distribution of the difference between the reconstructed hadronic top quark mass and the reconstructed  $m_{lb}$  mass. The amount of CM events is optimised by requiring  $m_{lb} < 200$  GeV and  $m_{b_{jj}} - m_{lb} > 0$  GeV.

At this point, the majority of the WM and UM events have an alternative top quark mass that is lower than the average reconstructed top quark mass for correctly matched events. As can be seen in Figure 5.15, the fraction of CM events can be drastically increased by requiring  $m_{b'_{jj}} > 200$  GeV. Although the amount of CM events is reduced by about 39%, the effect is much stronger for the other categories and the WM events are almost completely obliterated.

The selection requirements described above are summarised in Table 5.3 together with their effect on the different event categories. Considering all simulated events, about 59% is reconstructed correctly, 1% is not well-reconstructed and 40% cannot be matched. The purity of the sample increased from 93% to 97%.

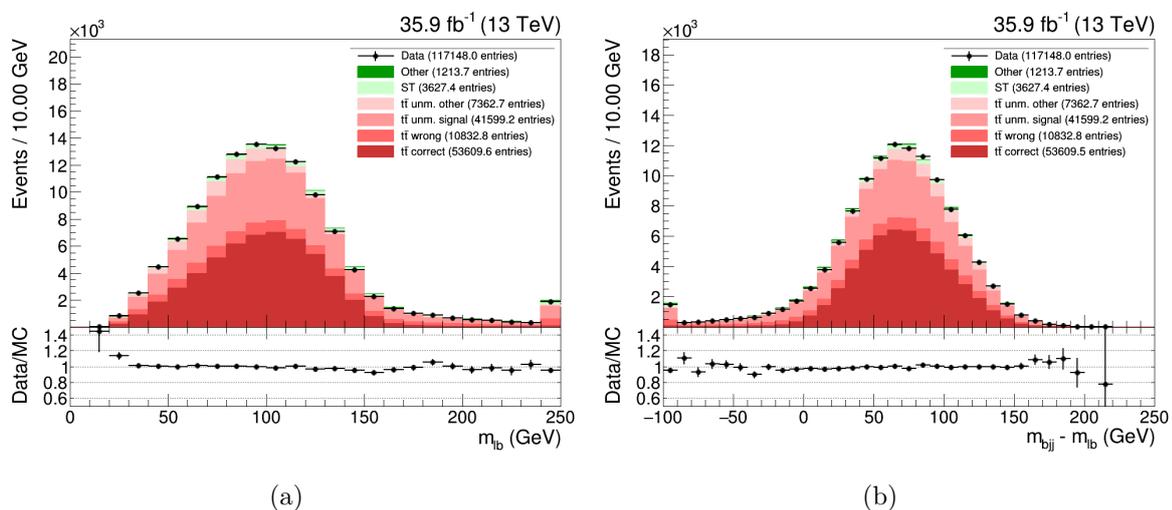
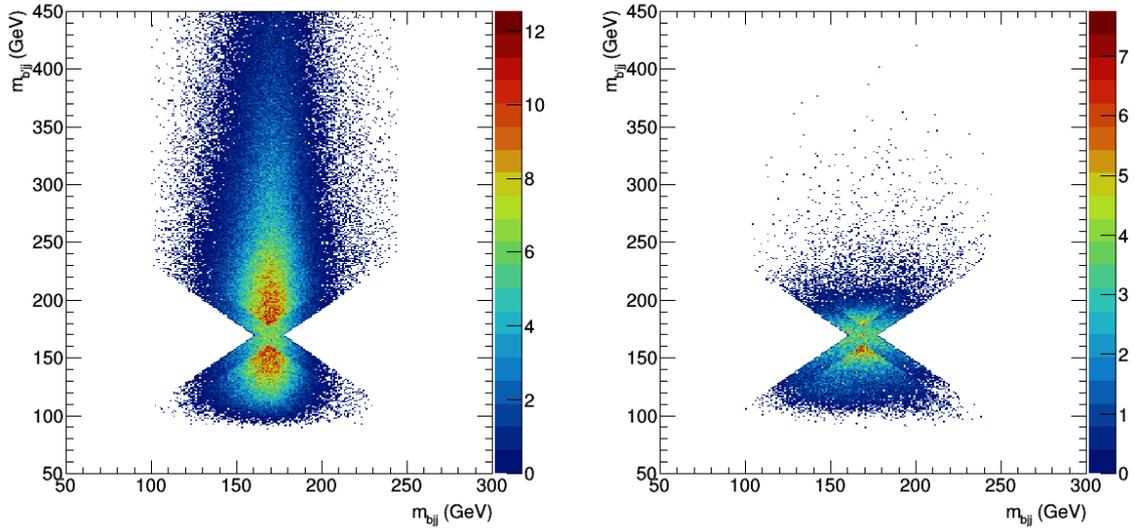
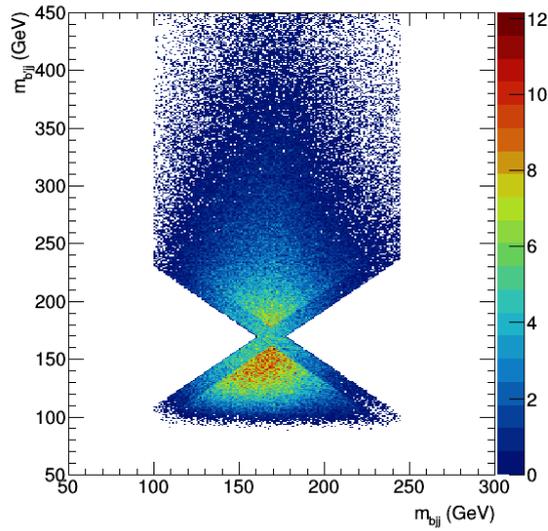


Figure 5.14: Distribution of (a) the reconstructed  $m_{lb}$  mass and (b) the difference between the reconstructed hadronic top quark mass and the reconstructed  $m_{lb}$  mass. The first (last) bin contains the underflow (overflow).



(a) CM

(b) WM



(c) UM

Figure 5.15: Distribution of the alternative top quark mass as a function of the top quark mass for (a) CM, (b) WM, and (c) UM events. The colour scale indicates the number of events in each bin normalised to the recorded integrated luminosity.

Table 5.3: Evolution of the number of selected events after the selection requirements on reconstructed particles, starting from the basic event selection described in Section 5.2. ( $35.9 \text{ fb}^{-1}$  of int. lumi.)

	$t\bar{t}$ CM	$t\bar{t}$ WM	$t\bar{t}$ UM semilep.	$t\bar{t}$ UM other	single top	other	Total exp.	Observed
Basic selection	50479 $\pm$ 225	26826 $\pm$ 164	118067 $\pm$ 344	14638 $\pm$ 121	10133 $\pm$ 101	4825 $\pm$ 69	224969 $\pm$ 474	227726
$100 < m_{b,jj} < 245 \text{ GeV}$	50067 $\pm$ 224	18930 $\pm$ 138	62218 $\pm$ 249	9341 $\pm$ 97	4825 $\pm$ 69	1905 $\pm$ 44	147287 $\pm$ 384	146641
KF $\chi^2 < 15$	47387 $\pm$ 218	17072 $\pm$ 131	41863 $\pm$ 205	7391 $\pm$ 86	3661 $\pm$ 61	1232 $\pm$ 35	118605 $\pm$ 344	117485
Swap	53623 $\pm$ 232	10836 $\pm$ 104	41863 $\pm$ 205	7391 $\pm$ 86	3661 $\pm$ 61	1232 $\pm$ 35	118605 $\pm$ 344	117485
$100 < m_{b,jj} < 245 \text{ GeV}$	53610 $\pm$ 232	10833 $\pm$ 104	41599 $\pm$ 204	7363 $\pm$ 86	3627 $\pm$ 60	1214 $\pm$ 35	118245 $\pm$ 344	117148
$m_{lb} < 200 \text{ GeV}$	53484 $\pm$ 231	9805 $\pm$ 99	39766 $\pm$ 199	7216 $\pm$ 85	3206 $\pm$ 57	1020 $\pm$ 32	114496 $\pm$ 338	113535
$m_{b,jj} - m_{lb} > 0 \text{ GeV}$	52931 $\pm$ 230	8936 $\pm$ 95	37378 $\pm$ 193	7022 $\pm$ 84	2899 $\pm$ 54	877 $\pm$ 30	110042 $\pm$ 332	109200
$m_{b'jj} > 200 \text{ GeV}$	32522 $\pm$ 180	719 $\pm$ 27	17197 $\pm$ 131	3621 $\pm$ 60	1311 $\pm$ 36	292 $\pm$ 17	55662 $\pm$ 236	53553



# 6

## Measurement of the Top Quark Width

---

The measurement procedure to determine the top quark decay width is based on a maximum likelihood (ML) method, where the probability density functions are constructed using a variable that is sensitive to changes in the decay width. This method treats every event separately, which allows to use different probability density functions for each event category. When a large number of events is considered, the ML function approaches a Gaussian distribution. This means that the logarithm of the likelihood has a parabolic shape in the neighbourhood of the ML estimator, which is in this case the top quark decay width  $\hat{\Gamma}_t$ . The procedure to set up the likelihood function is described in Section 6.1, which also explains how the estimator is calibrated. This method is extended to other variables and combinations of these in Section 6.2 and the systematic uncertainties that affect the measurement are discussed in Section 6.3. Finally, the top quark decay width is determined using data events in Section 6.4.

### 6.1 Measurement Procedure using One Variable

The top quark decay width  $\Gamma_t$  is estimated by applying the ML method, where probability density functions are constructed using the width of the invariant mass distribution of the hadronically decaying top quark. It was briefly discussed in Section 1.2 that the value of the top quark decay width has a dependency on the top quark mass. Therefore, a new variable is introduced to describe the shape of the top quark mass distribution without making assumptions about the value of the top quark mass itself. The reduced top quark mass  $m_r$  is defined as the reconstructed mass of the hadronically decaying top quark divided by the average top quark mass,

$$m_r = \frac{m_{bjj}}{\langle m_{bjj} \rangle_{CM}}, \quad (6.1)$$

where  $\langle m_{bjj} \rangle_{CM} = 169.3$  GeV is determined using only correctly reconstructed events. The distribution of the reduced top quark mass can be found in Figure 6.1. Comparing the normalised distributions of each data set, it is clear that not only the CM events are centred around one, which is true by definition, but also the data peak at that point. The larger tails are due to the selected background events, which have a much broader distribution.

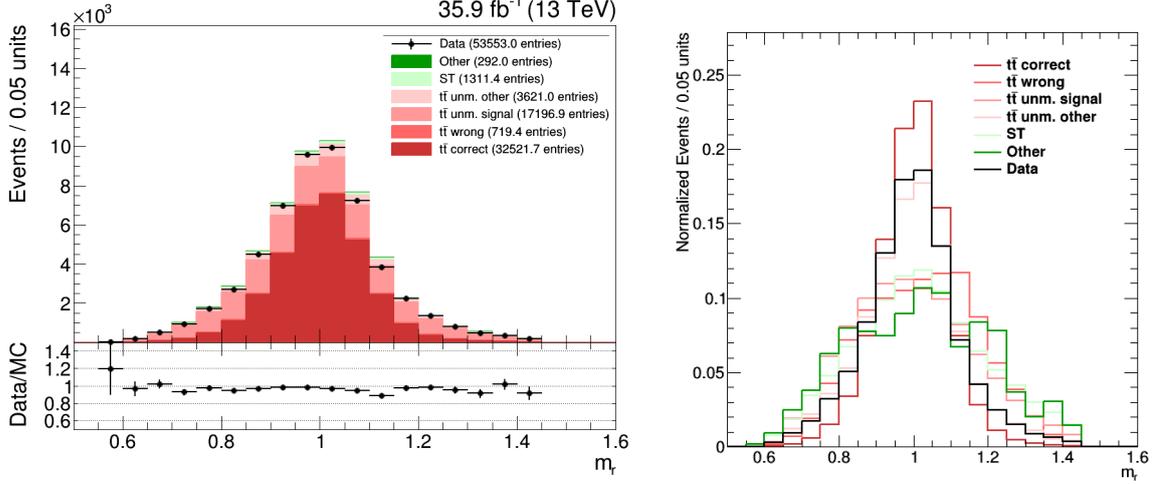


Figure 6.1: Distribution of the reduced top quark mass (left) and the normalised distribution for each data set separately (right).

The reweighting procedure described in Section 3.3 is used to obtain samples with different values of  $\Gamma_t$  and, unless mentioned otherwise, it is implicit that the  $t\bar{t}$  events are reweighted when distributions with non-nominal widths are considered. The reweighting to obtain a sample with a top quark width of  $s\Gamma_{t,\text{gen}}$  is performed for values of  $s \in [0.1, 1.5]$  in steps of 0.05 and for  $s \in [1.5, 6]$  in steps of 0.1.

### 6.1.1 Construction of a likelihood function

For each event category a histogram with 90 bins on the range  $[0.5, 2]$  is filled. This is done for all values of  $s$  specified above. Due to the low number of WM events after the event selection, this category is merged with the UM category, in order to reduce the effect of statistical fluctuations. For simplicity, these events will also be addressed as UM events. The histograms are then normalised and their shape is cast into a function by connecting the function values of the bin centres with a fluent line. Each histogram is thus transformed into a template function for a certain event category considering a certain top quark width. The templates are visualised in Figure 6.2 for the nominal top quark width ( $s = 1$ ).

Due to the normalisation, each template represents the probability to have a CM or UM event with a certain reduced top quark mass. As the data cannot be categorised, a total probability function per top quark width is constructed by combining the templates for each category. The relative weight of each template is determined by the number of simulated events in that category. The total probability for an event  $i$  to have a reduced top quark mass  $m_{r,i}$  when a top quark width  $\Gamma_t$  is assumed, is equal to

$$P_i(m_{r,i}|\Gamma_t) = f_{\text{CM}} P_{\text{CM}}(m_{r,i}|\Gamma_t) + f_{\text{UM}} P_{\text{UM}}(m_{r,i}|\Gamma_t), \quad (6.2)$$

where  $f_{\text{XM}}$  represents the fraction of the number of simulated XM events to the total amount of simulated events considered in the procedure. The total probability for

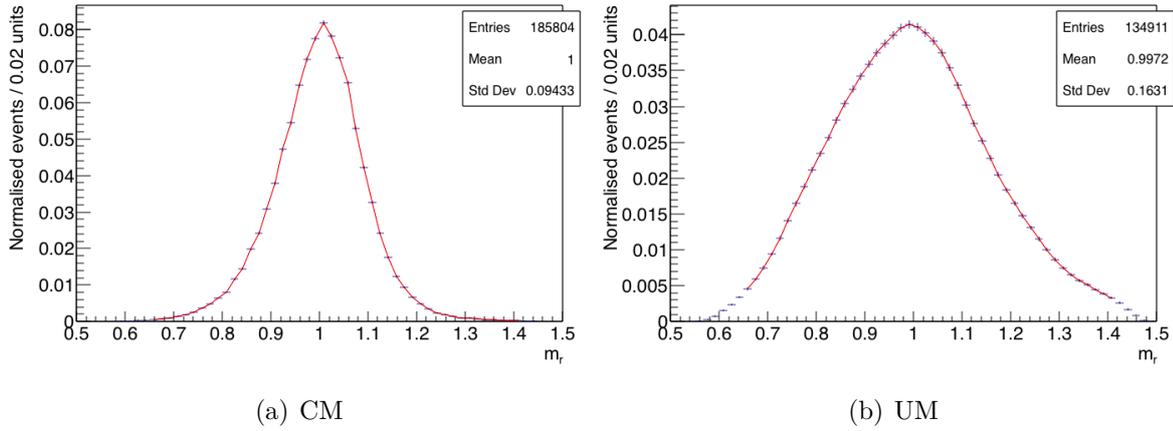


Figure 6.2: Distribution of the reduced top quark mass considering  $\Gamma_t = 1 \times \Gamma_{t,\text{gen}}$  for (a) correctly matched and (b) unmatched events. The red curves represent the templates, while the blue markers indicate the bin contents of the histograms and their respective uncertainties. The distributions are normalised on the interval 0.65 to 1.4.

the nominal top quark width is plotted in Figure 6.3. The likelihood function is then constructed by inserting Eq. (6.2) into the general expression for the logarithm of the likelihood,

$$\ell(\Gamma_t) = \log \mathcal{L}(\Gamma_t) = \sum_{i=1}^n \log P_i(m_{r,i} | \Gamma_t), \quad (6.3)$$

where  $n$  is the number of events.

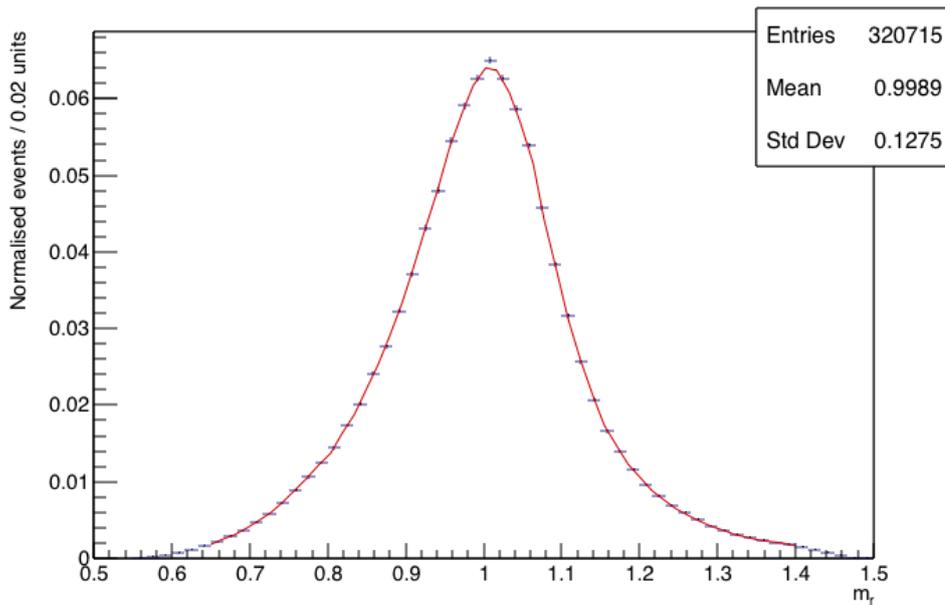


Figure 6.3: Distribution of the total probability for the reduced top quark mass considering  $\Gamma_t = 1 \times \Gamma_{t,\text{gen}}$ . The distribution is normalised on the interval 0.65 to 1.4.

A couple of optimisations to this procedure are detailed below. Their effect has already been included in the plots in Figures 6.2 and 6.3.

### Increase fraction of CM events

Only correctly reconstructed events are sensitive to the top quark decay width. Therefore, a higher weight of the CM template in the likelihood function, which corresponds to a higher fraction of CM events, will increase the precision of the measurement. It can be observed in Figure 6.2 that the distribution is much narrower for CM than for UM events. Considering only events with  $0.65 < m_r < 1.4$ , the fraction of CM events increases with about 0.5%. This is a relatively moderate effect due to the severe requirements on the reconstructed top quark mass in the event selection. When the range is restricted, the distributions are normalised such that the probability to have an  $m_r$  between 0.65 and 1.4 is equal to one.

The effect of the top quark width on the distribution of CM events is visualised in Figure 6.4. As was shown in Section 3.3.1, the size of  $\Gamma_t$  is directly related to the width of the distribution. Figure 6.4 shows that the sensitivity to  $\Gamma_t$  is contained in both the peak and the tails of the distribution.

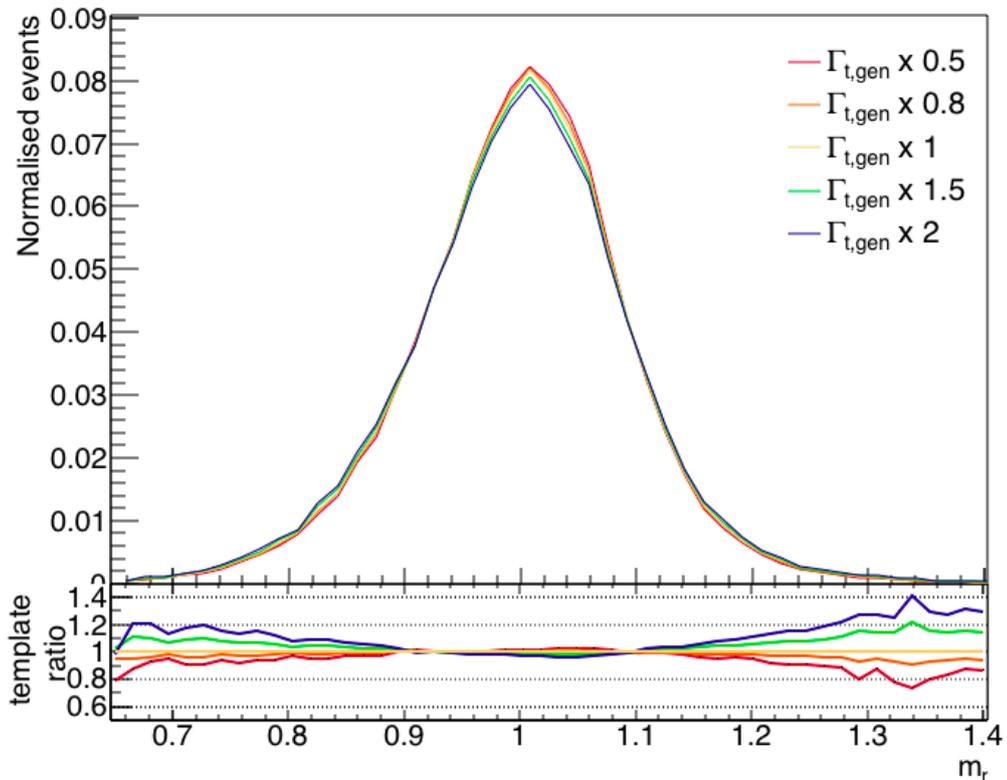


Figure 6.4: Differences between the templates for CM events considering several values of  $\Gamma_t = s \times \Gamma_{t,\text{gen}}$ .

