Measurement of the Top Quark Pair Production Cross Section in the Electron and Jets Channel at $\sqrt{s} = 7\,\text{TeV}$ with the CMS Detector at the LHC

Benjamin Klein

Dissertation

Faculty of Sciences
Ghent University

Promoter: Prof. Dr. Martin Grünewald
Department of Physics and Astronomy

2012

Ní mar a shíltear a bhítear.

(Irish proverb)

Dutch Summary / Nederlandstalige samenvatting

Sinds meer dan tweeduizend jaren streeft de mens ernaar de natuur en het universum te verstaan. In de laatste eeuw was er een grote vooruitgang in het inzicht in het werken van de natuur. Kwantumfysica werd ontworpen bij de aanvang van de twintigste eeuw en sinds het midden van de twintigste eeuw is deeltjesfysica een actief gebied van onderzoek.

Deeltjesfysica is het onderzoek van de fundamentele bouwstenen van materie en de krachten tussen deze bouwstenen. Het Standaard Model van de deeltjesfysica is een kwantumveldentheorie en beschrijft twaalf elementaire deeltjes, zes leptonen en zes quarks. Alle bekende materie in het universum bestaat uit deze elementaire deeltjes. Het Standaard Model beschrijft ook drie wisselwerkingen, de sterke wisselwerking, de zwakke wisselwerking en de elektromagnetische wisselwerking.

Van de twaalf elementaire deeltjes is de *top quark* het zwaarste deeltje met een massa van $173.5 \pm 0.6 \pm 0.8\,\mathrm{GeV}/c^2$. Het is meer dan dertig keer zwaarder dan de tweede zwaarste quark, de *bottom quark*. Deze uitzonderlijke eigenschap maakt het interessant voor studies en voor precisietests van het Standaard Model. De meting van de top quark-paar productie werkzame doorsnede is één van deze metingen en is het thema van dit proefschrift. Top quark gebeurtenissen worden bestudeerd waarin het top quark-paar in een elektron en vier jets desintegreren.

De Large Hadron Collider (LHC), een deeltjesversneller van de Europese organisatie voor kernonderzoek vlakbij Genève, Zwitserland, werd in 2008 in gebruik genomen. Met een totale energie van $\sqrt{s} = 7\,\mathrm{TeV}$ botsen protonen in vier botsingspunten en veroorzaken wisselwerkingen tussen de elementaire bouwstenen van de protonen. In deze botsingen worden onder andere ook top quarks geproduceerd. Grote deeltjesdetectoren meten deze gebeurtenissen en nemen hen op voor analyse. Eén van deze detectoren is de *Compact Muon Solenoid* (CMS) detector die gebruikt is om de gegevens voor deze analyse op te tekenen.

De gemeten werkzame doorsnede voor top quark productie is omgekeerd evenredig met de selectie-efficiëntie van top quark gebeurtenissen. Daarom is het belangrijk de selectie-efficiëntie heel precies te meten. Er zijn twee bronnen van efficiënties die in detail bestudeerd worden in dit proefschrift, de efficiëntie van het *trigger* systeem dat gebeurtenissen registreert tijdens de data-acquisitie, en de efficiëntie van de selectie van elektronen.

De meting van de werkzame doorsnede van top quark productie wordt doorgevoerd met een template fit techniek. De distributie van de *M3 massa*, de invariante massa van het drie-jet-systeem in een gebeurtenis met de hoogste transverse impuls,

wordt voor top quark en achtergrond gebeurtenissen afzonderlijk uit de simulatie geëxtraheerd of in de data gemeten. Het aantal top quark gebeurtenissen in de data wordt dan gemeten met een *maximum likelihood fit* van deze *templates* aan de data en dus ook de werkzame doorsnede voor top quark productie met de relatie $\sigma_{t\bar{t}} = N_{t\bar{t}}/(\epsilon \cdot \int \mathrm{d}t\mathcal{L})$. $\sigma_{t\bar{t}}$ is de werkzame doorsnede van de top quark-paar productie, $N_{t\bar{t}}$ is het aantal top quark-paar gebeurtenissen, ϵ is de selectie efficiëntie en $\int \mathrm{d}t\mathcal{L}$ is de geïntegreerde luminositeit, een maat voor de hoeveelheid data. Het finale resultaat van de meting is

$$\sigma_{t\bar{t}} = 155.5 \pm 3.7\,(\text{stat.})\,^{+12.6}_{-14.8}\,(\text{syst.}) \pm 3.4\,(\text{lumi.})\,\text{pb}.$$

Ook een combinatie van de meting van de werkzame doorsnede voor top quark productie met de signatuur van een elektron en jets en de signatuur van een muon en jets wordt doorgevoerd. De massa van de top quark wordt met een *joined likelihood* methode van de gemeten werkzame doorsnede geëxtraheerd.