### Smoothing of histograms

Since events in the UM category are not sensitive to the top quark width, variations in the templates for different widths are due to statistical fluctuations. In order to reduce these, the histograms are smoothed. This means that the bin contents of nearby bins are compared and, when large differences are observed in a non-consistent direction, the content is redistributed over the bins such that a smoother transition between bins is obtained. This is visualised in Figure 6.5.

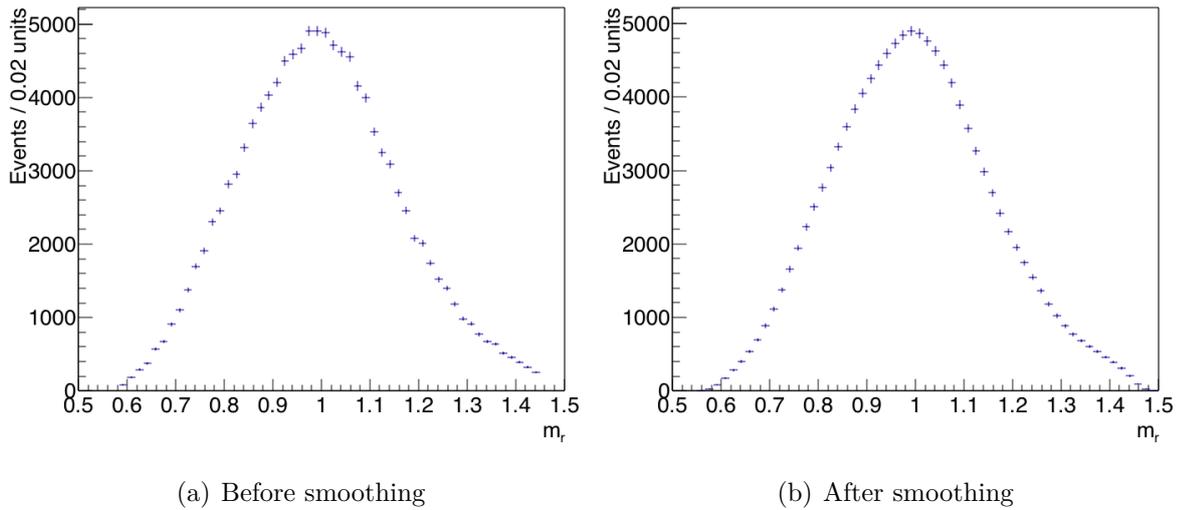


Figure 6.5: Effect of smoothing the reduced top quark mass distribution for UM events considering  $\Gamma_t = 1 \times \Gamma_{t,\text{gen}}$ .

When the templates for different top quark widths are superimposed, as in Figure 6.6, it is clear that the templates for all top quark widths are consistent within statistical fluctuations. As a result, only the templates for the nominal top quark width are used for the UM category so as to be sure to exclude deviations due to the finite number of events.

### Likelihood function

In this analysis the negative logarithm of the likelihood, denoted as  $\mathcal{L}$ , is used. Implementing the width independence of  $P_{\text{UM}}$  in Eq. (6.2),  $\mathcal{L}$  can be written as

$$\mathcal{L}(\Gamma_t) = -\ell(\Gamma_t) = -\sum_{i=1}^n \log \left[ f_{\text{CM}} P_{\text{CM}}(m_{r,i}|\Gamma_t) + f_{\text{UM}} P_{\text{UM}}(m_{r,i}) \right], \quad (6.4)$$

where the fractions are equal to 58.8% and 41.2% for CM and UM events, respectively. This function is plotted in Figure 6.7 for several values of  $\Gamma_t$  separately.

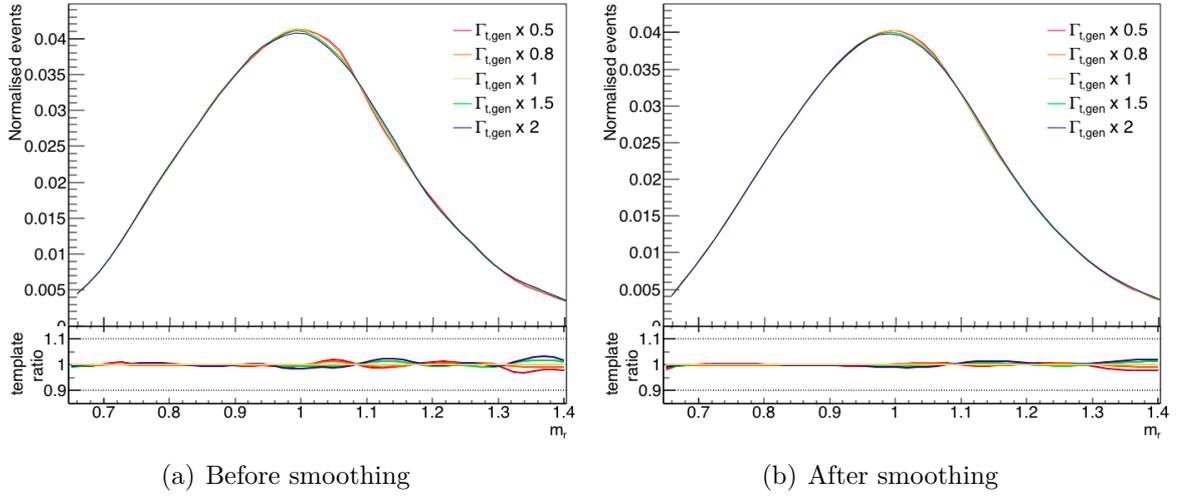


Figure 6.6: Differences between the templates for UM events considering several values of  $\Gamma_t = s \times \Gamma_{t,gen}$ .

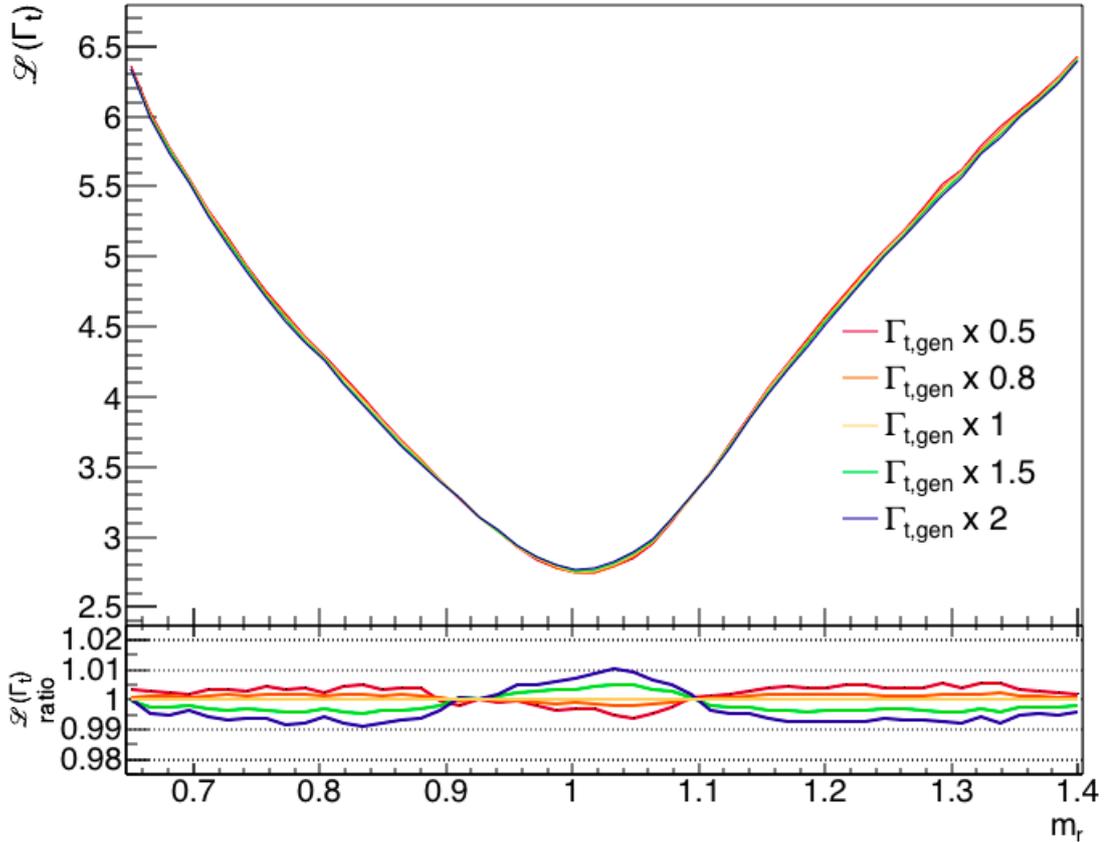


Figure 6.7: Distribution of the negative logarithm of the likelihood  $\mathcal{L}$  considering several top quark widths  $\Gamma_t$ .

### 6.1.2 Evaluation of the likelihood function

Given a set of events, the likelihood  $\mathcal{L}$  for a fixed value of the top quark width  $\Gamma_t$  is determined by adding the negative logarithms of the probability to have a certain reduced top quark mass  $m_{r,i}$  for each event  $i$ . This is done for all values of  $s$  defined above. As the sign of  $\mathcal{L}$  is inverted compared to the standard likelihood definition, the best estimator for the top quark width corresponds to the width value associated with the minimum of  $\mathcal{L}(\Gamma_t)$ . Considering only a limited amount of widths  $\Gamma_t = s \times \Gamma_{t,\text{gen}}$ , the likelihood values are plotted as a function of  $s$  and the minimum is fitted by a parabola. The function value of the top quark width estimator is subtracted from the likelihood distribution such that  $\Delta\mathcal{L}(\hat{\Gamma}_t) = 0$ . The points at one standard deviation from the estimator can be found where the parabola has a function value of  $1/2$ . The uncertainty on the estimator is then determined by taking the absolute value of the difference between one of these points and the estimator. As a parabola is completely symmetric around its minimum, either of the points will do. This is visualised in Figure 6.8 for simulated events, where  $\Delta\mathcal{L}(\Gamma_t) = 1/2$  is indicated by a horizontal grey line. The uncertainty traced by this line depends on the number of events contained in the simulated sample. If more events are considered, the differences between the likelihood values for each  $\Gamma_t$  will become more pronounced, which results in a narrower parabola. Similarly, the parabola will be broader for fewer events.

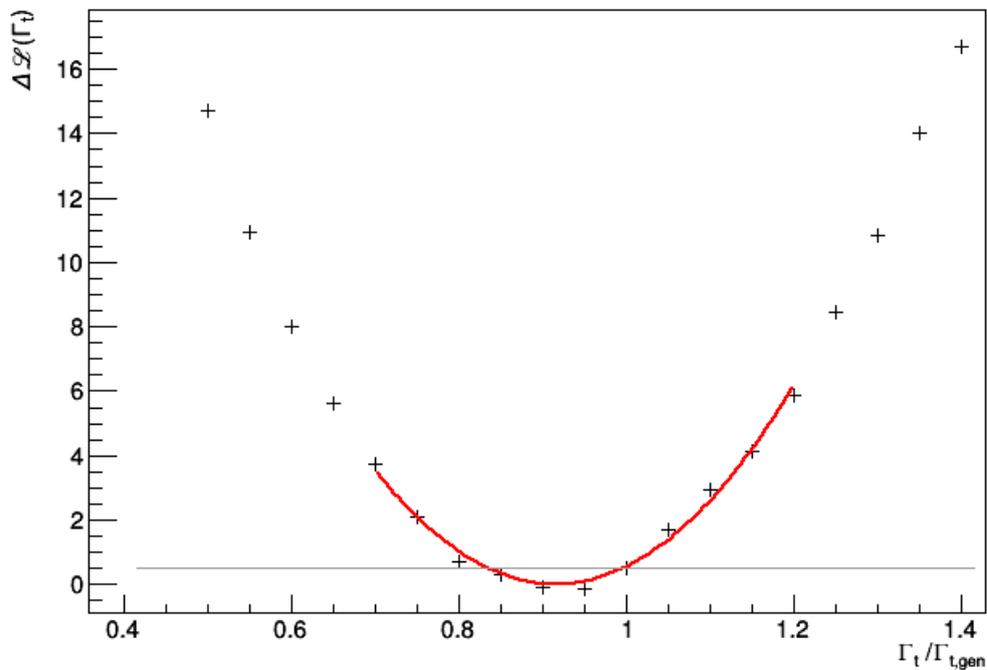


Figure 6.8: Distribution of the negative logarithm of the likelihood,  $\mathcal{L}$ , when the nominal simulated samples are used as input.

### 6.1.3 Calibration curve

The advantage of simulated samples is that the top quark decay width that was used to simulate the events is known. It can therefore be compared to the output of the measurement procedure. Figure 6.8 shows the output for a top quark width  $\Gamma_t = 1 \times \Gamma_{t,\text{gen}}$ . The minimum is found around  $(0.92 \pm 0.08) \times \Gamma_{t,\text{gen}}$ , so the estimator for the top quark width has a small bias. As both the input and the output value are known for this top quark width, a correction factor can be applied. In order to correct a random  $\Gamma_t = s \times \Gamma_{t,\text{gen}}$ , the likelihood  $\mathcal{L}$  is calculated for simulated events, where the top quark width of the  $t\bar{t}$  events is reweighted by a factor

$$s \in \{0.2, 0.4, 0.5, 0.6, 0.8, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5\}, \quad (6.5)$$

and the relative shift that is observed for each input width is evaluated. The output width is plotted as a function of the input width in Figure 6.9. Ignoring small values of  $s$ , all points can be fitted by

$$\Gamma_{t,\text{out}} = a + b \cdot \Gamma_{t,\text{in}}, \quad (6.6)$$

where  $a = -0.15 \pm 0.04$  and  $b = 1.09 \pm 0.02$ . Deviations from this linear behaviour for small values of  $s$  are due to the fact that the top quark width always has to be positive. Therefore, the likelihood function increases asymptotically near  $s = 0$ , which results in output values that are slightly larger than expected.

The calibration curve is validated using  $t\bar{t}$  samples that were generated with a different top quark width, see Table 3.2. The output values obtained using these samples are superimposed on the plot in Figure 6.9 using green markers. They correspond to the results using reweighting within one standard deviation. As the sample for  $s = 1$  is used as a starting point for the reweighting, the points overlap.

Including the bias correction, the estimator for the top quark decay width becomes

$$\hat{\Gamma}_t = \frac{\Gamma_{t,\text{out}}}{b} - \frac{a}{b}, \quad (6.7)$$

with an uncertainty

$$\sigma_{\hat{\Gamma}_t} = \frac{1}{b} \sqrt{\sigma_{\Gamma_{t,\text{out}}}^2 + \sigma_a^2 + \hat{\Gamma}_t^2 \sigma_b^2}, \quad (6.8)$$

where  $\sigma_a^2$  and  $\sigma_b^2$  are the uncertainties on the fit parameters of the calibration curve and  $\sigma_{\Gamma_{t,\text{out}}}$  is the uncertainty on the measured  $\Gamma_{t,\text{out}}$ , as defined in Section 6.1.2. The expected top quark decay width is then equal to  $(0.98 \pm 0.09) \times \Gamma_{t,\text{gen}}$ .

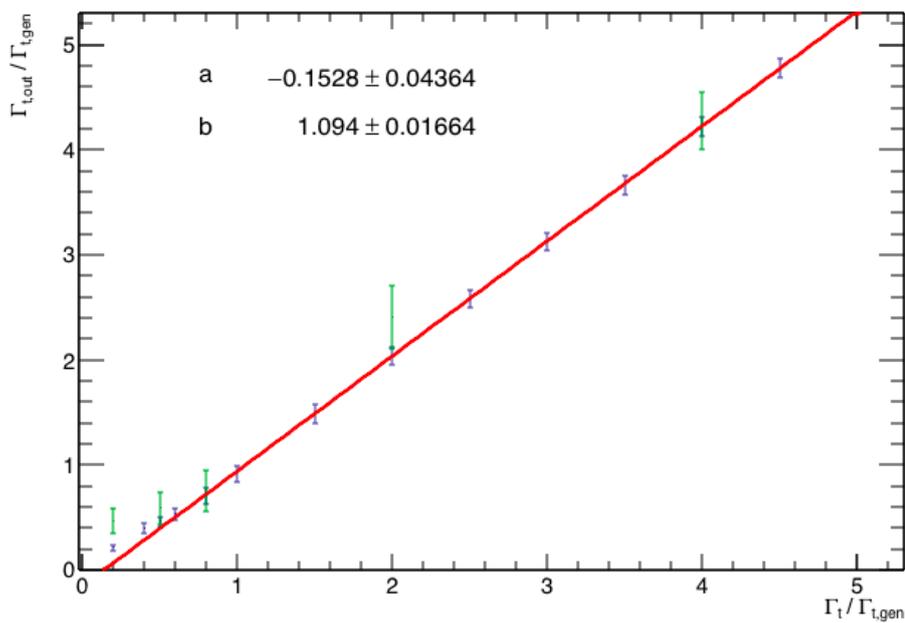


Figure 6.9: Calibration curve for the measurement procedure, plotting the output value of  $\Gamma_t / \Gamma_{t,\text{gen}}$  as a function of  $s$ . The blue markers represent top quark width estimates that were made when the  $t\bar{t}$  events are reweighted with a factor  $s$ , while the green markers correspond to the measurements on  $t\bar{t}$  events that were generated with a different top quark width.

### 6.1.4 Pull

In order to estimate if  $\sigma_{\hat{\Gamma}_t}$  as defined in Eq. (6.8) is a good estimate of the statistical uncertainty, pseudo experiments are generated and the pull is calculated. In general, the simulated samples contain many more events than the data sample, so the statistical uncertainty of the simulation is smaller than that of the data. Therefore, pseudo experiments are used, which contain only a subset of the simulated events such that their integrated luminosity is comparable to that of the data. The simulated events are selected randomly from the collection of events that pass all selection requirements. If the simulated samples are large, many pseudo experiments can be created without a significant overlap in the events that they contain.

For each pseudo experiment the likelihood  $\mathcal{L}$  is calculated and the top quark decay width and its uncertainty are estimated. The results for 400 pseudo experiments are plotted in Figure 6.10.

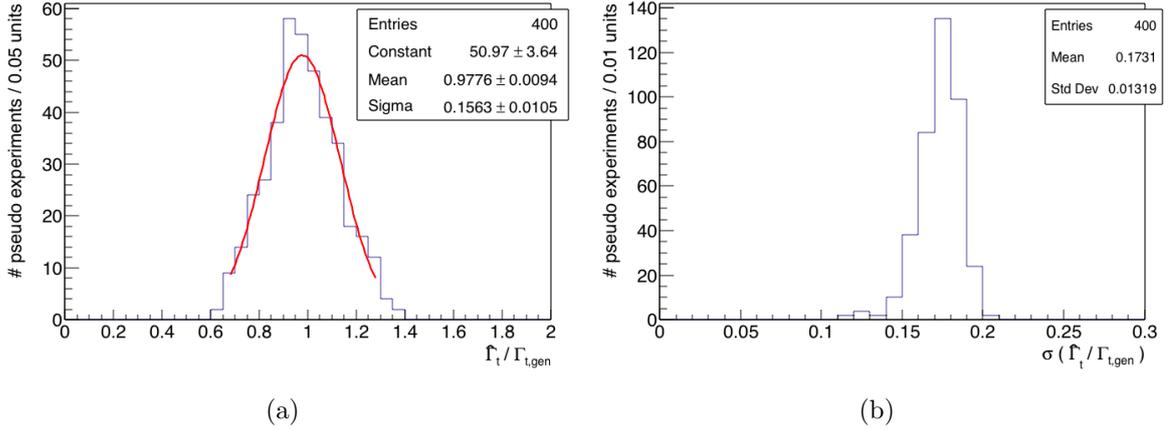


Figure 6.10: Distribution of the top quark decay width (a) and its uncertainty (b) for 400 pseudo experiments.

The average top quark width is equal to  $(0.98 \pm 0.01) \times \Gamma_{t,gen}$  and the pull is calculated by considering

$$\frac{\hat{\Gamma}_{t,i} - \langle \hat{\Gamma}_t \rangle}{\sigma_{\hat{\Gamma}_{t,i}}}, \quad (6.9)$$

for each pseudo experiment  $i$ , which is visualised in Figure 6.11. If  $\sigma_{\hat{\Gamma}_t}$  describes the statistical uncertainty well, it is expected that the pull should follow a Gaussian distribution centred around zero with a standard deviation of one. In this case, the standard deviation is equal to  $0.87 \pm 0.04$ , which implies that  $\sigma_{\hat{\Gamma}_t}$  overestimates the statistical uncertainty of the measurement and can thus be considered as a conservative estimate.