Alle resultaten zijn compatibel met andere metingen van de CMS and ATLAS collaboratie en ook met de voorspellingen van het Standaard Model van de deeltjesfysica.

Contents

I. Theory and Experiment **3**

1. The Standard Model of Particle Physics **5**
1.1. Field Theoretical Approach 5
1.2. Quarks and Leptons 6
1.3. Interactions between Particles 6
1.3.1. The Electromagnetic and Weak Interactions 7
1.3.2. The Strong Interaction 7
1.4. Free Parameters of the Standard Model 8
1.5. Unit Conventions 8
1.6. Measurement of the Interactions of Fundamental Particles 8
1.6.1. Proton-Proton Collisions 9
1.7. Top Quark Physics 11
1.7.1. Top Quark Production 12
1.7.2. Decay of Top Quark Pairs 14

2. The LHC and the CMS Experiment **17**
2.1. The Large Hadron Collider 17
2.1.1. The Experiments at the LHC 19
2.2. The Compact Muon Solenoid Experiment 20
2.2.1. The CMS Coordinate System 20
2.2.2. The Superconducting Solenoid Magnet 22
2.2.3. The Inner Tracking System 22
2.2.4. The Electromagnetic Calorimeter 22
2.2.5. The Hadronic Calorimeter 27
2.2.6. The Muon System 29
2.2.7. The Trigger System 31
2.3. Online Object Reconstruction 32
2.3.1. General Setup of the CMS Trigger System 32
2.3.2. Electron and Jet Reconstruction in the CMS Trigger System 34
2.4. Physics Objects Reconstruction within the Particle Flow Model . . . 36
2.4.1. Primary Vertex Reconstruction in CMS 37
2.4.2. Tracking in Particle Flow 39
2.4.3. Calorimeter Clustering in Particle Flow 40
2.4.4. Physics Objects Reconstruction 41

2.4.5. Electron Reconstruction 41
2.4.6. Jet Reconstruction in the CMS Experiment 42
2.4.7. Photon Conversion Rejection 46
2.5. Luminosity Measurement in CMS 47

3. Software Tools **49**
3.1. Monte Carlo Event Generation 49
3.1.1. MadGraph 51
3.1.2. Pythia 51
3.1.3. PowHeg 52
3.2. Geant 52
3.3. Root 52
3.4. The CMS Software Framework 53
3.4.1. Modules in the CMS Software Framework 53
3.4.2. Workflow in the CMSSW Framework 54
3.4.3. The Full CMS Detector Simulation 55
3.5. The Worldwide LHC Computing Grid 56
3.5.1. Organisational Structure of the Grid 57
3.5.2. Grid Workflow 57
3.5.3. The Tier-2 Centre T2_BE_IIHE 58

II. Event Selection and Selection Efficiencies **61**

4. Event Selection **63**
4.1. Offline Event Selection 63
4.1.1. Selection Requirements for Electron and Jets Events 64
4.1.2. Summary of the Event Selection 70
4.2. Simulation to Collision Data Tuning 73
4.2.1. Trigger and Lepton Identification Event Weights 74
4.2.2. Pile-up Interaction Weights 74
4.2.3. Correction of $W \rightarrow \ell\nu$ to Next-To-Leading-Order 74
4.2.4. Parton Distribution Function Reweighting 74
4.2.5. Jet Energy Resolution 76

5. Lepton and Trigger Efficiency Measurement in the CMS Experiment **77**
5.1. Online Event Selection 77
5.1.1. Electron & Jets Triggers in the CMS Experiment 77
5.1.2. Measurement of the Efficiency of Electron and Jets Triggers . 79
5.1.3. The Tag and Probe Method 80
5.1.4. Measurement of the Electron Trigger Efficiency 81
5.1.5. The Inclusive Cut and Count Method 85
5.1.6. Measurement of the Jet Trigger Efficiency 86
5.1.7. Summary on Trigger Efficiencies 90

5.2. Lepton Efficiency Determination . . . 90
5.2.1. Electron Identification . . . 91
5.2.2. Photon Conversion Rejection . . . 92
5.2.3. Electron Isolation . . . 95
5.2.4. Summary Electron Efficiencies . . . 98

III. Cross Section Measurement 101

6. Measurement of the Top Quark Pair Production Cross Section 103
6.1. Method . . . 103
6.1.1. Template Fit Technique for the Cross Section Measurement . 104
6.2. Shapes and Data Driven Multijet Estimate . . . 105
6.2.1. Choice of Distribution for Template Fit . . . 105
6.2.2. Simulated Shapes . . . 106
6.2.3. Data Driven Multijet Shape . . . 106
6.2.4. Summary of Template Shapes . . . 110
6.3. Validation of Template Fit . . . 111
6.3.1. Pseudo-Experiment Generation . . . 111
6.3.2. Validation of Fit Results on Simulation . . . 113
6.3.3. Effect of Binning on Template Fit . . . 116
6.4. Measurement of the Cross Section . . . 117
6.5. Confirmation of Fit Stability on Data . . . 119
6.6. Kinematic Distributions. . . . 120

7. Systematic Uncertainty Determination 129
7.1. Methodology for the Evaluation of Systematic Uncertainties . . . 129
7.2. Sources of Systematic Uncertainties . . . 130
7.2.1. Statistical Uncertainty on Template Shapes . . . 130
7.2.2. Pile-Up Interactions . . . 131
7.2.3. Luminosity Uncertainty . . . 131
7.2.4. Factorisation Scale . . . 133
7.2.5. Jet - Parton Matching Threshold . . . 133
7.2.6. Top Quark Mass . . . 135
7.2.7. W/Z Ratio and Single Top . . . 136
7.2.8. Jet Energy Scale and Jet Energy Resolution . . . 137
7.2.9. Electron Energy Scale . . . 139
7.2.10. Lepton Identification and Trigger Efficiency Measurements . 142
7.2.11. Multijet Shape Estimation . . . 142
7.2.12. Parton Distribution Functions . . . 145
7.3. Summary Systematic Uncertainties . . . 146

8. Combined Electron/Muon and Jets Cross Section Measurement 153
8.1. Summary Muon and Jets Measurement . . . 153

8.2. Combined Measurement. 154
8.2.1. Definition of Template Shapes for Simultaneous Fit 154
8.2.2. Stability of the Template Fit 155
8.2.3. Result Combined Channel 157
8.3. Systematic Uncertainties Combined Channel 157
8.3.1. Top Quark Mass 160
8.3.2. Multijet Background Contribution Estimation 160
8.3.3. Summary of Systematic Uncertainties 162
8.4. Conclusions on Combined Measurement 162

9. Top Quark Mass From Cross Section **169**

IV. Conclusions **173**

10. Comparison with other Measurements **175**

11. Conclusion **179**

V. Appendix **183**

A. Technical Setup **185**
A.1. Calibration Constants and Software Versions 185
A.2. Datasets with Simulated Events 185
A.2.1. Primary Datasets with Collision Data 185
A.2.2. Description of Simulated Datasets 185
A.2.3. Selection of Simulated Events 192

B. Trigger Monitoring **195**
B.1. Electron Trigger Efficiency in Various Trigger Menus 195
B.2. Jet Trigger Efficiency in Various Trigger Menus 200

List of Figures **205**

List of Tables **208**

Bibliography **220**

Introduction

For more than two millenia mankind strived to understand the nature of the world it inhabits. Large progress in the understanding of nature has been made in the last century with the emergence of the theory of Quantum Physics [1] and since the mid-twentieth century with the development of the field of particle physics.

Particle physics describes the study of the fundamental constituents and interactions of the universe excluding gravity. According to the *Standard Model* of particle physics, a quantum field theory, all matter in the universe consists of twelve fundamental particles, six quarks and six leptons. The heaviest of those constituents that has been observed in nature is the *top quark* with a mass of $173.5 \pm 0.6 \pm 0.8\,\mathrm{GeV}/c^2$ [2]. It is more than 30 times heavier than the second heaviest quark, the *bottom quark*. This exceptional role makes it very interesting for the study of its properties and to perform precision tests of the Standard Model.

The Large Hadron Collider at the European Organisation for Nuclear Research (CERN) started high energy proton-proton collision data taking at a centre-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$ in March 2010. At this centre-of-mass energy, the top quark production rate is more than a factor of 20 higher than in any former particle physics experiment. Therefore, one refers to the Large Hadron Collider also as a top quark factory. The large amount of top quarks produced allows for measurements with a new level of precision in high energy physics. Large detectors such as the Compact Muon Solenoid (CMS) are used to record physics events.