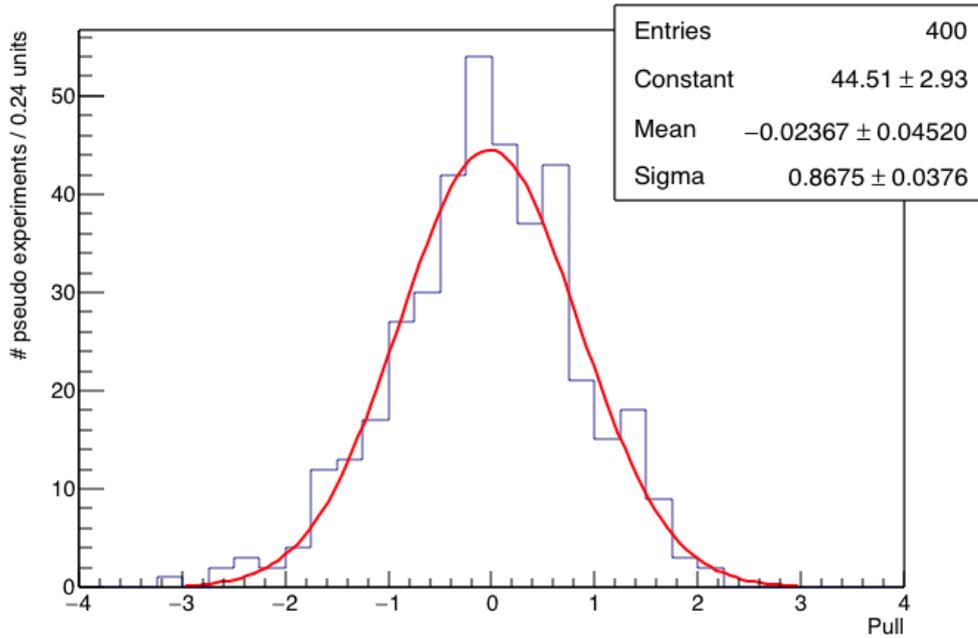


Figure 6.11: Distribution of the pull for 400 pseudo experiments, fitted by a Gaussian function.

## 6.2 Combination of Several Sensitive Variables

This section will first describe how probability density templates are created for another variable. Then, the likelihood functions for both variables are combined.

### 6.2.1 Template construction for another variable

The procedure outlined in Section 6.1 can be repeated using a variable related to the leptonically decaying top quark to make probability density templates. Therefore, the event categorisation is repeated for the reconstruction of top quarks that decay leptonically. In this case, events are correctly matched if the b jet candidate that is not assigned to the hadronic top quark can be matched to the b quark originating from the leptonic decay of the top quark and if the angular distance between the reconstructed and the generated muon is smaller than  $\Delta R < 0.1$ . With these requirements, only 11% of the generated events cannot be matched. Using the same event selection as outlined in Chapter 5, 93% of the matched events is correctly matched, so 83% of all events can be found in the CM category.

The reduced  $m_{\text{lb}}$  mass is defined as

$$m_{\text{lb},r} = \frac{m_{\text{lb}}}{\langle m_{\text{lb}} \rangle_{CM}}, \quad (6.10)$$

where  $\langle m_{\text{lb}} \rangle_{CM} = 97.3 \text{ GeV}$  is the average  $m_{\text{lb}}$  mass that is calculated using only events where the leptonically decaying top quark is correctly matched. The distribution of the

reduced  $m_{lb}$  mass is presented in Figure 6.13. Once again the shape of the distribution of the data is dominated by that of the CM events.

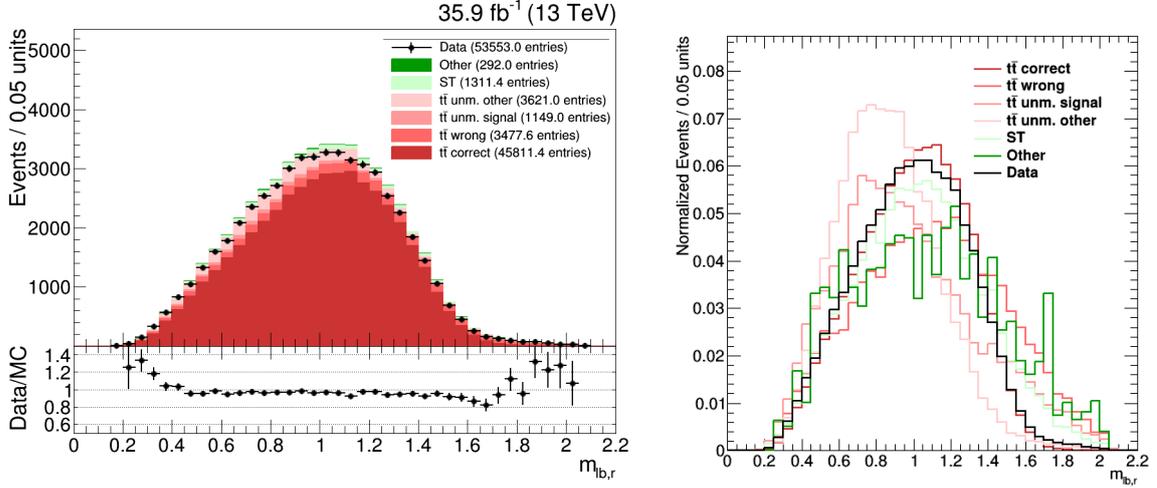


Figure 6.12: Distribution of the reduced  $m_{lb}$  mass (left) and the normalised distribution for each data set separately (right). The event categorisation is based on the matching of the leptonically decaying top quark.

As the distributions of  $m_{lb,r}$  have a different shape than those of  $m_r$ , other parameters are used to construct the initial histograms, namely the range  $[0, 2.5]$  is spanned by 80 bins. The histograms are normalised on an interval from 0.35 to 1.95 and template functions are constructed taking account of the considerations described in Section 6.1.1. The template functions for a width of  $\Gamma_t = 1 \times \Gamma_{t,\text{gen}}$  are visualised in Figure 6.13. These

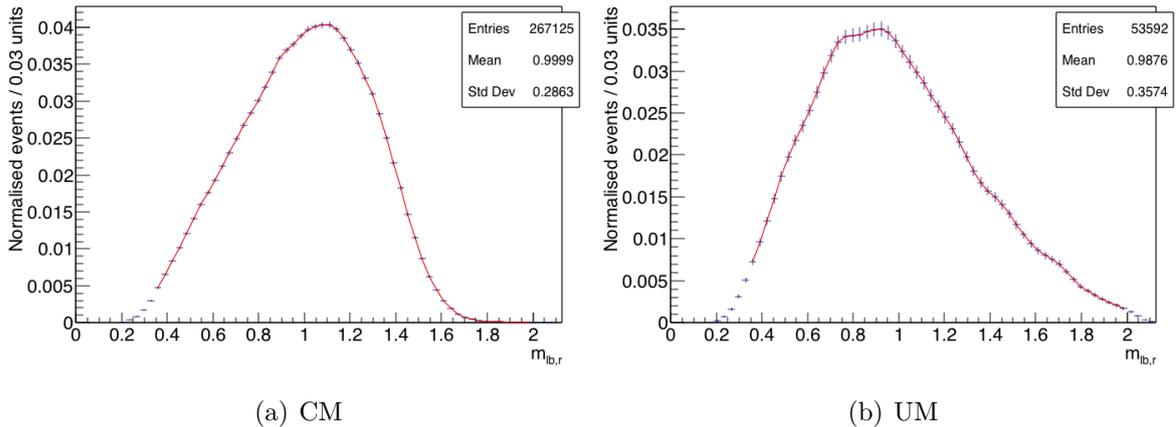


Figure 6.13: Distribution of the reduced  $m_{lb}$  mass considering  $\Gamma_t = 1 \times \Gamma_{t,\text{gen}}$  for (a) correctly matched and (b) unmatched events. The red curves represent the templates, while the blue markers indicate the bin contents of the histograms and their respective uncertainties. The distributions are normalised on the interval 0.35 to 1.95.

are combined into a likelihood function  $\mathcal{L}$  as in Equation (6.4), where the fractions are

equal to 83.5% and 16.5% for CM and UM events, respectively. Figure 6.14 shows the likelihood for several values of  $\Gamma_t$ .

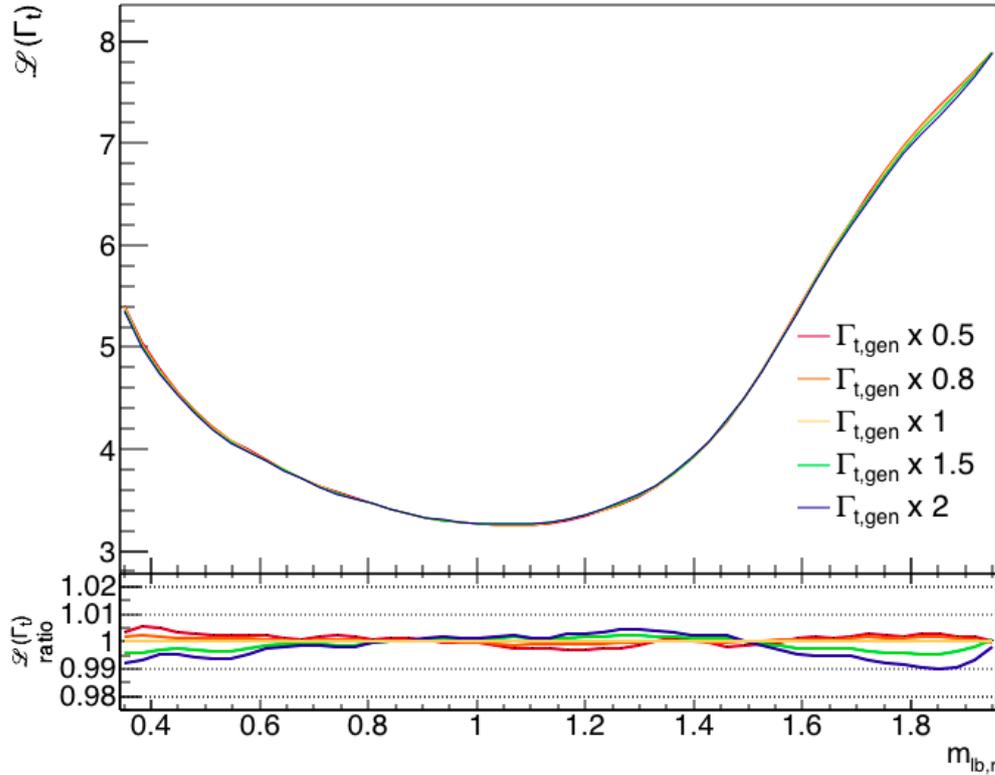


Figure 6.14: Distribution of the negative logarithm of the likelihood  $\mathcal{L}$ , where the sensitive variable is the reduced  $m_{lb}$  mass, considering several top quark widths  $\Gamma_t$ .

Compared to Figure 6.7, the likelihood functions in Figure 6.14 are much closer together. This indicates that  $m_{lb,r}$  is less sensitive to the top quark decay width than  $m_r$ . Therefore, only the results of the combination of these two variables are discussed in detail below.

## 6.2.2 Combination of both variables

As mentioned in Section 6.1.2, the shape of the  $\Delta\mathcal{L}(\Gamma_t)$  distribution depends on the number of events, which is related to the number of terms in the likelihood function. The sensitivity of the measurement procedure can thus be increased by multiplying the likelihoods of several variables. As this is equivalent to adding their logarithms, more terms are created without requiring more events. In reality, fewer events will be involved in the likelihood calculation, because the events have to abide by the requirements on both variables. In this case, events need to satisfy  $0.65 < m_r < 1.4$  and  $0.35 < m_{lb,r} < 1.95$ . The additional requirement on  $m_{lb,r}$  excludes less than 1% of the events involved in the purely hadronic measurement and an overall gain in the sensitivity is achieved.

Combining the sensitive variables  $m_r$  and  $m_{lb,r}$ , a top quark decay width of  $(0.79 \pm 0.07) \times \Gamma_{t,\text{gen}}$  is obtained for an input width of  $\Gamma_t = 1 \times \Gamma_{t,\text{gen}}$ . This is visualised in Figure 6.15. As expected, the likelihood distribution is narrower for this so-called combined measurement than for the hadronic-only measurement.

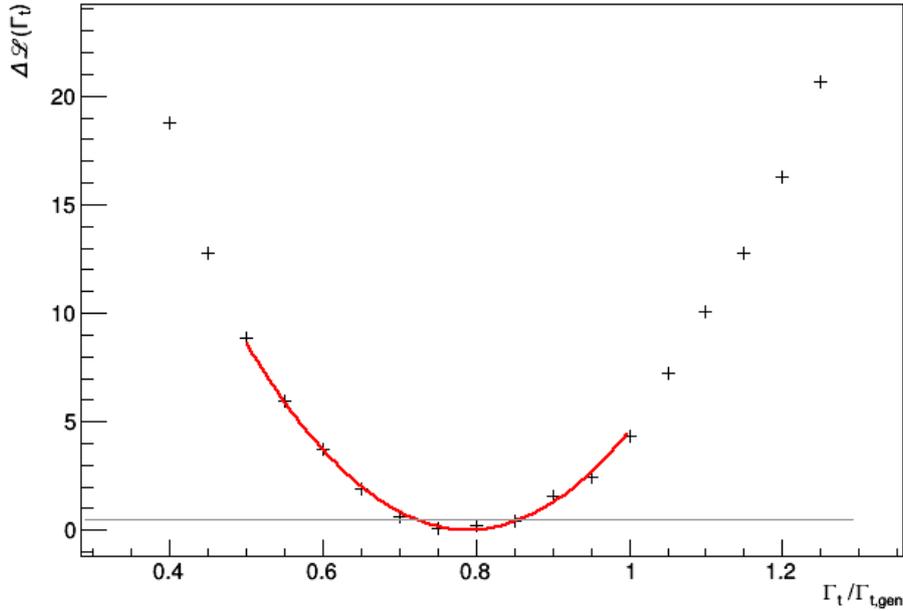


Figure 6.15: Distribution of the negative logarithm of the likelihood,  $\mathcal{L}$ , for the combined measurement, when the nominal simulated samples are used as input.

### 6.2.2.1 Calibration curve

Also for the combined measurement a shift of the top quark width value is observed. A dedicated calibration curve is determined following the criteria of Section 6.1.3. In this case the parameters of the curve  $\Gamma_{t,\text{out}} = a + b \cdot \Gamma_{t,\text{in}}$  are equal to  $a = -0.25 \pm 0.04$  and  $b = 1.08 \pm 0.01$ . When the calibration curve, which is shown in Figure 6.16, is applied to the result above, the top quark decay width obtained for the nominal simulated samples is equal to  $(0.96 \pm 0.07) \times \Gamma_{t,\text{gen}}$ .

### 6.2.2.2 Pull

For 400 pseudo experiments, the average top quark width is equal to  $(0.96 \pm 0.01) \times \Gamma_{t,\text{gen}}$ . The pull has a value of  $0.92 \pm 0.05$ , which implies that the statistical uncertainty is slightly overestimated.

The distributions of the top quark width and its uncertainty after applying the calibration curve are visualised in Figure 6.17 and the pull is shown in Figure 6.18.

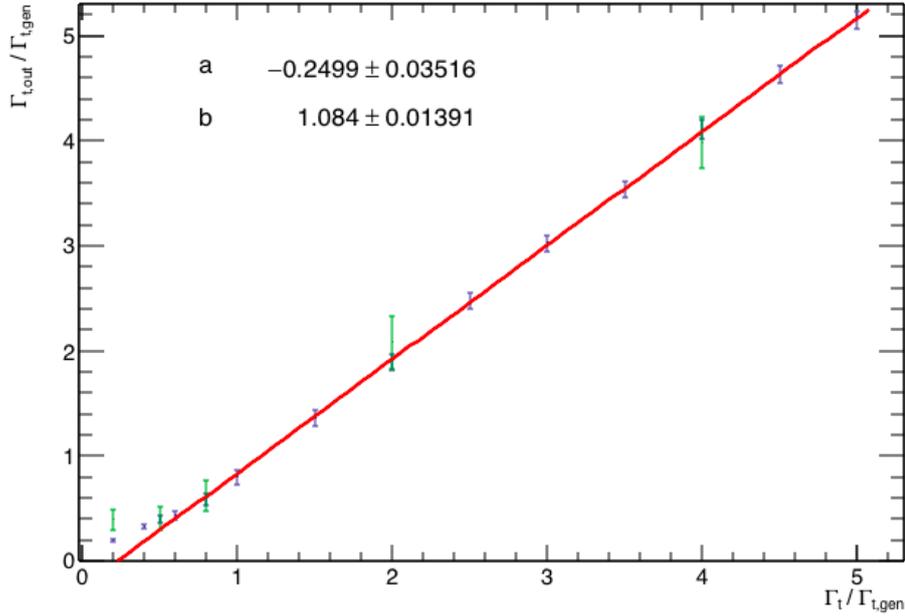


Figure 6.16: Calibration curve for the measurement procedure using both  $m_r$  and  $m_{lb,r}$  as sensitive variables. The output value of  $\Gamma_t/\Gamma_{t,\text{gen}}$  is plotted as a function of  $s$ . The blue markers represent top quark width estimates that were made when the  $\bar{t}\bar{t}$  events are reweighted with a factor  $s$ , while the green markers correspond to the measurements on  $\bar{t}\bar{t}$  events that were generated with a different top quark width.

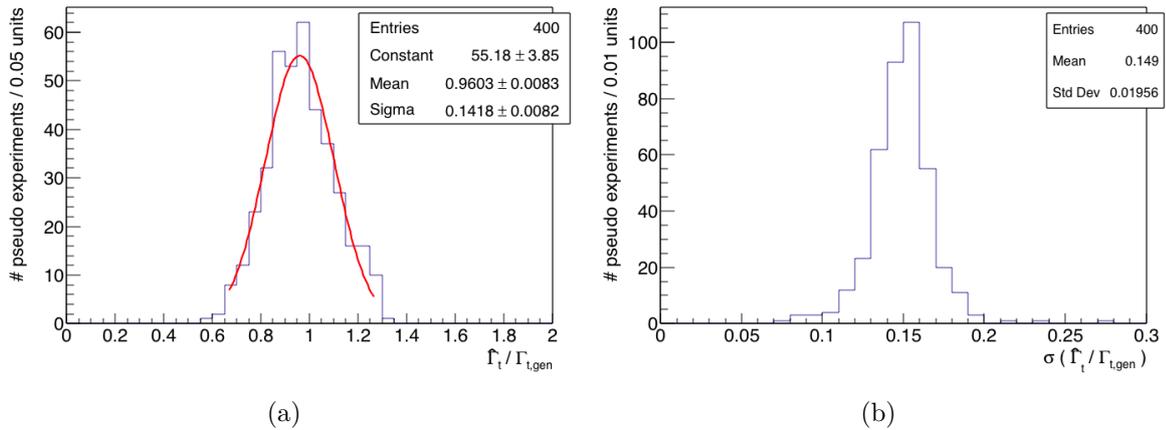


Figure 6.17: Distribution of the top quark decay width (a) and its uncertainty (b) for 400 pseudo experiments using both  $m_r$  and  $m_{lb,r}$  as sensitive variables.

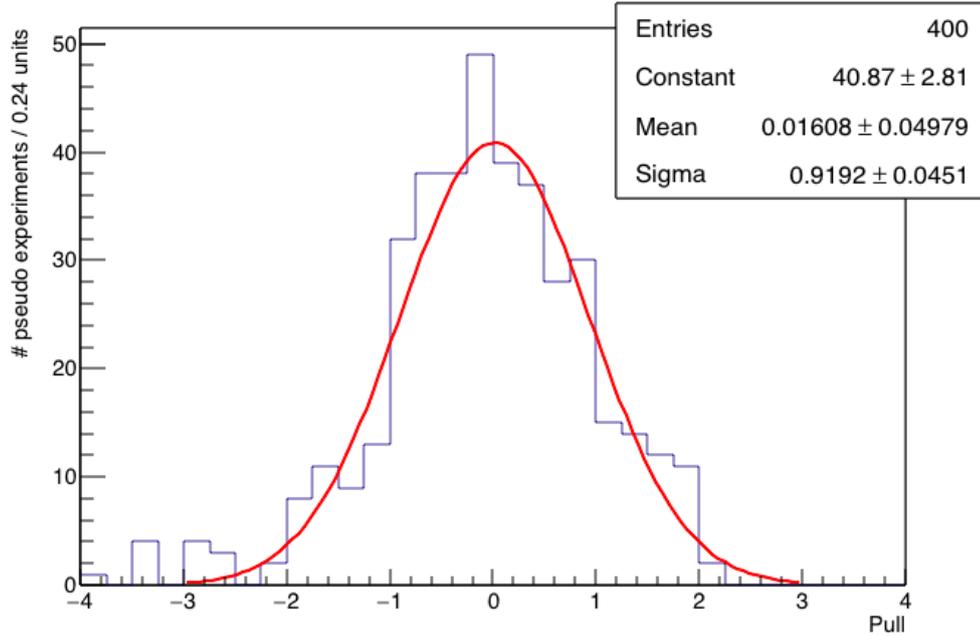


Figure 6.18: Distribution of the pull for 400 pseudo experiments when both  $m_r$  and  $m_{lb,r}$  are used as sensitive variables. The distribution is fitted by a Gaussian function.

## 6.3 Systematic Uncertainties

Apart from the number of events, the top quark width measurement is also influenced by the uncertainties on the model parameters of the simulation and the limited resolution of the detector. These uncertainties are described in more detail in Sections 6.3.1 and 6.3.2. Their effect on the top quark width measurement is estimated by repeatedly running the measurement procedure, each time implementing one change with respect to the nominal conditions. The result is compared to those in Sections 6.1 and 6.2, and the difference is taken as the systematic uncertainty due to this parameter. The results are summarised in Section 6.3.3 and the total systematic uncertainty, taking account of correlations between the systematic effects, is calculated.