One of the precision measurements is the determination of the top quark production cross section and is the topic of this thesis. High-precision theory predictions for the top quark pair production cross section are available, and a comparison with precision measurements of this observable in the experiments at the Large Hadron Collider allows for a test of the predictions of the Standard Model. In addition, a precise knowledge of the top quark pair production cross section is crucial to the search for and interpretation of physics beyond the Standard Model.

This thesis is structured in four parts. In the first part, the theoretical and technological principles of the experiments at the Large Hadron Collider are explained. The first chapter of this thesis gives an introduction to the Standard Model of particle physics and to the physics of the top quark. In the second chapter, the LHC and the CMS experiment are introduced, together with a description of the reconstruction methods for physics objects in the CMS data. Chapter three gives an overview of the software tools used to generate physics events and simulate their measurement in the CMS detector. In addition, Chapter three introduces the LHC Computing Grid, a network of computing centres distributed around the world to process data recorded by the LHC experiments.

In the second part of the thesis, the event selection and lepton and trigger efficiency measurements will be discussed. The fourth chapter summarises the event selection applied to collision data and to simulated samples. The physics processes that pass the event selection are described. Methods to tune the simulated event samples to collision data are also described in this chapter. In the fifth chapter, the measurement of lepton selection and trigger efficiencies is discussed, as this is a crucial ingredient of the top quark pair production cross section measurement.

The third part of this thesis consists of the measurement of the top quark pair production cross section and the estimation of the statistical and systematic uncertainties on the measurement. Chapter six describes how a maximum likelihood template fit is used to measure the top quark pair production cross section and summarises the results of the measurement. Kinematic distributions of various quantities in physics events are presented to give an overview of the agreement between simulated events and collision data. The systematic uncertainties on the measurement are studied in detail in Chapter seven. Chapter eight presents the combined measurement of the top quark pair production cross section in the electron/muon + jets decay channels using information on the muon + jets decay channel from Reference [3]. A method is presented to estimate the top quark mass based on the measured top quark pair production cross section in Chapter nine.

The fourth part of the thesis draws conclusions from the measurements. A comparison of the results of the cross section and mass measurements to the latest published results from the CMS and ATLAS collaboration is given in Chapter ten. A summary of the measurements and an outlook is presented in Chapter eleven.

Part I.

Theory and Experiment

1. The Standard Model of Elementary Particle Physics

The *Standard Model* of elementary particle physics is a quantum field theory that describes the fundamental particles and their interactions [4, 5, 6]. According to the current understanding of nature, all matter is built up of twelve fundamental fermions, six leptons and six quarks, and their anti-particles. They are listed in Table 1.1. Furthermore, four types of forces are responsible for the interaction of particles. The Standard Model of particle physics describes all twelve fundamental constituents of matter and three out of the four interactions, the strong, the weak and the electromagnetic interaction. The fourth interaction, gravity, has not yet been formulated in terms of a renormalisable quantum field theory.

Table 1.1.: Quantum numbers of fundamental particles described by the Standard Model [7].

Fermions	Generation			electr. charge	spin	weak isospin	colour
Leptons	ν_e	ν_μ	ν_τ	0	$\hbar/2$	$+1/2$	–
	e	μ	τ	$-e$		$-1/2$	
Quarks	u	c	t	$+2/3e$	$\hbar/2$	$+1/2$	r, g, b
	d	s	b	$-1/3e$		$-1/2$	

1.1. Field Theoretical Approach

The Standard Model is a field theory based on the definition of fermion fields and gauge symmetries. Requiring that the Lagrangian is gauge-invariant, local and renormalisable leads to a Lagrangian of the form

$$\mathcal{L}_{\text{Standard Model}} = \mathcal{L}_{\text{Gauge}} + \mathcal{L}_{\text{Matter}} + \mathcal{L}_{\text{Yukawa}} + \mathcal{L}_{\text{Higgs}}.$$

The gauge Lagrangian $\mathcal{L}_{\text{Gauge}}$ describes the kinetic energy of the gauge fields and their self-interaction whereas the matter Lagrangian $\mathcal{L}_{\text{Matter}}$ describes the kinetic energy of fermions and their interaction with the gauge fields. Mass terms for gauge bosons and fermions are forbidden by symmetries [8]. The Yukawa Lagrangian $\mathcal{L}_{\text{Yukawa}}$ is responsible for the creation of the fermion masses and the interaction

with the Higgs field and the Higgs Lagrangian $\mathcal{L}_{\text{Higgs}}$ defines the kinetic energy of the Higgs field, the gauge interactions and the Higgs potential. The non-zero vacuum-expectation value of the Higgs field breaks the electroweak local gauge symmetry and generates masses for the gauge bosons.

1.2. Quarks and Leptons

The fundamental particles of the Standard Model which are represented as fermion fields are described by several quantum numbers [7, 9]; the electric charge, the spin, the weak isospin and colour. Leptons are colourless and carry a spin of $\hbar/2$; $\hbar$ is the reduced Planck constant $\hbar = h/2\pi$ with the Planck constant [10] in SI units $h = 4.14 \cdot 10^{-15}\text{eV} \cdot \text{s}$. There are three charged lepton *flavours*, the electron e, the muon μ and the tau lepton τ, which carry an electric charge of $e = 1.602 \cdot 10^{-19}\,\text{C}$ [2]. They are grouped in doublets with the three neutral leptons, the neutrinos, which also exist like their charged counterparts in three flavours, the electron neutrino ν_e, the muon neutrino ν_μ and the tau neutrino ν_τ. Leptons interact only through the weak interaction in the case of the neutrinos and through the weak and the electromagnetic interaction in the case of the charged leptons. Neutrinos are assumed to be massless in the Standard Model. The rest masses of the charged leptons vary from $m_e = 0.5\,\text{MeV}/c^2$ for the electron up to $m_\tau = 1776.8\,\text{MeV}/c^2$ for the tau lepton.

There are six quarks, three down-type quarks, the *down* quark, the *strange* quark and the *bottom* quark, each of which carries a charge of $-1/3e$. In addition, there are three up-type quarks, the *up* quark, the *charm* quark and the *top* quark which carry an electric charge of $+2/3e$.

1.3. Interactions between Particles

The different particles interact with one another through the exchange of spin-1 gauge bosons. A summary of the gauge bosons and their quantum-mechanical properties can be found in Table 1.2.

Table 1.2.: The gauge bosons of the Standard Model and their characteristics. J denotes the angular momentum of the boson, P its parity, q its electrical charge and T_3 the third component of the weak isospin [7, 9].

particle	interaction	mass	J^P	q	T_3
Photon	e/m	–	1^-	0	0
Gluon	strong	–	1^-	0	-
Z^0	weak	91.18 GeV	1	0	0
$W^\pm$	weak	80.40 GeV	1	$\pm e$	± 1

1.3.1. The Electromagnetic and Weak Interactions

The electromagnetic force is mediated by massless photons. In the weak interaction massive vector-bosons are exchanged, the electrically neutral Z boson with a mass of $91.1876 \pm 0.0021\,\mathrm{GeV}/c^2$ [2] and a decay width of $2.49852 \pm 0.0023\,\mathrm{GeV}/c^2$ and the $W^{\pm}$ boson with positive or negative electric charge and a mass of $80.385 \pm 0.015\,\mathrm{GeV}/c^2$ [2] and a decay width of $2.085 \pm 0.042\,\mathrm{GeV}/c^2$. In the weak interaction the flavour eigenstates do not correspond to the mass eigenstates of the quarks. The eigenstates are connected by the unitary CKM[1] matrix which results from the diagonalisation of the fermion mass matrices in the field theoretical approach.

$$\begin{pmatrix} |d'\rangle \\ |s'\rangle \\ |b'\rangle \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix}$$

1.3.2. The Strong Interaction

The strong interaction is described by *Quantum Chromodynamics* (QCD) and is mediated by the exchange of massless gluons between particles that carry a colour charge. QCD is based on an SU(3) symmetry to describe three types of colour charges. SU(3) is a non-abelian group and therefore the gluons as gauge bosons carry themselves colour charges. A gluon-gluon self-interaction in the form of a triple-gluon and quadruple-gluon vertex is possible.

Divergences in the theory are eliminated by absorbing them into finite charges in a process called renormalisation. The coupling constant of the strong interaction becomes dependent on the renormalisation scale with $\alpha_s = \alpha_s(\frac{Q^2}{\mu_R^2})$ where Q^2 is the momentum transfer of the interaction and μ_R defines the renormalisation scale. At high energies a perturbative calculation becomes possible which shows that the strong coupling decreases logarithmically towards higher energies. At low energies, where the coupling constant takes on large values, a perturbative calculation is no longer possible. The large increase of the coupling constant at low energies leads to the effect of confinement. This means quarks can not be observed as free particles in nature and pairs or triplets of quarks/anti-quarks form colour-neutral hadrons. The transition of quarks and gluons to hadrons in the low energy regime is described as hadronisation.