### 6.3.1 Theoretical uncertainties

The events used in this thesis have been simulated according to the models described in Section 3.1. Many sources of modelling uncertainties have already been introduced there. This section will focus on the parameter variations that are applied to estimate the effect of mismodelling.

#### Matrix element and parton shower scale uncertainties

The factorisation scale  $\mu_F$  is varied within its order of magnitude, i.e. from  $\mu_F/2$  to  $2\mu_F$ . This is done using additional event weights provided in the simulated samples. Also the renormalisation scale is varied in this range and both scales can be varied up and down simultaneously as well. As the sum of the weighted events

is not necessarily equal to the unweighted number of events, a correction factor is introduced and the event weight is applied as

$$\text{weight}_i = w_i \cdot \frac{\sum_i w_{\text{orig},i}}{\sum_i w_i}, \quad (6.11)$$

where  $w_{\text{orig},i}$  is the event weight for the nominal  $\mu_F = \mu_R = M_{t,\text{gen}}$  and  $w_i$  is the event weight for alternative scenarios.

Variations in the  $Q^2$  scale also influence how much ISR and FSR is produced. This is investigated using extra samples where the strong coupling constant  $\alpha_s$  is evaluated in  $(2Q)^2$  and  $(Q/2)^2$ , respectively. The result of the FSR variation is scaled by a factor  $2^{-1/2}$  in order to better reflect the shower uncertainties [168].

The envelope of these variations, i.e. the largest discrepancy in each direction of the nominal result, is considered to be the systematic uncertainty due to fluctuations in the modelling of the PS.

### PDF uncertainties

Events are simulated using the NNPDF3.0 parton distribution functions at NLO, where the strong coupling constant is considered to be  $\alpha_s = 0.118$ . The effect of the PDF modelling uncertainties is estimated by creating alternative PDF sets, where one parameter is varied within its uncertainty. These variations are taken into account using event weights, which are applied as in Eq. (6.11). In this case, 50 parameters are varied up- and downwards, so 100 variations are considered. The uncertainty of the PDFs is equal to the root-mean-square (RMS) of these variations, i.e.

$$\text{RMS} = \sqrt{\frac{\sum_{i=1}^n (\text{variation}_i - \text{central})^2}{n - 1}}, \quad (6.12)$$

where  $n = 100$ .

In addition, also variations of  $\alpha_s$  are considered, i.e.  $\alpha_s = 0.117$  and  $0.119$ . The total systematic uncertainty due to PDF modelling is equal to the quadratic sum of the result in Eq. (6.12) and the  $\alpha_s$  variations.

### ME-PS matching

In the CUETP8M2T4 tune the `hdamp` parameter has a nominal value of  $1.58 \times M_{t,\text{gen}}$ . This value is varied up and down by one standard deviation, corresponding to `hdamp` =  $2.24 \times M_{t,\text{gen}}$  and `hdamp` =  $0.99 \times M_{t,\text{gen}}$ , respectively [132].

### b jet modelling

PYTHIA 8 uses the Bowler–Lund model [169] to describe the fragmentation into b hadrons. The fragmentation function is varied within its uncertainties [170, 171], which alters the ratio of the b hadron  $p_T$  to the jet  $p_T$ . This effect is investigated using event weights. Additionally, a comparison with the Peterson model [172] is made.

Further, the semileptonic decay of b hadrons,  $B \rightarrow lX$ , is investigated. The branching fraction is varied by  $-0.45\%$  and  $+0.77\%$ , which corresponds to the envelope of b hadron decay measurements and their uncertainties [8]. As for fragmentation, event weights are used to study the impact.

### Underlying event

The modelling of the UE is tested by varying a number of parameters in the tune. These influence the amount of MPIs and CR that occur in the UE.

### Colour reconnection

Besides the nominal  $t\bar{t}$  sample, using the default colour reconnection scheme, additional samples were created considering different approaches of colour reconnection, i.e. the default scheme with early resonance decays, the QCD-based model and the gluon-move model, all of which are described in Section 3.1.

### Top quark mass uncertainty

Due to the relation between the top quark mass and decay width, a variation of the top quark mass value is expected to directly influence the top quark width measurement. Therefore, the top quark mass of the nominal  $t\bar{t}$  sample is reweighted using the method described in Section 3.3.2, such that a mass variation of 1 GeV is simulated.

### Top quark $p_T$ uncertainty

The transverse momentum of the top quark is not well-modelled for generators with NLO accuracy [173], which results in a softer top quark  $p_T$  spectrum in data compared to simulation. A comparison between data and simulation results in the calculation of event weights, which can be used to estimate the effect of the mis-modelling. Both for the generated top and antitop quark, see definitions in Section 5.3.1, a scale factor is calculated as

$$\text{SF}(p_T) = e^{0.0615 - 0.0005 \cdot p_T}, \quad (6.13)$$

and the event weight is equal to  $w = \sqrt{\text{SF}(p_{T,t})\text{SF}(p_{T,\bar{t}})}$  [174].

It is advised not to use the event weights as a correction to the simulation, as this degrades the description of other event variables, such as the transverse mass of the  $t\bar{t}$  system,  $m_{t\bar{t}}$ . Instead, the analysis is performed with and without applying the event weights described above and the difference is taken as the systematic uncertainty.

## 6.3.2 Experimental uncertainties

All measurements have a limited precision. Therefore, the uncertainties on the applied scale factors and corrections, as well as quantities such as the luminosity, are investigated.

### LHC beam energy

Protons collide in the CMS experiment when both beams carry an energy of 6.5 TeV. This value has a relative uncertainty of about 0.1%, which is negligible on an analysis level [175].

### Luminosity uncertainty

Changes in the luminosity affect the amount of expected events and thus the relative weight of the simulated processes. The uncertainty on the luminosity

measurement for the 2016 run is 2.5% [176]. It has been checked that this is negligible for the measurement presented in this analysis.

### Jet energy corrections

The jet energy scale and resolution, as described in Section 4.3.3, are varied within their uncertainties. These corrections are applied before the event selection and only one variation is considered at a time. The uncertainties on the jet energy scale are generally smaller than 1% and those on the resolution are of the order of 1 – 5%.

### Pileup

During the 2016 run, the cross section for protons to collide in the LHC is 69.2 mb with an uncertainty of 4.6% [177, 178]. The deviation in the number of pileup interactions is estimated by varying the cross section within its uncertainty and reweighting the events accordingly.

### Lepton scale factors

The scale factors that are applied in the analysis have been determined with a certain uncertainty. The effect of these uncertainties is estimated by adding or subtracting one standard deviation from the nominal value of the scale factor. These are then used to reweight the events. The uncertainties are dependent on  $p_T$  and  $|\eta|$  and are usually between 1 – 3%.

### b tag scale factors

Also the uncertainties on the b jet identification and tagging efficiencies are propagated using event weights. The  $p_T$ - and  $|\eta|$ -dependent uncertainties predominantly have a value around 2%.

### Cross sections of simulated samples

In order to account for the uncertainties on the theoretical cross sections of the simulated processes, these are varied by 10% for single top quark production and by 30% for Drell–Yan and W boson production in association with jets. The cross sections of the ST t and tW channels are varied separately, while the cross sections for the DY + jets and W + jets samples are varied simultaneously.

### Precision of calibration curve parameters

The parameters of the calibration curve are determined with a certain uncertainty. Four alternative calibration curves are created by adding or subtracting one standard deviation to/from the nominal value of the parameters.

### 6.3.3 Summary of systematic uncertainties

The uncertainties described above are studied for both the hadronic and the combined measurement and are summarised in Tables 6.1 and 6.2, respectively. In order to get a better overview, uncertainties within the same categories have been added quadratically and, where appropriate, only the envelope of the uncertainties is shown. Detailed results for each component of the systematic uncertainties can be found in Appendix A. The nominal result for  $\hat{\Gamma}_{t,\text{nom}} = s \times \Gamma_{t,\text{gen}}$  is presented at the top of the tables. Each value represents the shift  $\hat{\Gamma}_{t,\text{sys}} - \hat{\Gamma}_{t,\text{nom}}$  induced by the systematic effect. The bottom line presents the total systematic uncertainty, obtained by quadratically adding the individual items. All values are expressed in GeV.

When the systematic effect is calculated using a different simulated sample than the nominal one, the statistical uncertainty on the shift is calculated as

$$\sigma = \sqrt{\sigma_{\hat{\Gamma}_{t,\text{sys}}}^2 + \sigma_{\hat{\Gamma}_{t,\text{nom}}}^2}. \quad (6.14)$$

If the statistical uncertainty on the shift is larger than the shift itself, which is the case for e.g. the underlying event, the size of the former is taken as the uncertainty. In case the systematic uncertainty is propagated on an event-weight basis, its statistical uncertainty is considered to be fully correlated with that of the nominal measurement.

When comparing the systematic uncertainties on the hadronic and the combined measurements, it is clear that the hadronic measurement, which uses a variable containing three jets, is much more affected by jet-related uncertainties. Some of these are reduced by adding the  $m_{\text{lb}}$  terms to the likelihood, especially the effect on the parton shower uncertainty is striking. In contrast, the combined measurement has higher lepton-related uncertainties, but these are negligible compared to the jet-related ones. Some of the uncertainties, such as the JES uncertainty, even get enhanced in the combined measurement, but overall the systematic effects are smaller for the combined measurement compared to the hadronic measurement.

In both cases theoretical modelling uncertainties are the dominant factor, especially those related to the PS scales, ME-PS matching and colour reconnection. A better modelling of these effects can thus further constrain the total systematic uncertainty and significantly improve the result.

Table 6.1: Summary of the systematic uncertainties for the hadronic measurement. The values represent an absolute shift compared to the nominal value presented at the top of the table and are expressed in GeV.

Nominal value	$1.279 \pm 0.112$			
Systematic	$+1\sigma$		$-1\sigma$	
	Shift	Unc	Shift	Unc
Calibration curve	0.056		-0.056	
JES	0.196		-0.027	
JER	0.167		-0.177	
Pileup SFs	0.113		-0.107	
b tagging SFs	0.055		-0.057	
Lepton SFs	0.003		-0.003	
Cross section variations	0.060		-0.060	
ME & PS scales	0.446	0.216	-0.404	0.162
ME-PS matching	0.273	0.209	-0.309	0.186
PDFs	0.062		-0.062	
b jet modelling	0.147		-0.083	
Underlying event	0.061	0.198	-0.066	0.197
Colour reconnection	0.447	0.195		
Top mass	0.319			
Top $p_T$ reweighting	0.155			
Total	0.867		-0.602	

Table 6.2: Summary of the systematic uncertainties for the combined measurement. The values represent an absolute shift compared to the nominal value presented at the top of the table and are expressed in GeV.

Nominal value	$1.258 \pm 0.096$			
Systematic	$+1\sigma$		$-1\sigma$	
	Shift	Unc	Shift	Unc
Calibration curve	0.046		-0.045	
JES	0.221		-0.165	
JER	0.131		-0.151	
Pileup SFs	0.084		-0.084	
b tagging SFs	0.041		-0.047	
Lepton SFs	0.013		-0.013	
Cross section variations	0.049		-0.054	
ME & PS scales	0.216	0.190	-0.270	0.140
ME-PS matching	0.322	0.181	-0.276	0.144
PDFs	0.073		-0.073	
b jet modelling	0.070		-0.105	
Underlying event			-0.157	0.164
Colour reconnection	0.307	0.189	-0.011	0.189
Top mass	0.269			
Top $p_T$ reweighting	0.129			
Total	0.650		-0.541	

## 6.4 Measurement of the Top Quark Decay Width

The measurement procedure described above has been set up and calibrated using simulated events only. Up to this point the analysis has been blinded in order not to let the strategy be influenced by the result. When the data are subjected to the same measurement routine, a top quark width of  $(0.73 \pm 0.14) \times \Gamma_{t,\text{gen}}$  is obtained for the hadronic measurement and  $(0.85 \pm 0.13) \times \Gamma_{t,\text{gen}}$  for the combined measurement. The distributions of the likelihood corresponding to these measurements are visualised in Figure 6.19. Note that the values quoted above have already been corrected by the calibration curve.

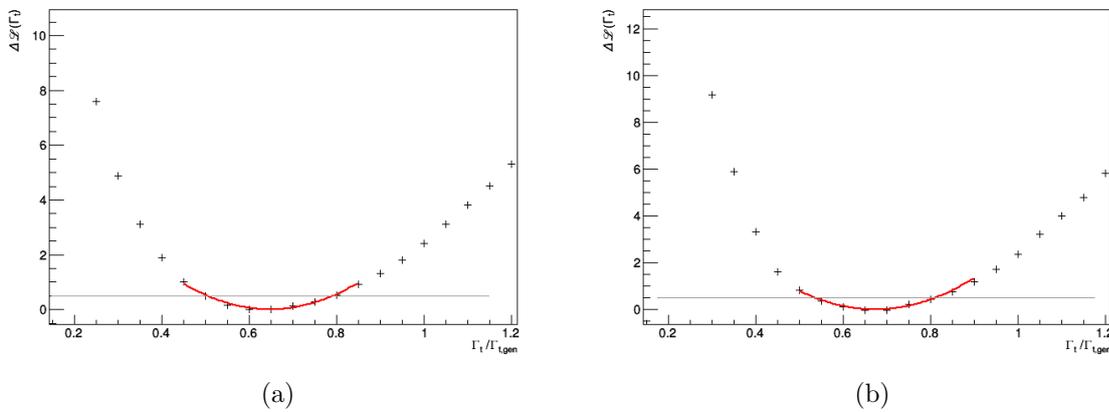


Figure 6.19: Distribution of the negative logarithm of the likelihood,  $\mathcal{L}$ , for the (a) hadronic and the (b) combined measurement using data.

Multiplying these results by the generated top quark width  $\Gamma_{t,\text{gen}} = 1.31$  GeV and taking account of the systematic uncertainties, the hadronic measurement corresponds to

$$\Gamma_t = 0.96 \pm 0.18 \text{ (stat.) } {}^{+0.87}_{-0.60} \text{ (sys.)} = 0.96 {}^{+0.89}_{-0.63} \text{ (total) GeV,} \quad (6.15)$$

and the combined measurement to

$$\Gamma_t = 1.12 \pm 0.17 \text{ (stat.) } {}^{+0.65}_{-0.54} \text{ (sys.)} = 1.12 {}^{+0.67}_{-0.57} \text{ (total) GeV.} \quad (6.16)$$

If a correction is implemented to account for the pull not being unity, the statistical uncertainties reduce to 0.16 GeV in both cases.

The results presented in Eqs. (6.15) and (6.16) are in agreement with the SM top quark width of 1.33 GeV within one standard deviation.



# 7

## Conclusions & Prospects

---

This thesis presents two measurements of the top quark decay width using top quark pair events in the semileptonic decay channel. Only events with exactly one muon and exactly four jets are considered. The top quark pair system is reconstructed and the reconstruction efficiency is estimated by comparing the reconstruction to the generator information for simulated events. These are then categorised according to whether they are correctly reconstructed or not. A kinematic fit is applied to increase the precision of the invariant top quark mass and further requirements are imposed on the selected events in order to increase the fraction of correctly reconstructed events. This way, the fraction of selected events that have a well-reconstructed hadronic top quark increases from about 26% to 59% and about 83% have a well-reconstructed leptonic top quark. Further, around 97% of the selected events are top quark pair events.

Likelihood templates are constructed using the variables  $m_r$  and  $m_{lb,r}$ , whose distributions are sensitive to the top quark decay width. This is done separately for the different event categories and different top quark widths are simulated by reweighting the  $t\bar{t}$  sample. For the measurement using  $m_r$ , a result of

$$\Gamma_t = 0.96 \pm 0.16 \text{ (stat.) } {}^{+0.87}_{-0.60} \text{ (sys.)} = 0.96 {}^{+0.88}_{-0.62} \text{ (total) GeV,} \quad (7.1)$$

is obtained. When this variable is complemented by  $m_{lb,r}$ , the result is improved to

$$\Gamma_t = 1.12 \pm 0.16 \text{ (stat.) } {}^{+0.65}_{-0.54} \text{ (sys.)} = 1.12 {}^{+0.67}_{-0.56} \text{ (total) GeV.} \quad (7.2)$$

These measurements are in agreement with the SM top quark decay width of 1.33 GeV. Comparing these results to the previous top quark width measurements, which are visualised in Figure 7.1, the combined measurement is more precise than the ATLAS result in the lepton+jets channel, i.e.  $\Gamma_t = 1.76 \pm 0.33 {}^{+0.79}_{-0.68}$  GeV. Further, the results presented in this thesis are compatible with the CMS measurement in the dilepton channel, which determined an interval  $0.60 < \Gamma_t < 2.50$  GeV at 95% CL.

The measurement procedure described in this thesis can easily be expanded to two dimensions. Similar to the variation of the top quark decay width, variations in another variable can be simulated. Considering the intimate connection between the top quark mass and decay width, a simultaneous measurement of both variables is a logical next step. Further, a two-dimensional measurement would give a clear view of how much the

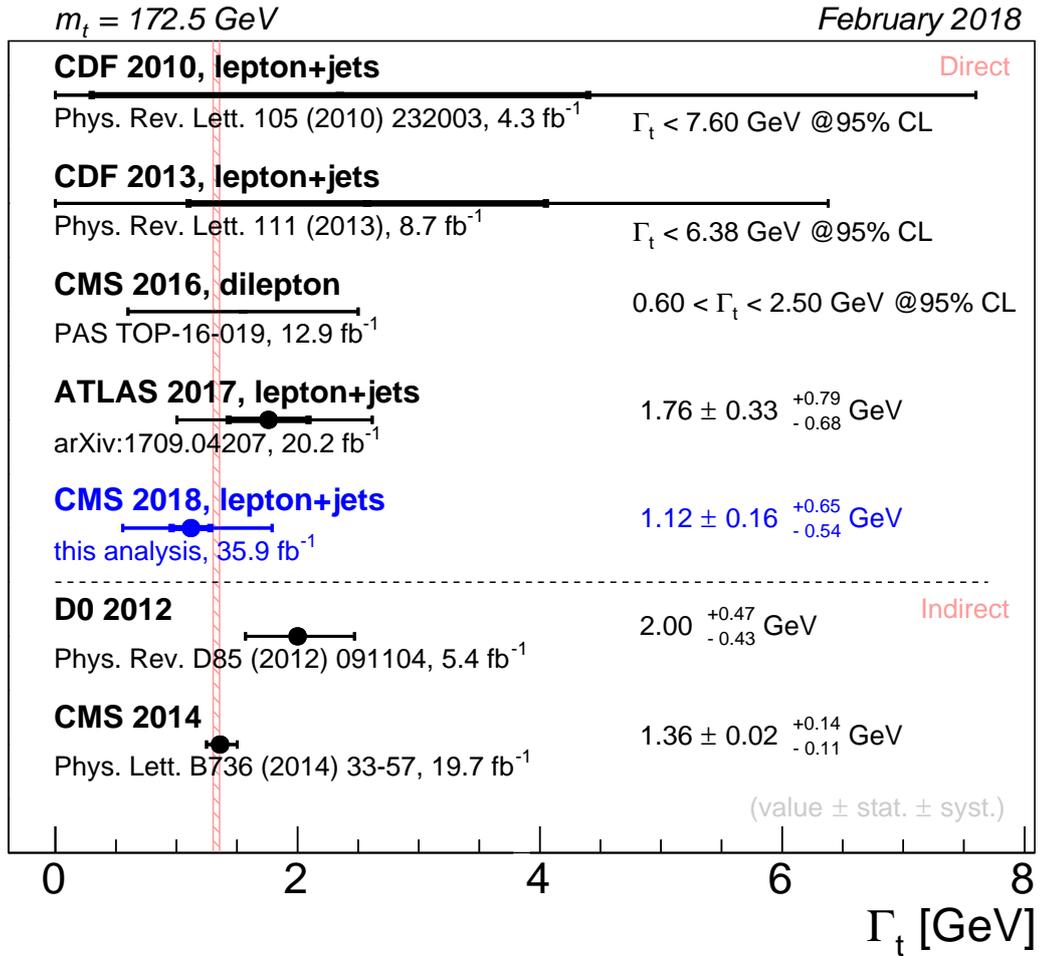


Figure 7.1: Summary of the top quark decay width measurements, including the most precise measurement presented in this thesis, performed by the ATLAS, CDF, CMS, and D0 experiments, all considering a top quark mass of 172.5 GeV. Direct measurements are grouped above the dashed line and indirect measurements below. The reference interval (red) indicates the theoretical prediction for the top quark decay width at NLO when considering  $m_t = 172 - 173 \text{ GeV}$ .

systematic uncertainties affecting both variables are correlated and thus lead to more precise measurements.

Instead of measuring another variable, a two-dimensional approach would also allow to constrain certain systematic uncertainties. Similar to the ideogram method used for top quark mass measurements, the second dimension could be used to estimate the jet energy scale, which has a significant effect on the present measurements. Likewise, a variable whose distribution is not influenced by the top quark decay width can be used to constrain other systematic effects.

During 2017 the CMS detector recorded about  $42 \text{ fb}^{-1}$  of integrated luminosity. If similar conditions as for the 2016 data set are assumed, using both sets together would double the amount of selected data events. This would reduce the statistical uncertainty by a factor  $\sqrt{2}$ , resulting in roughly  $0.12 \text{ GeV}$ . Compared to the systematic uncertainty, which is around  $0.60 \text{ GeV}$ , increasing the amount of data events does not have a considerable influence on the result.

In the same way, the improvements that could be obtained by considering single top quark production as part of the signal are negligible. These events would enhance the fraction of correctly reconstructed events by less than  $0.5\%$ , which is much smaller than the effect that is currently observed for some systematic uncertainties.