According to the string hadronisation model [11], two oppositely colour charged quarks that move apart cause a force field in the form of a band of gluons between the quarks. The further the two particles move apart due to their kinematic energy, the higher the potential energy between them becomes. At a distance of approximately 1 fm, it becomes energetically more favourable if a new quark-anti-quark pair is formed in between the two original quarks. This process is repeated until the kinematic energy of the quarks is no longer large enough for the creation of new

[1] Cabbibo-Kobayashi-Maskawa

quark-anti-quark pairs resulting in the formation of several colour neutral hadrons. In addition, during the hadronisation process the boosted colour charged particles emit gluon radiation which leads to a process called parton showering. Also the gluons from the parton shower create quark-anti-quark pairs that hadronise again. In the end a collimated stream of colour neutral hadrons is formed which is called a *jet*. The sum of the four vectors of all particles within a jet equals the four-vector of the parton that hadronised and formed the jet.

1.4. Free Parameters of the Standard Model

For several decades the Standard Model has been accurately predicting a large variety of fundamental interactions and has been validated in various experiments. Nevertheless, there are 18 free parameters in the Standard Model that cannot be deduced from theory and are measured in experiments. Those parameters are [9]

- the strength of the electromagnetic and the strong coupling constants α_{em} and α_s,
- the mass of the W, Z and Higgs bosons,
- the masses of the six quarks and three charged leptons and
- the four independent parameters to describe CKM matrix elements.

1.5. Unit Conventions

In particle physics it is common and often easier not to use the *International System of Units* [12] (SI System[2]) but instead a system where constants are absorbed in specific variables. In this thesis, the *Heavyside-Lorentz* System is used with the definitions $\hbar = c = 1$, $\epsilon_0 = 1$ and $\alpha = \frac{e^2}{4\pi} \simeq 1/137$. $\hbar$ is the reduced Planck constant, c the speed of light, ϵ_0 the vacuum permittivity and α the fine structure constant.

1.6. Measurement of the Interactions of Fundamental Particles

In particle physics one is interested in measuring the interactions and behaviour of fundamental particles to draw conclusions on the laws of physics. The particle sources can be of natural origin as detected by cosmic ray experiments. Alternatively, the particles can be artificially accelerated as in a typical high energy physics particle accelerator. The particles can then be directed at a fixed target. Nevertheless, to reach higher interaction energies between the particles, as it is desirable for

[2]French: Système international d'unités

the search and discovery of new particles, particle colliders are employed. Bunches of particles are accelerated in opposite directions to collide at a pre-defined interaction point. The particles can be either pairs of electrons, protons, or their anti-particles, heavy ions or combinations thereof.

1.6.1. Proton-Proton Collisions

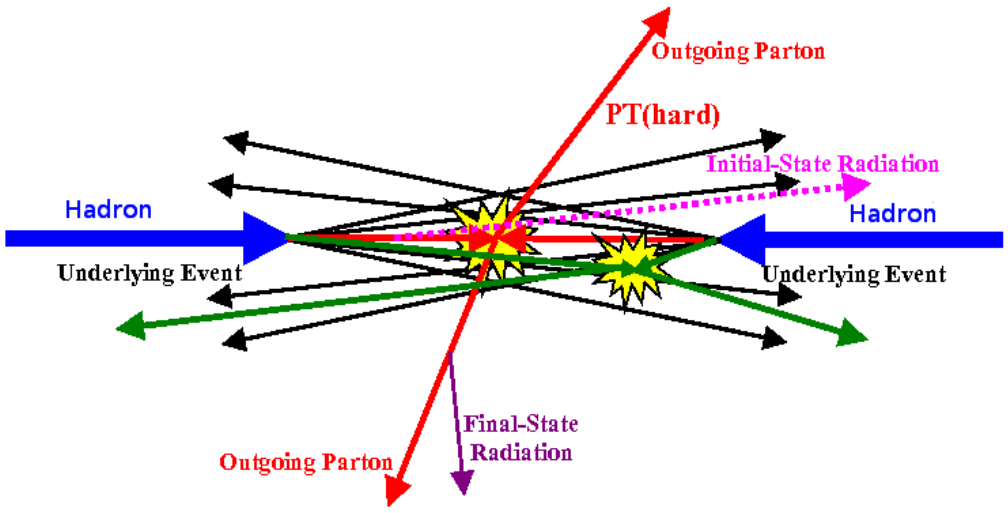


Figure 1.1.: Interaction between two hadrons in a collider experiment, taken from [13].

Protons consist of two up quarks and one down quark, called valence quarks. In addition, gluons are constantly emitted and absorbed within the proton which may also split into quark-anti-quark pairs, the sea quarks. This process happens in a very short amount of time and is therefore allowed by the *uncertainty principle* [14] by Werner Heisenberg. The valence quarks, sea quarks and gluons are the constituents of the protons and in proton-proton collisions two or more constituents of the two protons interact as depicted in Figure 1.1. The interaction of the proton constituents is referred to as the hard scattering process. The no longer colour neutral remnants of the initial protons emit gluon radiation which leads to the formation of new colour neutral hadrons that form the *underlying event*. The colour charged hadrons participating in the hard scattering process also emit gluon radiation before and after the interaction which is referred to as initial and final state radiation.

An important variable for the description of the hard scattering process in a

collider experiment is the *centre-of-mass energy* which is defined as

$$\sqrt{s} = \sqrt{(p_1 + p_2)^2},$$

where p_1 and p_2 are the four vectors of the two interacting particles. The centre-of-mass energy of the interactions in the data considered for this thesis was $\sqrt{s} = 7\,\text{TeV}$ in the year 2011.

The fraction of the total proton energy that is carried by an individual proton constituent is quantified by the *Bjorken-x* variable. The momentum distribution of the partons in the proton is measured in deep-inelastic scattering in electron-proton colliders, for example in the experiments at the Tevatron [15] or HERA accelerator [16], and then extrapolated to the centre-of-mass energies of the Large Hadron Collider. The information is provided in the form of parton distribution function (PDF) libraries. CTEQ 6.6 [17] is a popular library used in high energy physics. The prediction of the parton distribution functions for gluons, quarks and anti-quarks is shown in Figure 1.2.

For calculations of parton-parton interactions a factorisation of the problem is applied. The short distance hard scattering partonic cross section is treated separately from the long distance longitudinal momentum distribution functions. The factorisation is applied at the threshold called factorisation scale μ_F^2. The high energetic hard scattering process is calculated in perturbative QCD as described in Section 1.3.2 while the long distance pieces of the interaction are factorised into the parton distribution functions. It is a convention to chose the arbitrary values of the factorisation scale μ_F^2 and the renormalisation scale μ_R^2 to be the same $\mu_F^2 = \mu_R^2 = \mu^2$.

The instantaneous luminosity L is a measure of the number of particle interactions per time and is defined as

$$L = \frac{\gamma\ f\ k_B\ N_p^2}{4\,\pi\,\epsilon_n\,\beta^*}\ F,$$

for a proton-proton collider. γ is the Lorentz factor, f the revolution frequency, k_B the number of particle bunches, N_p the number of protons per bunch, ϵ_n the normalised transverse emittance, β^* the betatron function at the intersection point and F the reduction factor due to the crossing angle of the proton beams at the intersection points [18]. The value of the integrated luminosity over time $\int \mathrm{d}t\, L$ is a measure of the amount of delivered or recorded data in a given time interval.

In collider experiments, instead of using the polar angle Θ in coordinate systems it is common to describe this information by the rapidity

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}$$

or by the pseudo-rapidity

$$\eta = -\ln \tan \frac{\Theta}{2}.$$

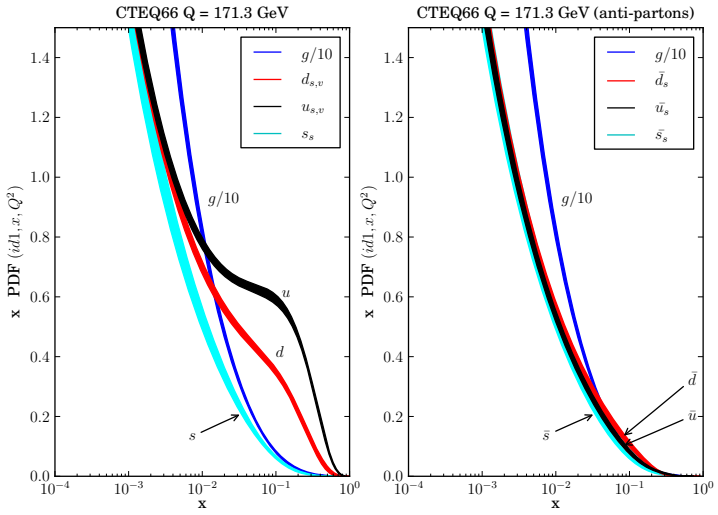


Figure 1.2.: Prediction of parton distribution functions as a function of Bjorken-x according to CTEQ 6.6 library [17] for gluons and quarks (left) and anti-quarks (right).