The fraction of correctly reconstructed events can potentially be further improved using a multivariate analysis technique, but the largest progress can be made by actively tackling the systematic uncertainties. Especially jet-related modelling uncertainties, such as the shower scales, the matching between matrix element and parton shower, and colour reconnection, have a large impact. These need to be studied in more detail in order to reduce the intrinsic uncertainties on these effects. As a first step, larger simulated samples would give a better insight in how the different modelling behaves compared to the nominal sample. At the moment several effects are overestimated due to the large statistical uncertainty caused by the limited sample sizes. E.g. for the hadronic measurement the statistical uncertainty on the shift due to the underlying event is about three times larger than the shift itself. For other systematic uncertainties, the modelling uncertainties need to be further constrained using the data. This is exemplified in Figure 7.2, where the effect of the uncertainty due to final state radiation on the distributions of the sensitive variables  $m_r$  and  $m_{\text{lb},r}$  is visualised. Comparing the data to the three distributions, the uncertainties seem to be overestimated.

While acquiring more data does not directly lead to a more precise result, it opens possibilities for a smarter event selection. For example, when the instantaneous luminosity increases, this is reflected in the amount of pileup interactions. It can be investigated if the size of the systematic uncertainty is related to the number of pileup vertices in the event and, if so, an upper limit on the amount of primary vertices can be imposed. Similar techniques can be used to constrain other systematic effects. A very crude version of this was applied in this thesis using the additional event selection criteria in Chapter 5. As an example, Figure 7.3 shows the normalised distribution of the alternative top quark mass  $m_{b'jj}$  before applying these requirements. The effect of less or more FSR on the distribution is visualised for CM events separately and

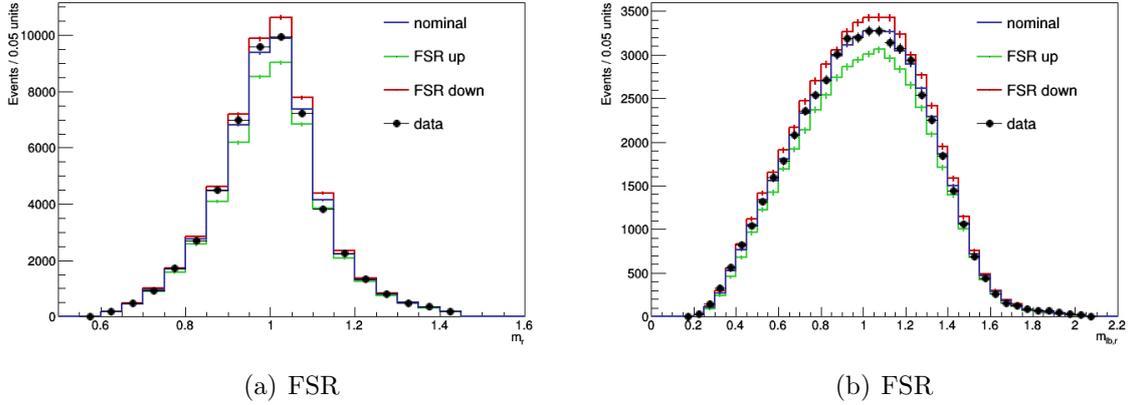


Figure 7.2: Effect of the uncertainty due to FSR on the distributions of (a)  $m_r$  and (b)  $m_{lb,r}$ .

for the combination of all events. It is clear that CM events with a low  $m_{b'_{jj}}$  are more affected by FSR than those with a higher  $m_{b'_{jj}}$ . A similar effect is seen for the events in the peak of the distribution when all event categories are considered. This provided a motivation to reject all events with  $m_{b'_{jj}} < 200$  GeV. An improvement can be achieved if events inducing a large systematic effect are identified. The correlation of the kinematic properties of these events with the size of several systematic uncertainties can be investigated to construct the optimal event selection process.

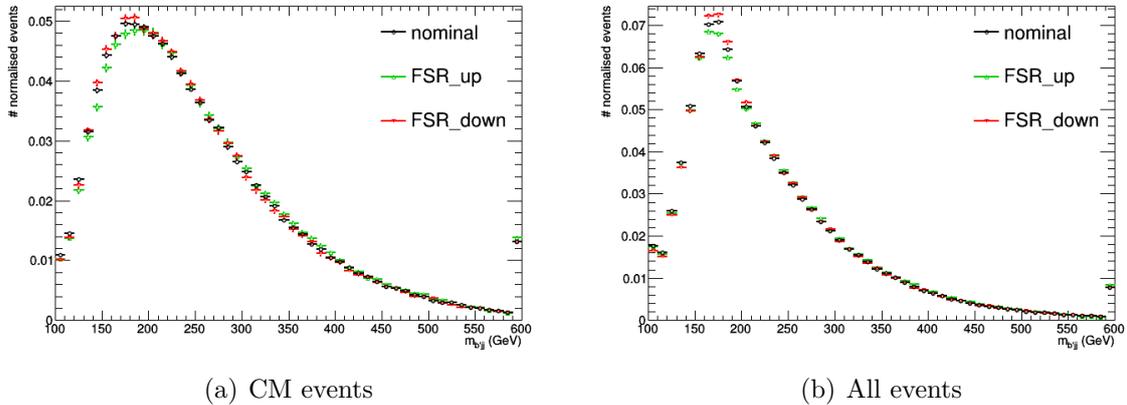


Figure 7.3: Normalised distribution of the alternative top quark mass  $m_{b'_{jj}}$  for (a) CM events and (b) all events. The first (last) bin contains the underflow (overflow).

A proof of the validity of this argument can be found in this analysis. Despite selecting less than a third of the number of events that were selected for the ATLAS measurement, the statistical uncertainty in this analysis is a factor 2 smaller. The better reconstruction of this subset of events thus leads to a smaller statistical spread in the data. Further, an improvement in the systematic uncertainty is obtained even though more sources of systematic effects, such as the uncertainty on the top quark mass and

more colour reconnection models, are included. In addition, the uncertainties due to the shower scales and matching, which are the main uncertainty sources for this analysis, are considered to be correlated in ATLAS measurement, while these are treated as completely independent uncertainties in this analysis.

Ultimately, a more precise measurement of the top quark width can be achieved at future lepton colliders, where a lot of jet-related uncertainties are expected to play a subdominant role. Firstly, the uncertainties related to the compositeness of the proton should no longer be taken into account. Concretely, there is no need to determine uncertainties due to the parton density functions and there are no proton remnants to produce underlying event interactions. Secondly, no gluons are radiated in the initial state of the hard-scattering process. The combination of these effects also reduces the impact of colour reconnection. In addition, the reduced amount of final-state jets that are not connected to the hard-scattering process would positively impact the reconstruction efficiency.

Potential future accelerators that would be suitable for top quark studies are the International Linear Collider (ILC) [179], the Compact Linear Collider (CLIC) [180], and the Triple-Large Electron-Positron collider (TLEP, also known as FCC-ee) [181]. All of these are electron-positron ( $e^+e^-$ ) colliders and are expected to take data at centre-of-mass energies between 200 – 1000 GeV, 350 – 3000 GeV, and 90 – 350 GeV, respectively. These energy ranges are well-suited to perform a so-called threshold scan. Leptons not being composite particles, the centre-of-mass energy of a lepton collision can easily be determined by adding the energies of the colliding leptons. Therefore, collisions can be produced at predetermined centre-of-mass energies. If this is done for several energies around the production threshold of a certain process, a threshold scan is performed. The centre-of-mass energy needed to produce a top quark pair is around twice the top quark mass,  $\sqrt{s} \sim 2 \times m_t$ . Around this point the production cross section increases, as is visualised in Figure 7.4 [182], where the parameters of the TESLA project [183] are used, which was later incorporated into the ILC. This upswing occurs earlier for lower than for higher masses. Moreover, the value of the top quark decay width influences the slope of the curve. Larger widths result in a more gradual slope, while the incline is more abrupt for smaller width values. Thus, measuring the top quark pair production cross section at several centre-of-mass energies can give a lot of information about the top quark properties. Because of the large correlations, it is advised to make a simultaneous measurement of these variables [182]. In order to increase the precision on the measurements, other quantities, such as the position of the peak in the momentum distribution, can be used to disentangle the effects of the variables. Assuming that an integrated luminosity of about  $30 \text{ fb}^{-1}$  is collected in each scan point, a simultaneous measurement using the quantities described above is expected to attain a precision of about 20 MeV for the top quark mass and about 30 MeV for its decay width [182, 184]. The uncertainties increase to 35 MeV and 40 MeV, respectively, if only the cross section would be used as a sensitive quantity.

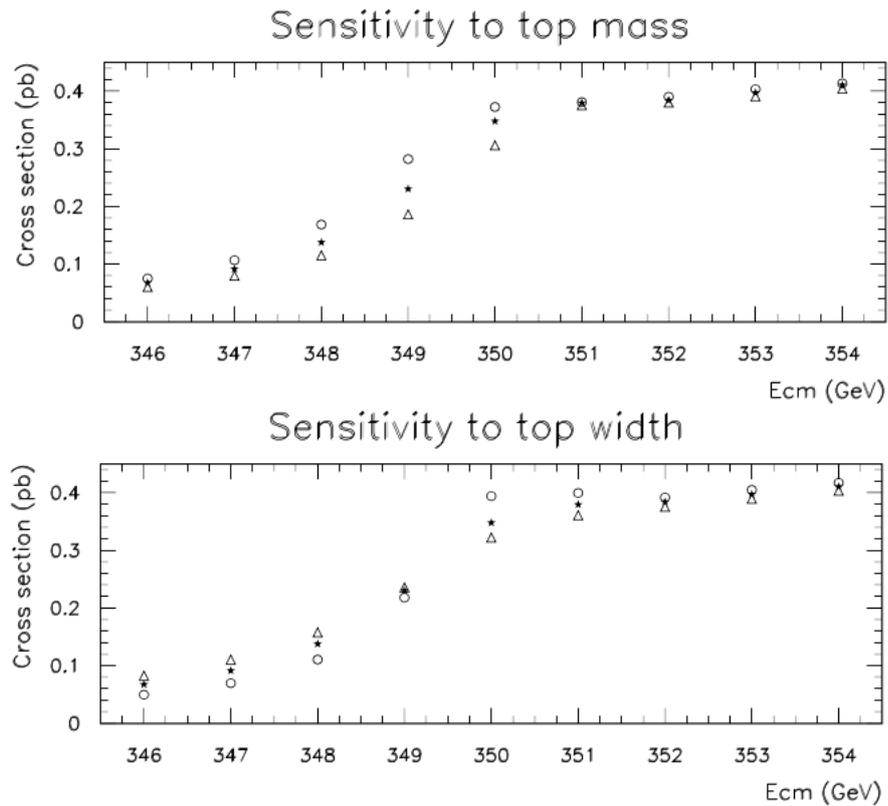


Figure 7.4: Top quark pair production cross section as a function of the centre-of-mass energy. The sensitivity to the top quark mass (decay width) is estimated by varying the mass (width) up and down by 200 MeV (400 MeV). Upward variations are indicated by triangles and downward variations by circles. Image adapted from [182].

# Summary

---

The standard model (SM) of particle physics describes the fundamental particles and their interactions. It is one of the most extensively tested theories during the last century. However, several observations indicate that the SM is not the end of the story. As it is now, the SM cannot pinpoint the nature of the majority of mass in the universe, nor does it explain how gravity works on quantum scales. Being the heaviest particle in the SM, the top quark is expected to be most sensitive to beyond the SM effects, such as interactions with as yet unknown particles. Besides direct searches for new physics, precise measurements of the properties of SM particles and interactions can be compared to their predictions, thus favouring or excluding certain beyond the SM theories.

This thesis investigates the decay properties of the top quark. In the SM the top quark decays almost uniquely into a W boson and a b quark. The probability for this process to happen is reflected in the top quark decay width, which is predicted to have a value around 1.33 GeV. If the top quark is able to decay into other particles as well, as is possible in several extensions of the SM, the top quark decay width will be larger than the SM prediction.

A direct measurement of the top quark decay width is performed using proton collisions produced by the Large Hadron Collider (LHC) at a centre-of-mass energy of 13 TeV. The data were recorded by the CMS experiment and correspond to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . A dedicated selection process singles out events where top quarks are produced in pairs. Only the semimuonic decay channel, i.e. where one of the W bosons decays into a pair of quarks and the other into a muon and its corresponding neutrino, are considered. Therefore, events are required to have exactly one muon and exactly four jets, whereof two should originate from b quarks. The top quark pair events are simulated, as well as background processes that produce the same final state. The top quark pair system is reconstructed and the reconstruction efficiency is estimated by comparing the reconstruction to the generator information for simulated events. These are then categorised according to whether they are correctly reconstructed or not. Further selection requirements based on reconstructed variables are imposed so as to increase the fraction of well-reconstructed events.

Several measurements of the top quark decay width are performed using a maximum likelihood method, where templates are constructed using the distributions of variables that are sensitive to the top quark decay width. The best measurement obtained in this analysis results in a top quark decay width of  $\Gamma_t = 1.12 \pm 0.16$  (stat.)  $^{+0.65}_{-0.54}$  (syst.) GeV, which is consistent with the value predicted by the SM. This result represents the most precise direct measurement of the top quark decay width in the semileptonic decay

channel, reducing the uncertainties of the previous best result by 30%. Further, the measurement is consistent with the best result in the dilepton channel.

## Samenvatting

---

### **Directe meting van de vervalbreedte van de top quark in het muon+jets kanaal gebruikmakende van het CMS experiment aan de LHC**

Het standaard model (SM) van de deeltjesfysica beschrijft de elementaire deeltjes en hun interacties. Het is één van de meest uitgebreid geteste modellen van de voorbije eeuw. Desondanks wijzen enkele observaties erop dat het SM niet het volledige verhaal vertelt. In zijn huidige vorm kan het SM het merendeel van de massa in het heelal niet beschrijven, noch kan het verklaren hoe de zwaartekracht werkt op kwantumschalen. Aangezien de top quark het zwaarste deeltje van het SM is, wordt verwacht dat deze het gevoeligste is aan effecten van nieuwe fysica die nog niet door het SM beschreven wordt, zoals interacties met tot nog toe ongekende deeltjes. Precisiemetingen van de eigenschappen van de SM deeltjes en interacties vormen een aanvulling op directe zoektochten naar nieuwe fysica, aangezien deze vergeleken kunnen worden met hun verwachte waarden. Op deze manier kunnen theorieën die het SM uitbreiden gesteund of verworpen worden.

Dit proefschrift onderzoekt het verval van de top quark. Volgens het SM vervalft deze bijna uitsluitend in een W boson en een b quark. De waarschijnlijkheid voor dit proces zit vervat in de vervalbreedte van de top quark, die een verwachte waarde van ongeveer 1.33 GeV heeft. Als de top quark ook in andere deeltjes kan vervallen, zoals sommige uitbreidingen van het SM voorspellen, zal de vervalbreedte een grotere waarde hebben.

Gebruikmakende van protonbotsingen met een massamiddelpuntsenergie van 13 TeV die geproduceerd worden door de Large Hadron Collider, wordt een directe meting van de vervalbreedte van de top quark gedaan. De gebruikte data werd verzameld door het CMS experiment en komt overeen met een geïntegreerde luminositeit van  $35.9 \text{ fb}^{-1}$ . Deze analyse beschouwt enkel gebeurtenissen waar top quarks in paren worden geproduceerd en deze op een semi-muonische manier vervallen. Dit wil zeggen dat een van de W bosonen vervalft in een quark-antiquark paar en het andere in een muon en een muon-neutrino. Gebeurtenissen moeten dus exact één muon en exact vier jets bevatten, waarvan er twee afstammen van een b quark. Net als top quark paar gebeurtenissen worden ook achtergrondprocessen die dezelfde deeltjes produceren gesimuleerd. De gebeurtenissen worden dan gereconstrueerd en de reconstructie-efficiëntie wordt geschat door de reconstructie te vergelijken met de extra informatie die vervat zit in de gesimuleerde gebeurtenissen. Afhankelijk van of de gebeurtenissen goed gereconstrueerd zijn of niet worden deze onderverdeeld in verschillende categorieën. Met

behulp van een specifieke selectieprocedure die gebruik maakt van enkele gereconstrueerde variabelen wordt de fractie van correct gereconstrueerde gebeurtenissen geoptimaliseerd.

Verschillende metingen van de vervalbreedte van de top quark worden uitgevoerd gebruikmakende van een maximum likelihood methode, waarbij waarschijnlijkheidsverdelingen geconstrueerd worden aan de hand van de verdeling van variabelen die gevoelig zijn aan de vervalbreedte van de top quark. De beste meting die zo verkregen wordt is gelijk aan  $\Gamma_t = 1.12 \pm 0.16$  (stat.)  $^{+0.65}_{-0.54}$  (syst.) GeV, wat in overeenstemming is met de waarde die door het SM voorspeld wordt. Dit resultaat is de meest precieze meting van de vervalbreedte van de top quark in het semileptonische kanaal en reduceert de onzekerheden van het vorige beste resultaat met 30%. Verder is deze meting consistent met het beste resultaat in het dileptonische kanaal.

## Bibliography

---

- [1] University of Toronto, Environmental Health and Safety Department, *Radiation Protection Training Manual*, (last consulted: 24 September 2017).  
<https://ehs.utoronto.ca/our-services/radiation-safety/radiation-protection-manual/>
- [2] F. Mandl and G. G. Shaw, *Quantum field theory; 2nd ed.* Wiley, New York, NY, 2010.
- [3] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory; 1995 ed.* Westview, Boulder, CO, 1995.
- [4] D. J. Griffiths, *Introduction to elementary particles; 2nd rev. version.* Physics textbook. Wiley, New York, NY, 2008.
- [5] P. Langacker, *The standard model and beyond.* Series in High Energy Physics, Cosmology and Gravitation. Taylor and Francis, Boca Raton, FL, 2010.
- [6] D. J. Griffiths, *Introduction to Quantum Mechanics; 2nd ed.* Pearson, Upper Saddle River, NJ, 2005.
- [7] National Institute of Standards and Technology, Gaithersburg, MD., *The NIST Reference on Constants, Units and Uncertainties*, 2015, (last consulted: 15 October 2017). [Online]. Available:  
<https://physics.nist.gov/cuu/index.html>
- [8] Particle Data Group Collaboration, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. C **40** (2016), no. 10, 100001,  
doi:10.1088/1674-1137/40/10/100001.
- [9] F. Muheim, *Weak interaction*, 2006, Particle physics lectures, University of Edinburgh, UK, (last consulted: 3 October 2017).  
<http://www2.ph.ed.ac.uk/~muheim/teaching/np3/lect-weak.pdf>
- [10] DESY Press Release, *The most precise picture of the proton, HERA experiments H1 and ZEUS publish combined analysis*, Jul, 2015, (last consulted: 2 October 2017). [http://www.desy.de/news/news\\_search/index\\_eng.html?openDirectAnchor=829](http://www.desy.de/news/news_search/index_eng.html?openDirectAnchor=829)

- [11] The 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. B **716** (2012), no. 1, 1 – 29, doi:10.1016/j.physletb.2012.08.020.
- [12] The CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012), no. 1, 30 – 61, doi:10.1016/j.physletb.2012.08.021.
- [13] The ATLAS and CMS Collaboration, G. Aad *et al.*, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at  $\sqrt{s} = 7$  and 8 TeV*, JHEP **08** (2016) 045, doi:10.1007/JHEP08(2016)045, arXiv:1606.02266.
- [14] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (Aug, 1964) 321–323, doi:10.1103/PhysRevLett.13.321.
- [15] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (Oct, 1964) 508–509, doi:10.1103/PhysRevLett.13.508.
- [16] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (Nov, 1964) 585–587, doi:10.1103/PhysRevLett.13.585.
- [17] B. Isildak, *Measurement of the differential dijet production cross section in proton-proton collisions at  $\sqrt{s} = 7$  TeV*. PhD thesis, Bogazici U., 2011. arXiv:1308.6064.
- [18] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (Feb, 1957) 1413–1415, doi:10.1103/PhysRev.105.1413.
- [19] J. Wacker, *Why does the Higgs field have a Mexican hat shape?*, Dec, 2016, Quora (last consulted: 8 October 2017). <https://www.quora.com/Why-does-the-Higgs-field-have-a-Mexican-hat-shape>
- [20] R. Argurio, ed., *The String Theory Universe*, Fortschritte der Physik, vol. 64. Apr, 2016. Proceedings of the 21st European String Workshop and 3rd COST MP1210 Meeting, Leuven, Belgium. <https://cds.cern.ch/record/2263090>.
- [21] The SNO Collaboration, Q. R. Ahmad *et al.*, *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301, doi:10.1103/PhysRevLett.89.011301, arXiv:nucl-ex/0204008.
- [22] The Super-Kamiokande Collaboration, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562–1567, doi:10.1103/PhysRevLett.81.1562, arXiv:hep-ex/9807003.

- [23] A. de Gouvêa, *Neutrino Mass Models*, Annual Review of Nuclear and Particle Science **66** (2016), no. 1, 197–217, doi:10.1146/annurev-nucl-102115-044600, <https://doi.org/10.1146/annurev-nucl-102115-044600>.
- [24] SLD Electroweak Group, DELPHI, ALEPH, SLD, SLD Heavy Flavour Group, OPAL, LEP Electroweak Working Group, L3 Collaboration, S. Schael *et al.*, *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006) 257–454, doi:10.1016/j.physrep.2005.12.006, arXiv:hep-ex/0509008.
- [25] S. P. Martin, *A Supersymmetry primer*, doi:10.1142/9789812839657, 10.1142/9789814307505, arXiv:hep-ph/9709356. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].
- [26] N. Polonsky, *Supersymmetry: Structure and phenomena. Extensions of the standard model*, Lect. Notes Phys. Monogr. **68** (2001) 1–169, doi:10.1007/3-540-44642-7, arXiv:hep-ph/0108236.
- [27] *The CDF experiment*, <http://www-cdf.fnal.gov/physics/physics.html>.
- [28] *The D0 experiment*, <http://www-d0.fnal.gov/>.
- [29] *The Tevatron Collider*, <http://tevatron.fnal.gov>.
- [30] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [31] The CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.
- [32] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.
- [33] The Gfitter Group Collaboration, M. Baak, J. Cuth, J. Haller, A. Hoecker, R. Kogler, K. Moenig, M. Schott, and J. Stelzer, *The global electroweak fit at NNLO and prospects for the LHC and ILC*, Eur. Phys. J. C **74** (Jul, 2014) 3046. 26 p. <http://cds.cern.ch/record/1743339>.
- [34] LHC Top Physics Working Group, *NLO single-top channel cross sections, ATLAS-CMS recommended predictions for single-top cross sections using the Hathor v2.1 program*, Sep, 2017, (last consulted: 13 October 2017). <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec>
- [35] The CDF Collaboration, *Precision Top-Quark Mass Measurement at CDF*, Phys. Rev. Lett. **109** (Oct, 2012) 152003, doi:10.1103/PhysRevLett.109.152003.
- [36] The CDF Collaboration, *Top quark mass measurement using the template method at CDF*, Phys. Rev. D **83** (Jun, 2011) 111101, doi:10.1103/PhysRevD.83.111101.
- [37] The CDF Collaboration, *Measurement of the top quark mass in the all-hadronic mode at CDF*, Phys. Lett. B **714** (2012), no. 1, 24 – 31, doi:10.1016/j.physletb.2012.06.007.

- [38] The CDF Collaboration, *Top-quark mass measurement in events with jets and missing transverse energy using the full CDF data set*, Phys. Rev. D **88** (Jul, 2013) 011101, doi:10.1103/PhysRevD.88.011101.
- [39] The D0 Collaboration, *Precise measurement of the top-quark mass from lepton + jets events at D0*, Phys. Rev. D **84** (Aug, 2011) 032004, doi:10.1103/PhysRevD.84.032004.
- [40] The D0 Collaboration, *Measurement of the top-quark mass in  $p\bar{p}$  collisions using events with two leptons*, Phys. Rev. D **86** (Sep, 2012) 051103, doi:10.1103/PhysRevD.86.051103.
- [41] The ATLAS Collaboration, *Measurement of the top quark mass in the  $t\bar{t} \rightarrow$  lepton+jets and  $t\bar{t} \rightarrow$  dilepton channels using  $\sqrt{s} = 7$  TeV ATLAS data*, Eur. Phys. J. C **75** (Jul, 2015) 330, doi:10.1140/epjc/s10052-015-3544-0.
- [42] The ATLAS Collaboration, *Measurement of the top-quark mass in the fully hadronic decay channel from ATLAS data at  $\sqrt{s} = 7$  TeV*, Eur. Phys. J. C **75** (Apr, 2015) 158, doi:10.1140/epjc/s10052-015-3373-1.
- [43] The ATLAS Collaboration, *Measurement of the top quark mass in the  $t\bar{t} \rightarrow$  dilepton channel from  $\sqrt{s} = 8$  TeV ATLAS data*, Phys. Lett. B **761** (2016), no. Supplement C, 350 – 371, doi:10.1016/j.physletb.2016.08.042.
- [44] The ATLAS Collaboration, *Top-quark mass measurement in the all-hadronic  $t\bar{t}$  decay channel at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, JHEP **09** (2017) 118, doi:10.1007/JHEP09(2017)118, arXiv:1702.07546.
- [45] The ATLAS Collaboration, *Measurement of the top quark mass in the  $t\bar{t} \rightarrow$  lepton+jets channel from  $\sqrt{s} = 8$  TeV ATLAS data*, Tech. Rep. ATLAS-CONF-2017-071, CERN, Geneva, Sep, 2017.
- [46] *Measurement of the top quark mass in topologies enhanced with single top-quarks produced in the  $t$ -channel in  $\sqrt{s} = 8$  TeV ATLAS data*, Tech. Rep. ATLAS-CONF-2014-055, CERN, Geneva, Sep, 2014.
- [47] The CMS Collaboration, S. Chatrchyan *et al.*, *Measurement of the  $t\bar{t}$  production cross section and the top quark mass in the dilepton channel in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*, JHEP **07** (2011) 049, doi:10.1007/JHEP07(2011)049, arXiv:1105.5661.
- [48] The CMS Collaboration, *Measurement of the top-quark mass in  $t\bar{t}$  events with dilepton final states in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*, Eur. Phys. J. C **72** (Oct, 2012) 2202, doi:10.1140/epjc/s10052-012-2202-z.
- [49] The CMS Collaboration, *Measurement of the top-quark mass in all-jets  $t\bar{t}$  events in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*, Eur. Phys. J. C **74** (Apr, 2014) 2758, doi:10.1140/epjc/s10052-014-2758-x.

- [50] The CMS Collaboration, *Measurement of the top-quark mass in  $t\bar{t}$  events with lepton+jets final states in  $pp$  collisions at  $\sqrt{s} = 7$  TeV*, JHEP **12** (2012) 105, doi:10.1007/JHEP12(2012)105, arXiv:1209.2319.
- [51] The CMS Collaboration, *Measurement of the top quark mass using proton-proton data at  $\sqrt{s} = 7$  and 8 TeV*, Phys. Rev. D **93** (Apr, 2016) 072004, doi:10.1103/PhysRevD.93.072004.
- [52] The CMS Collaboration, *Measurement of the top quark mass using single top quark events in proton-proton collisions at  $\sqrt{s} = 8$  TeV*, Eur. Phys. J. C **77** (May, 2017) 354, doi:10.1140/epjc/s10052-017-4912-8.
- [53] The CMS Collaboration, *Measurement of the top quark mass with lepton+jets final states in  $pp$  collisions at  $\sqrt{s} = 13$  TeV*, Tech. Rep. CMS-PAS-TOP-17-007, CERN, Geneva, 2017.
- [54] The ATLAS and CMS Collaboration, *Combination of ATLAS and CMS results on the mass of the top quark using up to 4.9 inverse femtobarns of data*, Tech. Rep. CMS-PAS-TOP-13-005, CERN, Geneva, 2013.
- [55] The CDF and D0 Collaboration, Tevatron Electroweak Working Group, *Combination of CDF and D0 results on the mass of the top quark using up to 9.7  $\text{fb}^{-1}$  at the Tevatron*, arXiv:1407.2682.
- [56] The ATLAS, CDF, CMS and D0 Collaboration, *First combination of Tevatron and LHC measurements of the top-quark mass*, arXiv:1403.4427.
- [57] LHC Top Physics Working Group, *LHCTopWG Summary Plots*, Sep, 2017, (last consulted: 14 October 2017). <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots>
- [58] I. I. Y. Bigi, Y. L. Dokshitzer, V. A. Khoze, J. H. Kuhn, and P. M. Zerwas, *Production and Decay Properties of Ultraheavy Quarks*, Phys. Lett. B **181** (1986) 157–163, doi:10.1016/0370-2693(86)91275-X.
- [59] M. Jezabek and J. H. Kuhn, *The Top width: Theoretical update*, Phys. Rev. **D48** (1993) R1910–R1913, doi:10.1103/PhysRevD.49.4970, 10.1103/PhysRevD.48.R1910, hep-ph/9302295. [Erratum: Phys. Rev.D49,4970(1994)].
- [60] A. Czarnecki and K. Melnikov, *Two loop QCD corrections to top quark width*, Nucl. Phys. B **544** (1999) 520–531, doi:10.1016/S0550-3213(98)00844-X, arXiv:hep-ph/9806244.
- [61] K. G. Chetyrkin, R. Harlander, T. Seidensticker, and M. Steinhauser, *Second order QCD corrections to  $\Gamma(t \rightarrow Wb)$* , Phys. Rev. D **60** (1999) 114015, doi:10.1103/PhysRevD.60.114015, arXiv:hep-ph/9906273.

- [62] J. Gao, C. S. Li, and H. X. Zhu, *Top Quark Decay at Next-to-Next-to Leading Order in QCD*, Phys. Rev. Lett. **110** (2013), no. 4, 042001, doi:10.1103/PhysRevLett.110.042001, arXiv:1210.2808.
- [63] CDF Collaboration, T. Aaltonen *et al.*, *First Direct Bound on the Total Width of the Top Quark in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$ -TeV*, Phys. Rev. Lett. **102** (2009) 042001, doi:10.1103/PhysRevLett.102.042001, arXiv:0808.2167.
- [64] CDF Collaboration, T. Aaltonen *et al.*, *Direct Top-Quark Width Measurement from Lepton + Jets Events at CDF II*, Phys. Rev. Lett. **105** (2010) 232003, doi:10.1103/PhysRevLett.105.232003, arXiv:1008.3891.
- [65] CDF Collaboration, T. A. Aaltonen *et al.*, *Direct Measurement of the Total Decay Width of the Top Quark*, Phys. Rev. Lett. **111** (2013), no. 20, 202001, doi:10.1103/PhysRevLett.111.202001, arXiv:1308.4050.
- [66] The CMS Collaboration, *Bounding the top quark width using final states with two charged leptons and two jets at  $\sqrt{s} = 13$  TeV*, Tech. Rep. CMS-PAS-TOP-16-019, CERN, Geneva, 2016.
- [67] The ATLAS Collaboration, M. Aaboud *et al.*, *Direct top-quark decay width measurement in the  $t\bar{t}$  lepton+jets channel at  $\sqrt{s} = 8$  TeV with the ATLAS experiment*, arXiv:1709.04207. Submitted to *EPJC*.
- [68] D0 Collaboration, V. M. Abazov *et al.*, *Determination of the width of the top quark*, Phys. Rev. Lett. **106** (2011) 022001, doi:10.1103/PhysRevLett.106.022001, arXiv:1009.5686.
- [69] D0 Collaboration, V. M. Abazov *et al.*, *An Improved determination of the width of the top quark*, Phys. Rev. D **85** (2012) 091104, doi:10.1103/PhysRevD.85.091104, arXiv:1201.4156.
- [70] CMS Collaboration, V. Khachatryan *et al.*, *Measurement of the ratio  $\mathcal{B}(t \rightarrow Wb)/\mathcal{B}(t \rightarrow Wq)$  in  $pp$  collisions at  $\sqrt{s} = 8$  TeV*, Phys. Lett. B **736** (2014) 33–57, doi:10.1016/j.physletb.2014.06.076, arXiv:1404.2292.
- [71] J. A. Aguilar-Saavedra, *Top flavor-changing neutral interactions: Theoretical expectations and experimental detection*, Acta Phys. Polon. B **35** (2004) 2695–2710, arXiv:hep-ph/0409342.
- [72] A. L. Read, *Presentation of search results: the CLs technique*, Journal of Physics G: Nuclear and Particle Physics **28** (2002), no. 10, 2693.
- [73] CERN, *LHC Guide*, Mar, 2017, <https://cds.cern.ch/record/2255762>
- [74] The LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.
- [75] The ALICE Collaboration, *The ALICE Experiment at the CERN LHC*, JINST **3** (2008) S08002.

- [76] The LHCf Collaboration, *The LHCf Detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006.
- [77] The TOTEM Collaboration, *The TOTEM Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.
- [78] The CMS-TOTEM Collaboration, M. Albrow, M. Arneodo, V. Avati, J. Baechler, N. Cartiglia, M. Deile, M. Gallinaro, J. Hollar, M. Lo Vetere, K. Oesterberg, N. Turini, J. Varela, and D. Wright, *CMS-TOTEM Precision Proton Spectrometer*, Tech. Rep. CERN-LHCC-2014-021. TOTEM-TDR-003. CMS-TDR-13, Sep, 2014.
- [79] CERN, *Pulling together: Superconducting electromagnets*, (last consulted: 3 June 2017). <http://home.cern/about/engineering/pulling-together-superconducting-electromagnets>
- [80] X. C. Vidal and R. Cid, *LHC layout, Taking a closer look at LHC*, (last consulted: 3 June 2017). [https://www.lhc-closer.es/taking\\_a\\_closer\\_look\\_at\\_lhc/0.lhc\\_layout](https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.lhc_layout)
- [81] CERN, *Radiofrequency cavities*, (last consulted: 3 June 2017). <http://home.cern/about/engineering/radiofrequency-cavities>
- [82] The CMS Collaboration, *CMS Luminosity – Public Results*, (last consulted: 19 February 2018). <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>
- [83] *CMS installs the world’s largest silicon detector*, Mar, 2008. <https://cds.cern.ch/record/1734187>
- [84] K. Gill, *CMS Tracker services*, Jun, 2013, presented at the Forum on Tracking Detector Mechanics, Oxford. [https://indico.cern.ch/event/233332/contributions/1546083/attachments/388966/540961/Oxford\\_Gill.pdf](https://indico.cern.ch/event/233332/contributions/1546083/attachments/388966/540961/Oxford_Gill.pdf)
- [85] J. Freeman, *Innovations for the CMS HCAL*, International Journal of Modern Physics A **25** (2010), no. 12, 2421–2436, doi:10.1142/S0217751X10049682, <http://www.worldscientific.com/doi/pdf/10.1142/S0217751X10049682>.
- [86] J. Anderson, J. Freeman, S. Los, and J. Whitmore, *Upgrade of the CMS Hadron Outer Calorimeter with SIPMs*, Physics Procedia **37** (2012) 72 – 78, doi:10.1016/j.phpro.2012.02.358.
- [87] The CMS Collaboration, *Technical proposal for the upgrade of the CMS detector through 2020*, Tech. Rep. CERN-LHCC-2011-006. LHCC-P-004, Jun, 2011.
- [88] The CMS Collaboration, B. Bilki, *Commissioning of Upgrade Forward Hadron Calorimeters of CMS*, Tech. Rep. CMS-CR-2016-417, CERN, Geneva, Nov, 2016.
- [89] The CMS Collaboration, *Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data*, JINST **5** (2010) T03012.

- [90] The CMS Collaboration, G. Abbiendi, *The CMS muon system in Run2: preparation, status and first results*, PoS **EPS-HEP2015** (Oct, 2015) 237. 8 p.
- [91] The CMS Collaboration, C. Battilana, *The CMS muon system status and upgrades for LHC run-2 and performance of muon reconstruction with 13 TeV data*, Tech. Rep. CMS-CR-2016-437, CERN, Geneva, Dec, 2016.
- [92] T. A. Bawej, U. Behrens, J. Branson, O. Chaze, S. Cittolin, G. L. Darlea, C. Deldicque, M. Dobson, A. Dupont, S. Erhan, A. K. Forrest, D. Gigi, F. Glege, G. Gomez Ceballos, R. Gomez-Reino Garrido, J. G. Hegeman, A. G. Holzner, L. Masetti, F. Meijers, E. Meschi, R. Mommsen, S. Morovic, V. O'Dell, L. Orsini, C. M. E. Paus, A. Petrucci, M. Pieri, A. Racz, H. Sakulin, C. Schwick, B. B. Stieger, K. Sumorok, J. Veverka, and P. Zejdl, *The New CMS DAQ System for Run 2 of the LHC*, Tech. Rep. CMS-CR-2014-082, CERN, Geneva, May, 2014.
- [93] The CMS Collaboration, L. Cadamuro, *The CMS Level-1 Trigger for LHC Run II*, Tech. Rep. CMS-CR-2016-412, CERN, Geneva, Nov, 2016.
- [94] The CMS Collaboration, A. Tapper, *The CMS Level-1 Trigger for LHC Run II*, Tech. Rep. CMS-CR-2016-303, CERN, Geneva, Oct, 2016.
- [95] The CMS Collaboration, *The CMS trigger system*, JINST **12** (2017) P01020.
- [96] The CMS Collaboration, A. Perrotta, *Performance of the CMS High Level Trigger*, Tech. Rep. CMS-CR-2015-096, CERN, Geneva, May, 2015.
- [97] L. Rossi and O. Brüning, *The High Luminosity Large Hadron Collider: the new machine for illuminating the mysteries of Universe*. Advanced series on directions in high energy physics. World Scientific, Hackensack, NJ, 2015.
- [98] F. Bordry, *From LHC to post LHC era*, Nov, 2017, presented at the 7th HL-LHC Collaboration Meeting, Madrid. [https://indico.cern.ch/event/647714/contributions/2632835/attachments/1557114/2449313/Fk\\_Bordry\\_From\\_LHC\\_to\\_post\\_LHC\\_era\\_Madrid\\_November\\_2017.pdf](https://indico.cern.ch/event/647714/contributions/2632835/attachments/1557114/2449313/Fk_Bordry_From_LHC_to_post_LHC_era_Madrid_November_2017.pdf)
- [99] The CMS Collaboration, *The Phase-2 Upgrade of the CMS Tracker*, Tech. Rep. CERN-LHCC-2017-009. CMS-TDR-014, CERN, Geneva, Jun, 2017.
- [100] The CMS Collaboration, A. H. Dierlamm, *The CMS Outer Tracker for HL-LHC*, Tech. Rep. CMS-CR-2018-014, CERN, Geneva, Jan, 2018.
- [101] The CMS Collaboration, G. Sguazzoni, *Upgrades of the CMS Outer Tracker for HL-LHC. Upgrades of the CMS Outer Tracker detector for the HL-LHC*, Tech. Rep. CMS-CR-2016-040, CERN, Geneva, Mar, 2016.
- [102] D. Contardo, M. Klute, J. Mans, L. Silvestris, and J. Butler, *Technical Proposal for the Phase-II Upgrade of the CMS Detector*, Tech. Rep. CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02, Geneva, Jun, 2015.

- [103] The CMS Collaboration, *CMS Data Quality Information*, (last consulted: 19 February 2018).  
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/DataQuality>
- [104] The Tracker DPG group, *CMS Silicon Strip Performance Results 2016*, (last consulted: 9 February 2018).  
<https://twiki.cern.ch/twiki/bin/view/CMS/StripsOfflinePlots2016>
- [105] D. Zeppenfeld, *Event generation and parton shower*, <http://www.gk-eichtheorien.physik.uni-mainz.de/Dateien/Zeppenfeld-3.pdf>
- [106] The NNPDF Collaboration, R. D. Ball *et al.*, *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, doi:10.1007/JHEP04(2015)040, 1410.8849.
- [107] S. Carrazza, A. Ferrara, D. Palazzo, and J. Rojo, *APFEL Web: a web-based application for the graphical visualization of parton distribution functions*, Journal of Physics G: Nuclear and Particle Physics **42** (2015) 057001.
- [108] 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.
- [109] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5 : Going Beyond*, JHEP **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.
- [110] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [111] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:0409146.
- [112] 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.
- [113] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [114] S. Alioli, S.-O. Moch, and P. Uwer, *Hadronic top-quark pair-production with one jet and parton showering*, JHEP **01** (2012) 137, doi:10.1007/JHEP01(2012)137, arXiv:1110.5251.
- [115] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO single-top production matched with shower in POWHEG: s- and t-channel contributions*, JHEP **09** (2009) 111,

- doi:10.1007/JHEP02(2010)011, 10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: JHEP02,011(2010)].
- [116] E. Re, *Single-top  $Wt$ -channel production matched with parton showers using the POWHEG method*, Eur. Phys. J. C **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.
- [117] M. Dobbs, S. Frixione, E. Laenen, and K. Tollefson, *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.
- [118] 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)].
- [119] 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)].
- [120] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. B **126** (1977) 298–318, doi:10.1016/0550-3213(77)90384-4.
- [121] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
- [122] M. Bahr *et al.*, *Herwig++ Physics and Manual*, Eur. Phys. J. C **58** (2008) 639–707, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.
- [123] S. Gieseke, C. Rohr, and A. Siodmok, *Colour reconnections in Herwig++*, Eur. Phys. J. C **72** (2012) 2225, doi:10.1140/epjc/s10052-012-2225-5, arXiv:1206.0041.
- [124] J. Alwall *et al.*, *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. C **53** (2008) 473–500, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.
- [125] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, *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.
- [126] P. Nason, *Shower Monte Carlo at Next-to-Leading Order*, Oct, 2008, presented at the Cavendish HEP Seminars, University of Cambridge. <http://www.hep.phy.cam.ac.uk/seminars/talks/2008/michaelmas/PaoloNason-14102008.pdf>

- [127] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, JHEP **04** (2009) 002, doi:10.1088/1126-6708/2009/04/002, arXiv:0812.0578.
- [128] G. Heinrich, S. P. Jones, M. Kerner, G. Luisoni, and E. Vryonidou, *NLO predictions for Higgs boson pair production with full top quark mass dependence matched to parton showers*, arXiv:1703.09252.
- [129] B. Andersson, *The Lund Model*, vol. 7 of *Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology*. Cambridge University Press, 2005.
- [130] T. Sjöstrand, *Old ideas in hadronisation: The Lund string*, Apr, 2009, presented at the ‘New ideas in hadronisation’ workshop, Durham. <http://home.thep.lu.se/~torbjorn/talks/durham09.pdf>
- [131] The CMS Collaboration, *Event generator tunes obtained from underlying event and multiparton scattering measurements*, Eur. Phys. J. C **76** (2016), no. 3, 155, doi:10.1140/epjc/s10052-016-3988-x.
- [132] The CMS Collaboration, *Investigations of the impact of the parton shower tuning in Pythia 8 in the modelling of  $t\bar{t}$  at  $\sqrt{s} = 8$  and 13 TeV*, Tech. Rep. CMS-PAS-TOP-16-021, CERN, Geneva, 2016.
- [133] S. Argyropoulos and T. Sjöstrand, *Effects of color reconnection on  $t\bar{t}$  final states at the LHC*, JHEP **11** (2014) 043, doi:10.1007/JHEP11(2014)043, arXiv:1407.6653.
- [134] T. Sjöstrand *et al.*, *Pythia 8.2 Online Manual: Colour Reconnection*, (last consulted: 17 July 2017). <http://home.thep.lu.se/~torbjorn/pythia82html/ColourReconnection.html>
- [135] C. Bierlich, G. Gustafson, L. Lönnblad, and A. Tarasov, *Effects of Overlapping Strings in  $pp$  Collisions*, JHEP **03** (2015) 148, doi:10.1007/JHEP03(2015)148, arXiv:1412.6259.
- [136] J. R. Christiansen and P. Z. Skands, *String Formation Beyond Leading Colour*, JHEP **08** (2015) 003, doi:10.1007/JHEP08(2015)003, arXiv:1505.01681.
- [137] The GEANT4 Collaboration, S. Agostinelli, J. Allison, *et al.*, *Geant4 – a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003), no. 3, 250 – 303, doi:10.1016/S0168-9002(03)01368-8.
- [138] The GEANT4 Collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270, doi:10.1109/TNS.2006.869826.
- [139] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.

- [140] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann, *HATHOR: HAdronic Top and Heavy quarks crOSS section calculatoR*, Comput. Phys. Commun. **182** (2011) 1034–1046, doi:10.1016/j.cpc.2010.12.040, arXiv:1007.1327.
- [141] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz, P. Rieck, and P. Uwer, *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, Comput. Phys. Commun. **191** (2015) 74–89, doi:10.1016/j.cpc.2015.02.001, arXiv:1406.4403.
- [142] N. Kidonakis, *Theoretical results for electroweak-boson and single-top production*, PoS **DIS2015** (2015) 170, arXiv:1506.04072.
- [143] The CMS Collaboration, A. M. Sirunyan *et al.*, *Particle-flow reconstruction and global event description with the CMS detector*, JINST **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- [144] The CMS Collaboration, D. Barney, *Animated CMS slice for Powerpoint*, <https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=5581>.
- [145] S. Cucciarelli, M. Konecki, D. Kotliński, and T. Todorov, *Track reconstruction, primary vertex finding and seed generation with the Pixel Detector*, CMS NOTE-2006/026.
- [146] The CMS Collaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, JINST **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.
- [147] W. Adam, B. Mangano, T. Speer, and T. Todorov, *Track reconstruction in the CMS Tracker*, CMS NOTE-2006/041.
- [148] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987), no. 2-3, 444 – 450, doi:10.1016/0168-9002(87)90887-4. <http://www.sciencedirect.com/science/article/pii/0168900287908874>.
- [149] S. Cucciarelli, M. Konecki, D. Kotliński, and T. Todorov, *Track-Parameter Evaluation and Primary-Vertex Finding with the Pixel Detector*, Tech. Rep. CMS-NOTE-2003-026, CERN, Geneva, Sep, 2003.
- [150] The CMS Collaboration, *Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV*, JINST **10** (2015) P08010, doi:10.1088/1748-0221/10/08/P08010, arXiv:1502.02702.
- [151] The 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.

- [152] The CMS Collaboration, *Muon Identification and Isolation efficiency on full 2016 dataset*, tech. rep., Mar, 2017.
- [153] M. Della Negra, L. Foà, A. Hervé, and A. Ball, *CMS physics Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 2006.
- [154] W. Adam, R. Frühwirth, A. Strandlie, and T. Todor, *Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC*, Tech. Rep. CMS-NOTE-2005-001, CERN, Geneva, 2005.
- [155] The CMS Collaboration, *Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV*, JINST **10** (2015) P06005, doi:10.1088/1748-0221/10/06/P06005, arXiv:1502.02701.
- [156] The CMS EGM group, *Cut Based Electron ID for Run 2*, (last consulted: 27 April 2017). <https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedElectronIdentificationRun2>
- [157] I. Kravchenko and R. Kamalieddin, *Effective area correction factors for electron isolation*, (last consulted: 27 April 2017). [https://github.com/ikrav/cmssw/blob/egm\\_id\\_80X\\_v1/RecoEgamma/ElectronIdentification/data/Summer16/effAreaElectrons\\_cone03\\_pfNeuHadronsAndPhotons\\_80X.txt](https://github.com/ikrav/cmssw/blob/egm_id_80X_v1/RecoEgamma/ElectronIdentification/data/Summer16/effAreaElectrons_cone03_pfNeuHadronsAndPhotons_80X.txt)
- [158] 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.
- [159] G. P. Salam, *Towards Jetography*, Eur. Phys. J. C **67** (2010) 637–686, doi:10.1140/epjc/s10052-010-1314-6, arXiv:0906.1833.
- [160] The CMS Collaboration, V. Khachatryan *et al.*, *Jet energy scale and resolution in the CMS experiment in  $pp$  collisions at 8 TeV*, JINST **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.
- [161] The CMS Collaboration, *Jet energy scale and resolution performances with 13 TeV data*, Jun, 2016. <http://cds.cern.ch/record/2160347>
- [162] The CMS JME group, *Jet Identification*, (last consulted: 27 April 2017). <https://twiki.cern.ch/twiki/bin/view/CMS/JetID>
- [163] The CMS Collaboration, *Identification of  $b$ -quark jets with the CMS experiment*, JINST **8** (2012) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:1211.4462. arXiv:hep-ex/1211.4462. CMS-BTV-12-001. CERN-PH-EP-2012-262.
- [164] The CMS Collaboration, *Identification of heavy-flavour jets with the CMS detector in  $pp$  collisions at 13 TeV*, arXiv:1712.07158. Submitted to JINST.
- [165] The CMS MUO group, *Reference muon ID, isolation and trigger efficiencies for Run 2*, (last consulted: 30 August 2017). <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEfsRun2>

- [166] T. Sjöstrand *et al.*, *Pythia 8.2 Online Manual: Particle Properties*, (last consulted: 17 July 2017). <http://home.thep.lu.se/~torbjorn/pythia82html/ParticleProperties.html>
- [167] J. D’Hondt, S. Lowette, O. L. Buchmüller, S. Cucciarelli, F.-P. Schilling, M. Spiropulu, S. Paktinat-Mehdiabadi, D. Benedetti, and L. Pape, *Fitting of Event Topologies with External Kinematic Constraints in CMS*, Tech. Rep. CMS-NOTE-2006-023, CERN, Geneva, Jan, 2006.
- [168] P. Skands, S. Carrazza, and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 Tune*, Eur. Phys. J. C **74** (2014), no. 8, 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.
- [169] M. G. Bowler,  *$e^+e^-$  Production of heavy quarks in the string model*, Zeitschrift für Physik C Particles and Fields **11** (Jun, 1981) 169–174, doi:10.1007/BF01574001.
- [170] The ALEPH Collaboration, *Study of the fragmentation of  $b$  quarks into  $B$  mesons at the  $Z$  peak*, Phys. Lett. B **512** (2001), no. 1-2, 30 – 48, doi:https://doi.org/10.1016/S0370-2693(01)00690-6.
- [171] The DELPHI Collaboration, *A study of the  $b$ -quark fragmentation function with the DELPHI detector at LEP I and an averaged distribution obtained at the  $Z$  Pole*, Eur. Phys. J. C **71** (Feb, 2011) 1557, doi:10.1140/epjc/s10052-011-1557-x.
- [172] C. Peterson, D. Schlatter, I. Schmitt, and P. M. Zerwas, *Scaling violations in inclusive  $e^+e^-$  annihilation spectra*, Phys. Rev. D **27** (Jan, 1983) 105–111, doi:10.1103/PhysRevD.27.105.
- [173] M. Czakon, P. Fiedler, D. Heymes, and A. Mitov, *NNLO QCD predictions for fully-differential top-quark pair production at the Tevatron*, JHEP **05** (2016) 034, doi:10.1007/JHEP05(2016)034, arXiv:1601.05375.
- [174] The CMS Collaboration, *Measurement of differential cross sections for top quark pair production using the lepton + jets final state in proton-proton collisions at 13 TeV*, Phys. Rev. D **95** (May, 2017) 092001, doi:10.1103/PhysRevD.95.092001.
- [175] E. Todesco and J. Wenninger, *Large Hadron Collider momentum calibration and accuracy*, Phys. Rev. Accel. Beams **20** (Aug, 2017) 081003, doi:10.1103/PhysRevAccelBeams.20.081003.
- [176] The CMS Collaboration, *CMS Luminosity Measurements for the 2016 Data Taking Period*, Tech. Rep. CMS-PAS-LUM-17-001, CERN, Geneva, 2017.
- [177] The ATLAS Collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at  $\sqrt{s} = 13$  TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (2016), no. 18, 182002, doi:10.1103/PhysRevLett.117.182002, arXiv:1606.02625.
- [178] The CMS Collaboration, *Measurement of the inelastic proton-proton cross section at  $\sqrt{s} = 13$  TeV*, Tech. Rep. CMS-PAS-FSQ-15-005, CERN, Geneva, 2016.

- 
- [179] T. Behnke, J. E. Brau, B. Foster, J. Fuster, M. Harrison, J. M. Paterson, M. Peskin, M. Stanitzki, N. Walker, and H. Yamamoto, *The International Linear Collider Technical Design Report - Volume 1: Executive Summary*, arXiv:1306.6327.
- [180] M. Aicheler, P. Burrows, M. Draper, T. Garvey, P. Lebrun, K. Peach, N. Phinney, H. Schmickler, D. Schulte, and N. Toge, *A Multi-TeV Linear Collider Based on CLIC Technology*, doi:10.5170/CERN-2012-007.
- [181] TLEP Design Study Working Group Collaboration, M. Bicer *et al.*, *First Look at the Physics Case of TLEP*, JHEP **01** (2014) 164, doi:10.1007/JHEP01(2014)164, arXiv:1308.6176.
- [182] M. Martinez and R. Miquel, *Multiparameter fits to the  $t$  anti- $t$  threshold observables at a future  $e^+ e^-$  linear collider*, Eur. Phys. J. C **27** (2003) 49–55, doi:10.1140/epjc/s2002-01094-1, arXiv:hep-ph/0207315.
- [183] ECFA/DESY LC Physics Working Group Collaboration, J. A. Aguilar-Saavedra *et al.*, *TESLA: The Superconducting electron positron linear collider with an integrated x-ray laser laboratory. Technical design report. Part 3. Physics at an  $e^+ e^-$  linear collider*, arXiv:hep-ph/0106315.
- [184] H. Baer, T. Barklow, K. Fujii, Y. Gao, A. Hoang, S. Kanemura, J. List, H. E. Logan, A. Nomerotski, M. Perelstein, *et al.*, *The International Linear Collider Technical Design Report - Volume 2: Physics*, arXiv:1306.6352.



## Breakdown of Systematic Uncertainties

---

# A

Contrary to Tables 6.1 and 6.2, where uncertainties within the same categories have been added quadratically or the envelope of the uncertainties is shown, this appendix presents an overview of all components of the systematic uncertainties. The values in Tables A.1 to A.4 represent the systematic uncertainties for the variable  $s = \Gamma_t/\Gamma_{t,\text{gen}}$ , so they should be multiplied by  $\Gamma_{t,\text{gen}} = 1.31$  GeV in order to calculate the systematic uncertainty on the top quark decay width. Tables A.1–A.2 contain respectively the experimental and theoretical uncertainties on the hadronic measurement and Tables A.3–A.4 those on the combined measurement.

Similar as in Section 6.3.3, each value represents the shift  $\hat{\Gamma}_{t,\text{sys}} - \hat{\Gamma}_{t,\text{nom}}$  induced by the systematic effect. If a certain systematic uncertainty has a variation that does not correspond to a  $\pm 1\sigma$  variation, such as the effect of applying a certain model or not, or if both values are shifted to the same side, only one value is presented.

On the bottom line all shifts presented in the table are quadratically added. For both measurements, the sum of the theoretical uncertainties is at least twice as large as the sum of the experimental uncertainties.

Table A.1: Effect of the experimental systematic uncertainties for the hadronic measurement. The values represent an absolute shift compared to the nominal value presented at the top of the table due to a  $\pm 1\sigma$  variation.

Nominal value	$0.976 \pm 0.086$	
Systematic	$+1\sigma$	$-1\sigma$
<i>Calibration curve</i>		
Slope	0.015	-0.015
Constant	0.040	-0.040
JES	0.150	-0.020
JER	0.127	-0.135
Pileup SFs	0.086	-0.082
b tagging SFs	0.042	-0.044
<i>Lepton SFs</i>		
Id	0.001	-0.001
Isolation	0.000	0.000
Trigger	0.002	-0.002
Tracking	0.000	0.000
<i>Cross section variations</i>		
ST t channel	0.013	-0.014
ST tW channel	0.026	-0.027
Other	0.035	-0.035
Total	0.227	-0.176

Table A.2: Effect of theoretical systematic uncertainties for the hadronic measurement. The values represent an absolute shift compared to the nominal value presented at the top of the table. When the systematic uncertainty has an on/off effect instead of a  $\pm 1\sigma$  variation or when both values are shifted to the same side, only one value is presented.

Nominal value	$0.976 \pm 0.086$			
Systematic	$+1\sigma$		$-1\sigma$	
	Shift	Unc	Shift	Unc
ISR	-0.018	0.151	-0.309	0.124
FSR	0.340	0.165	-0.257	0.116
<i>ME scales</i>				
$\mu_R = 1, \mu_F = 2$	0.004			
$\mu_R = 1, \mu_F = 0.5$			-0.007	
$\mu_R = 2, \mu_F = 1$			-0.084	
$\mu_R = 2, \mu_F = 2$			-0.076	
$\mu_R = 0.5, \mu_F = 1$	0.122			
$\mu_R = 0.5, \mu_F = 0.5$	0.124			
ME-PS matching	0.208	0.159	-0.236	0.142
<i>PDF</i>				
Replicas	0.047		-0.047	
$\alpha_s$	0.005		-0.007	
<i>b fragmentation</i>				
Bowler-Lund	0.018			
Peterson	0.102			
b semileptonic BR	0.043		-0.063	
Underlying event	0.047	0.151	-0.050	0.151
<i>Colour reconnection</i>				
MPI (ERD)	0.035	0.149		
QCD-based (ERD)	0.341	0.160		
Gluon move	0.197	0.159		
Gluon move (ERD)	0.008	0.150		
Top mass	0.244			
Top $p_T$ reweighting	0.118			
Total	0.622		-0.424	

Table A.3: Effect of the experimental systematic uncertainties for the combined measurement. The values represent an absolute shift compared to the nominal value presented at the top of the table due to a  $\pm 1\sigma$  variation.

Nominal value	$0.960 \pm 0.073$	
Systematic	$+1\sigma$	$-1\sigma$
<i>Calibration curve</i>		
Slope	0.012	-0.012
Constant	0.032	-0.032
JES	0.169	-0.126
JER	0.100	-0.116
Pileup SFs	0.064	-0.064
b tagging SFs	0.032	-0.036
<i>Lepton SFs</i>		
Id	0.008	-0.008
Isolation	0.000	0.000
Trigger	0.006	-0.006
Tracking	0.000	0.000
<i>Cross section variations</i>		
ST t channel	0.010	-0.010
ST tW channel	0.019	-0.023
Other	0.030	-0.033
Total	0.215	-0.194

Table A.4: Effect of theoretical systematic uncertainties for the combined measurement. The values represent an absolute shift compared to the nominal value presented at the top of the table. When the systematic uncertainty has an on/off effect instead of a  $\pm 1\sigma$  variation or when both values are shifted to the same side, only one value is presented.

Nominal value	$0.960 \pm 0.073$			
Systematic	$+1\sigma$		$-1\sigma$	
	Shift	Unc	Shift	Unc
ISR	-0.087	0.122	-0.206	0.107
FSR	0.165	0.145	-0.175	0.100
<i>ME scales</i>				
$\mu_R = 1, \mu_F = 2$			-0.002	
$\mu_R = 1, \mu_F = 0.5$	0.000			
$\mu_R = 2, \mu_F = 1$			-0.073	
$\mu_R = 2, \mu_F = 2$			-0.069	
$\mu_R = 0.5, \mu_F = 1$	0.093			
$\mu_R = 0.5, \mu_F = 0.5$	0.102			
ME-PS matching	0.246	0.138	-0.210	0.110
<i>PDF</i>				
Replicas	0.055		-0.055	
$\alpha_s$	0.004		-0.005	
<i>b fragmentation</i>				
Bowler-Lund	0.034		-0.043	
Peterson	0.021			
b semileptonic BR	0.035		-0.067	
Underlying event	-0.119	0.122	-0.120	0.125
<i>Colour reconnection</i>				
MPI (ERD)	0.015	0.144		
QCD-based (ERD)	0.234	0.140		
Gluon move	0.189	0.136		
Gluon move (ERD)			-0.008	0.144
Top mass	0.205			
Top $p_T$ reweighting	0.098			
Total	0.447		-0.364	



## Effect of the Systematic Uncertainties on the Sensitive Variable Distributions

---

# B

The systematic uncertainties influence the distributions of the sensitive variables. Two effects can be differentiated, i.e. differences in the amount of events in each bin and differences in the shape of the distributions.

The largest systematic uncertainties affecting the measurements presented in this thesis are initial and final state radiation, the jet energy scale, the top quark mass, the ME-PS matching and the underlying event and colour reconnection.

The effect of the above-mentioned systematic uncertainties on the distributions of the sensitive variables is visualised in Figures B.1 to B.4. Whereas ISR, FSR, and the ME-PS matching mostly have a rate effect, the shape of the distributions affected by JES and the top quark mass are profoundly different. Both seem to have a similar effect, i.e. higher masses/energy scales shift the peak of the distribution to larger values and vice versa. This is a logical consequence of the definitions of  $m_r$  and  $m_{lb,r}$ , since the uncertainties increase/decrease the value of the numerator, while the denominator remains constant.

For colour reconnection, the largest differences are observed for the gluon move (ERD) model and, to a lesser extent, for the QCD (ERD) model. The underlying event, on the other hand, only has a very slight influence on the distributions. This supports the idea that this systematic uncertainty would contribute much less if its statistical uncertainty can be reduced.

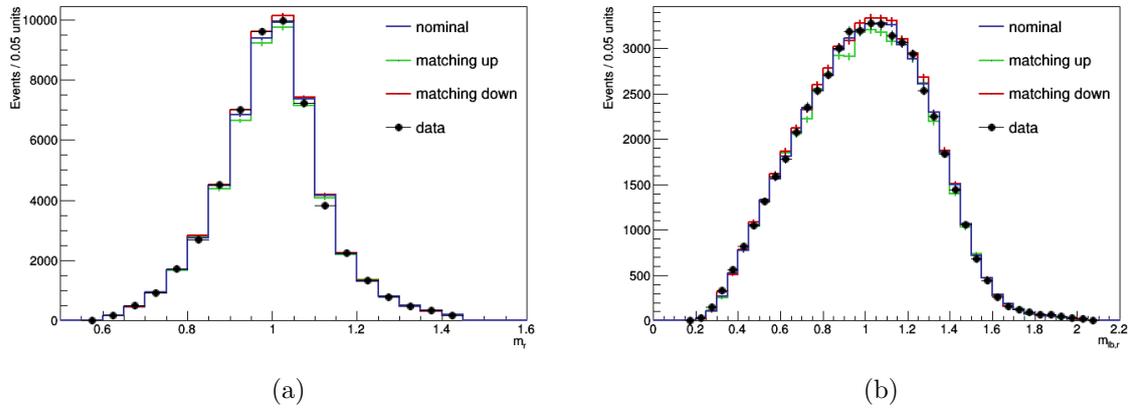


Figure B.1: Effect of the uncertainty due to ME-PS matching on the distribution of (a)  $m_r$  and (b)  $m_{lb,r}$ .

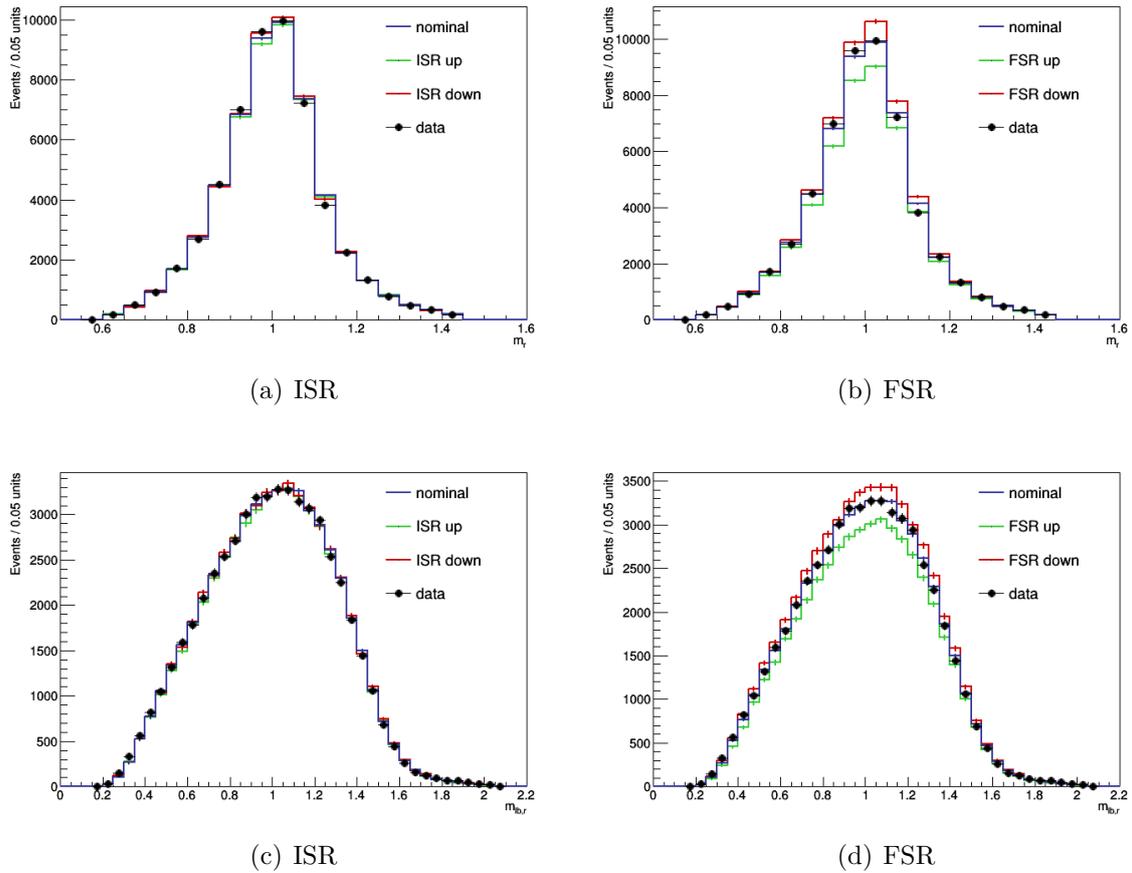
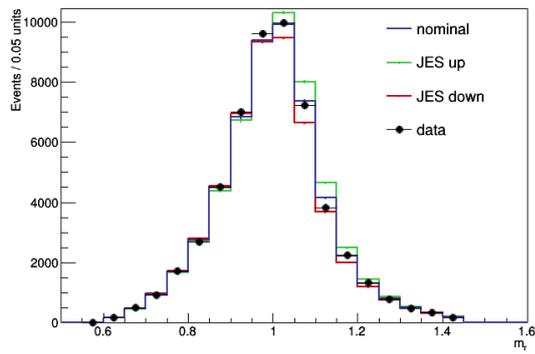
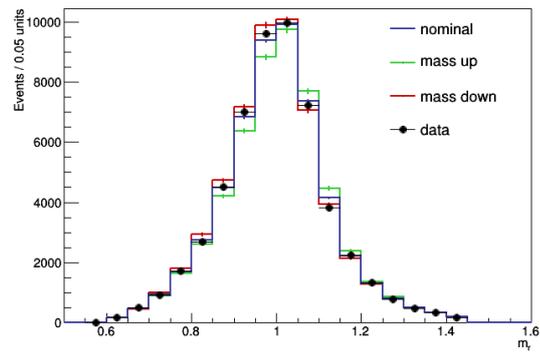


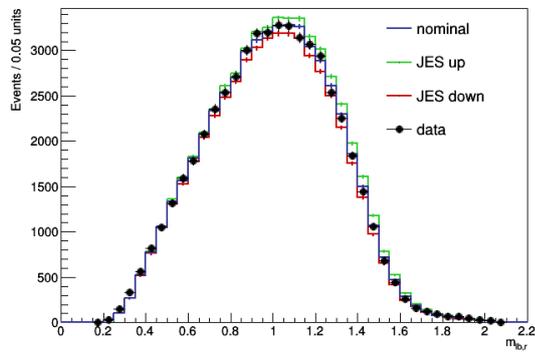
Figure B.2: Effect of the uncertainty due to ISR (left) and FSR (right) on the distributions of  $m_r$  (top) and  $m_{lb,r}$  (bottom).



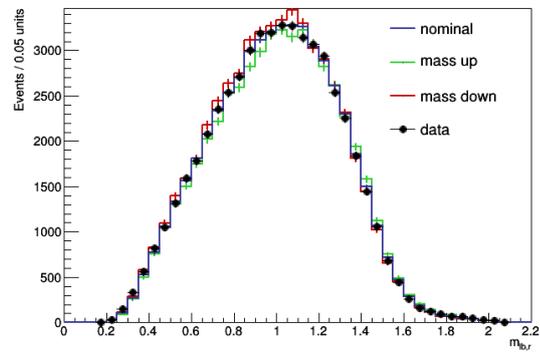
(a) JES



(b) Mass

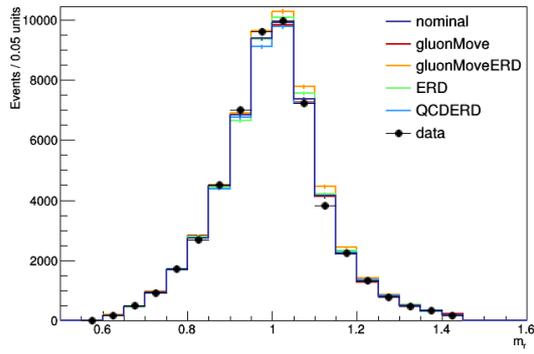


(c) JES

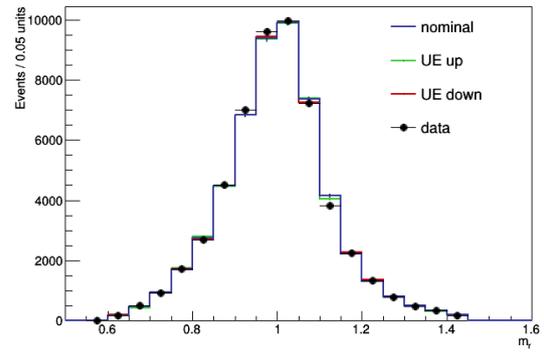


(d) Mass

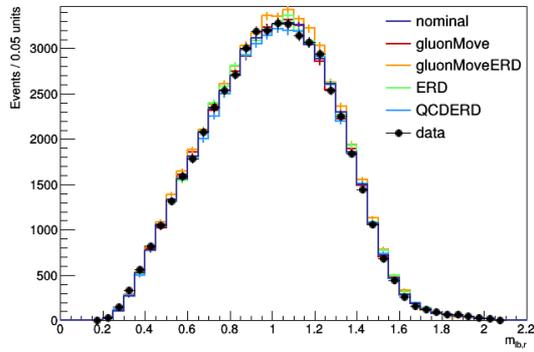
Figure B.3: Effect of the uncertainty due to JES (left) and the top quark mass (right) on the distributions of  $m_r$  (top) and  $m_{lb,r}$  (bottom).



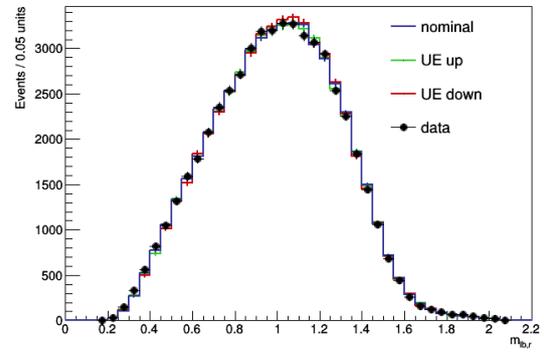
(a) CR



(b) UE



(c) CR



(d) UE

Figure B.4: Effect of the uncertainty due to colour reconnection (left) and the underlying event (right) on the distributions of  $m_r$  (top) and  $m_{lb,r}$  (bottom).

## Author's Contributions

---

This thesis presents the first direct measurement of the top quark decay width from the CMS experiment using top quark pair events in the lepton+jets decay channel, for which I am the sole contributor. I conducted all parts of the analysis described in Chapters 5–6, from the selection and reconstruction procedure to the final result. Many of the techniques developed for my analysis were also used in the analysis that measured the top quark decay width in the dilepton channel, which is published as *Bounding the top quark width using final states with two charged leptons and two jets at  $\sqrt{s} = 13$  TeV, CMS Collaboration, 2016, CMS-PAS-TOP-16-019.*

An example is the reweighting procedure described in Section 3.3. Further, I combined the sensitive variable in that analysis with another one that is only available in the lepton+jets channel. The analysis described in this thesis has been endorsed by the CMS collaboration.

In addition, I performed a phenomenological study to compare the performance of the novel matrix element method (MEM) to a standard multivariate analysis technique (MVA) in the search for a top squark decaying to a top quark and a neutralino, a process that is predicted by some supersymmetry models. Whereas the results are comparable in the dilepton channel, the MVA has an overall better performance in the other decay channels. Seeing as the MEM is very time-consuming and CPU intensive, I concluded that using an MVA technique is the better option for new physics searches. This result influenced the plans of other members of my research group to a great degree.

Based on my expertise with the MEM, I wanted to investigate whether this method had a positive influence on a measurement of the top quark decay width. Due to technical issues to treat the value of the top quark decay width as a free parameter during the calculation of the matrix element in the `MadGraph/MadEvent` software, the decision was made not to continue this train of thought.

In the course of my PhD I was heavily involved in the commissioning and operation of the silicon strip tracker, which was operated at  $-15^{\circ}\text{C}$  for the first time during the LHC run-2. As a consequence, all operational parameters needed to be recalibrated in order to reflect the new conditions. In addition, I helped to install new temperature and humidity sensors, placed insulating foam so as to reduce the amount of air from whence moisture can condensate, and I helped sealing the tracker volume and service channels with vapour barriers. The latter proved to be very functional when a water leak occurred and no tracker equipment got damaged despite the puddle of water on top of the seal. Further, I was involved in commissioning the tracker dry gas system,

performing leak and pressure tests.

After the commissioning, I regularly took up detector expert on-call shifts. During these shifts I coordinated all daily activities concerning the silicon strip tracker, monitored the conditions of the detector to ensure safe operation and was responsible for the quality of the data taken. In intervals between data-taking, I assisted in the maintenance of the detector, such as the replacement of faulty electronic components, in both the experimental and the service cavern. Furthermore, I contributed to the upkeep of the data quality monitoring (DQM) software, which also got me involved in the DQM-on-call shifts. Monitoring the quality of the data when it is produced enables to react quickly on potential problems, thus ensuring good tracking information, which is vital for most analyses.

Additionally, I helped calibrating the detector control unit (DCU) chips, which monitor the temperature, voltages, and leakage current of the sensors. The latter can be used as a measure for the amount of radiation damage the silicon sensor has endured. When properly calibrated, DCUs can thus give an estimation of the sustained damage and further life expectancy of the sensors.

Given the timescale needed to plan large projects, preparations for a new tracking detector are already being made. The plan is to include track information into the level-1 trigger for the high-luminosity upgrade of the LHC. This so-called track trigger depends on a dedicated geometry that enables to quickly differentiate between high- and low- $p_T$  tracks. I contributed in the validation of this procedure, making plots that give an overview of the properties of the track-trigger objects. These are used to determine the efficiency of several proposed geometries and reconstruction algorithms and some of the plots can be re-used for the data quality monitoring of the new detector. In this way I contributed to the technical design report for the new ‘phase-2 outer tracker’, *The Phase-2 Upgrade of the CMS Tracker, CMS Collaboration, 2017, CERN-LHCC-2017-009*.

Besides participating in numerous reviews of scientific papers, I have also been involved with outreach. When the 2013 Nobel prize of physics was awarded to F. Englert and P. Higgs for the discovery of the H boson, I explained in layman’s terms what this discovery entailed and why it was important in the television programme Terzake broadcasted by the VRT (national channel comparable to the BBC in the UK). This report was awarded the diversity prize in the category ‘News’. Further, I have been a guide at the exhibition ‘*CERN: 60 years of Science for Peace*’ in Brussels in 2015 and last year, I was a member of the jury for the scientific prize of the project ‘*90 degrees South: Your Experiment at the South Pole*’. For this project, advanced primary and secondary school children designed experiments that would give different results in Belgium and on the South Pole and the jury had to decide which experiment was the most successful.

## List of Abbreviations

---

ALICE	A Large Ion Collider Experiment, one of the experiments at the LHC
ATLAS	A Toroidal LHC ApparatuS, one of the experiments at the LHC
BSM	Beyond the standard model
BW	Breit–Wigner, statistical distribution that describes the distribution of (amongst others) the mass of a particle
CDF	Collider Detector at Fermilab, one of the experiments at the Tevatron
CERN	European Organisation for Nuclear Research, located at the Franco-Swiss border near Geneva. (Lit. Conseil Européen pour la Recherche Nucléaire)
CHS	Charged hadron subtraction
CLIC	Compact Linear Collider, potential future $e^+e^-$ accelerator
CM	Correctly matched events
CMS	Compact Muon Solenoid, one of the experiments at the LHC
CR	Colour reconnection
CSC	Cathode Strip Chamber, type of muon detector at the CMS experiment
CSVv2	Combined Secondary Vertex v2 algorithm
CT-PPS	CMS-TOTEM Precision Proton Synchrotron, joint collaboration of CMS and TOTEM
DGLAP	Equations describing the probability that a parton will branch. These were developed in parallel by Dokshitzer, Gribov and Lipatov, and Altarelli and Parisi.
D0	One of the experiments at the Tevatron, located in the D0 region
DT	Drift Tube, type of muon detector at the CMS experiment

---

EB	ECAL Barrel
ECAL	Electromagnetic calorimeter
EE	ECAL Endcap
ERD	Early resonance decays
ES	ECAL preshower
FCC-ee	Future Circular electron-positron Collider, potential future $e^+e^-$ accelerator, also known as TLEP
FSR	Final state radiation
FWHM	Full-width-half-maximum, width of a distribution at the function value that is half the value of its peak. This is equal to about 2.35 standard deviations.
GSF	Gaussian sum filter
HB	HCAL Barrel
HCAL	Hadron calorimeter
HE	HCAL Endcap
HF	HCAL Forward
HL-LHC	High-Luminosity LHC
HLT	High-level trigger
HO	HCAL Outer
ILC	International Linear Collider, potential future $e^+e^-$ accelerator
IP	Impact parameter, transverse distance to the $z$ axis
ISR	Initial state radiation
JEC	Jet energy corrections
JER	Jet energy resolution
JES	Jet energy scale
KF	Kinematic fit

---

L1	Level-1 trigger
LEIR	Low Energy Ion Ring, accelerator located at CERN
LHC	Large Hadron Collider, accelerator located at CERN
LHCb	Large Hadron Collider beauty, one of the experiments at the LHC
LHCf	Large Hadron Collider forward, one of the experiments at the LHC, located near the ATLAS experiment.
LO	Leading order
LS1	Long Shutdown 1, timeframe between the LHC Run 1 (2010-2013) and Run 2 (2015-ongoing) when the accelerator was not operated and upgrades to both accelerator and experiments were performed
ME	Matrix element
ML	Maximum likelihood method
MLM	ME-PS matching scheme developed by M. L. Mangano
MPI	Multiparton interaction
MVA	Multivariate analysis
NLO	Next-to-leading order
NNLO	Next-to-next-to-leading order
PD	Pixel detector
PDF	Parton distribution function
PF	Particle flow, reconstruction algorithm
PS	Parton shower
PS	Proton Synchrotron, accelerator located at CERN
PU	Pileup
PV	Primary vertex
QCD	Quantum chromodynamics
QFT	Quantum field theory

---

RC	Random cone algorithm
RF	Radiofrequency
RMS	Root-mean-square
RPC	Resistive Plate Chamber, type of muon detector at the CMS experiment
SL	Superlayer, collection of four stacked layers of DT cells
SM	Standard model of particle physics
SPS	Super Proton Synchrotron, accelerator located at CERN
ST	Single top
$SU(n)$	Special unitary group containing all matrices of dimension $n$
SV	Secondary vertex
TEC	Tracker Endcap disc
TESLA	TeV Electron Superconducting Linear Accelerator, potential future $e^+e^-$ accelerator project that got incorporated into the ILC
Tevatron	Circular particle accelerator located at Fermilab, near Chicago, USA, that accelerates protons and antiprotons to the TeV energy scale
TIB	Tracker Inner Barrel
TID	Tracker Inner Disc
TLEP	Triple-Large Electron-Positron collider, potential future $e^+e^-$ accelerator, also known as FCC-ee
TOB	Tracker Outer Barrel
TOTEM	TOTAL Elastic and diffractive cross section Measurement, one of the experiments at the LHC, located near the CMS experiment
$U(n)$	Unitary group containing all matrices of dimension $n$
UE	Underlying event
UM	Unmatched events
WM	Wrongly matched events
WP	Working point

## Glossary

---

Bremsstrahlung	Electromagnetic radiation that is produced when a charged particle, e.g. an electron, is accelerated or decelerated by the presence of a strong electric field, e.g. that of a nucleus.
Electromagnetic shower	or cascade, produced by subsequent photon radiation (through Bremsstrahlung) and pair production when an electron or photon crosses matter. Each step generates particles with slightly lower energies. The longitudinal size of the shower depends on the radiation length of the material and its transverse size is described by the Moliere radius.
Interaction length	Average distance a hadron can cross in a material before a nuclear interaction occurs. Symbol: $\lambda_I$
Moliere radius	Transverse width of an electromagnetic shower. The Moliere radius $R_0$ is proportional to the radiation length.
Multiple scattering	Multiple collisions of a charged particle with the nuclei of the material it is crossing. Each collision has an influence on the direction the particle is travelling. If the scattering angle is small, the difference between the incident and outgoing angle can be described by a Gaussian function.
Pair production	The transformation of a photon into an electron-positron pair. In order for this process to occur, the energy of the photon must be at least twice the rest energy of the electron, i.e. $2 \times 511$ keV.
Radiation length	Length of material an electron has to cross to lose $1/e \simeq 0.368$ of its original energy. Symbol: $X_0$



## Acknowledgements

---

A thesis does not write itself (if you don't believe me, ask the author) and considering the many hours analysing and writing, it risks being a very lonely business. Luckily, I've had the pleasure to meet some amazing people throughout these years who prevented me becoming a complete recluse. You have encouraged me and pushed me forwards and for this you deserve my eternal gratitude.

First of all, I would like to thank my supervisor Petra Van Mulders, who was always available for a topical (or non-topical) chat. The problems I encountered drove us both to despair sometimes and we've been sent back to the drawing board many times, but in the end it paid off. The very fact you are reading this today is proof of that. Starting a PhD is not a trivial thing. Besides an enthusiastic student, it takes someone to believe in you and give you the opportunity, for which I would like to thank Jorgen D'Hondt. Both of you had a great impact on the final result. I enjoyed discussing my analysis with you and your comments greatly improved the quality of this manuscript. Further, I am grateful to the members of my jury for the pleasant discourse and the interesting questions during my private defence.

The last couple of years I spent most of my waking hours at the IIHE, where I encountered many exceptional people. A special word of thanks has to go to Marleen. Not only did you guide me through the never-ending maze that is bureaucracy, you were always there to help in whichever way possible. In addition, I would like to thank Olivier, Abdel, Abdel, Samir, Romain, Stéphane, Shkelzen, and Adriano for their technical support (even in the weekends!) and listening to my complaints about people spamming the m-machines.

Undoubtedly, my office mates have had the main influence on my general mood and I am grateful to have been able to spend so much time with them. Thank you, Kevin (the time lord), Isis (the master of plots), Isabelle (the non-crazy cat lady), Shimaa (the caring), Doug (the master of puns), Simon (the bearer of baked goods), and by extension Joni (the maker of baked goods—and ice cream), Petra (the leader of rocky island), Quentin (the background soundbox), Denys (the incredible snowman), Kirill (Beatle no. 1), Seth (the machine educator), Jarne (the invisible), Dom (the silent), Leonidas (the loud), Giannis (/Gianna), and Emil (the new kid on the block). You are truly extraordinary and brightened up the many days I spent in our gigantic office. Also Annik, Gerrit, James, Natalie, Tae Jeong, Nadir, Grégory, Alexis, Stijn, and Michael had the pleasure to feel the effects global\* warming very close up (\*local, very local in fact). Karen and Dries, you shared a wall with our office, but unfortunately for us

we couldn't share your sun blinds. Luckily no energy is wasted, as the radishes and pumpkins growing in our 'little' greenhouse are proving at this very moment. (Giving a whole new meaning to a Root-based analysis.)

Kevin and Isis, we spent almost a third of our lifetime together as students and PhD candidates. And what a journey it has been! Both in the literal (from Brussels to Switzerland and back again) and the figurative sense. Isis, driving to CERN with you always felt like a holiday and I will never forget limboing through the Surkov frame with you. Also you, Isabelle, made my stays at Point 5 memorable.

In addition, I would like to thank the 'Point 5 regulars', Christian, Ferdinando, Francesco, Derek, Nicola, Stefanos, and I am very indebted to Erik, who let me explore the tracker up-close and was always willing to feed me information from his bottomless source of knowledge. I truly became a 'tracker girl' and I still feel an almost irresistible urge to pick up the phone when someone has the same ringtone as the DOC phone.

Also without you, Rebeca, my trips to CERN would never have been the same. (What would the life of a tracker DOC be without a kind shift leader and an amazing RFM?) Thank you for your valuable input to my analysis and for the times you guided me around CMS. Hopefully, we can repeat this some day.

On a different note, thank you, Isabelle and Doug, for liking tea as much as I do and introducing a paradigm shift from 'coffee break' to 'tea break'. Nevertheless, coffee drinkers, Isis and Simon, are indispensable to make them truly delightful.

My greatest thanks and my deepest apologies go to my family and friends, who I have shamelessly ignored for too long a time. Thanks for believing in me and encouraging me to pursue my dreams. Tom, your support during these past few months means more to me than I can put into words. Without you this journey would not nearly be as enjoyable as it is now.

Last, but not least, I would like to spare a thought for the protons who have given their life so that I could do this research. Your contribution will never be forgotten.