1.7. Top Quark Physics

The top quark is the heaviest observed fundamental particle with a world average mass of $173.5 \pm 0.6 \pm 0.8\,\mathrm{GeV}/c^2$ [2]. The top quark was first observed [19] at the *Tevatron* [15] accelerator at Fermilab [20], near Chicago, IL in 1995. The top quark is the second member of the weak isospin doublet formed together with the bottom quark, it carries an electric charge of $Q = 2/3$ and a value of the third component of the weak isospin of $T_3 = +1/2$. It has been predicted in 1977 after the discovery of the bottom quark [21].

In contrast to all other observed quarks, the lifetime of the top quark, approximately $5{\cdot}10^{-24}\,\mathrm{s}$, is too short to form any bound states such as top flavoured hadrons through the process of hadronisation as described in Section 1.3.2 or bound states like $t\bar{t}$ quarkonium [2].

The properties of the top quark are measured as a precision test of the Standard Model predictions. Furthermore, the mass of the top quark is one of the free parameters of the Standard Model and hence is not predicted by theory. Top quarks are a background to many physics processes beyond the Standard Model, for example searches of a so far unobserved hypothetical fourth generation quark doublet t' and b' or of decays of hypothetical supersymmetric (SUSY) particles like the decay of a gluino into a stop quark and a top quark: $\tilde{g} \to \tilde{t}t$.

1.7.1. Top Quark Production

Top quarks can be produced either in quark-anti-quark pairs through the strong interaction or as individual top quarks through the weak interaction.

The production of top quark pairs at high energies is described by perturbative QCD. The principles of the description of the processes of the strong interaction are described in Section 1.3.2 and Section 1.6.1. From perturbative QCD follows the total top quark pair production cross section

$$\begin{aligned}\sigma^{t\bar{t}}(\sqrt{s}, m_t) &= \sum_{i,j=q,\bar{q},g} \int \mathrm{d}x_i \mathrm{d}x_j f_i(x_i, \mu^2) \bar{f}_j(x_j, \mu^2) \\ &\quad \times \hat{\sigma}^{i,j\to t\bar{t}}(\rho, m_t^2, x_i, x_j, \alpha_s(\mu^2), \mu^2).\end{aligned}$$

ρ is defined as $\rho = 4m_t^2/\sqrt{\hat{s}}$ with the effective centre-of-mass energy squared $\hat{s} = x_i x_j s$. x_i and x_j are the Bjorken-x variables of the interacting partons, $f(x, \mu)$ are the parton distribution functions discussed in Section 1.6.1 with the energy scale variable μ.

Top quark pair production is initiated by quark-anti-quark interactions or gluon-gluon fusion. At the Tevatron accelerator the top quark production is dominated by quark-anti-quark annihilation of the valence quarks in the protons and anti-protons. At the LHC with the much higher centre-of-mass energy the momentum fraction of virtual gluons in the protons becomes high enough for top quark pair production and domiantes the production process. The corresponding Feynman diagrams are presented in Figure 1.3.

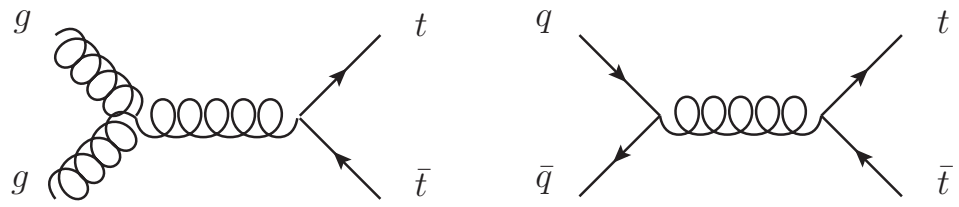


Figure 1.3.: Feynman diagram top quark pair production, taken from [3].

The production cross section for various Standard Model processes as a function of the centre-of-mass energy is shown in Figure 1.4. One can see that the top quark pair production cross section at the LHC increases by more than a factor of 20 compared to the cross section at Tevatron energies. One can also see that the total proton-proton cross section is eight orders of magnitude larger than the top quark production cross section which demonstrates the challenge of the selection of top quark pair events and rejecting this huge background. The dominant part of the total cross section is the jet production in $q\bar{q}/gg \to q\bar{q}/gg$ scattering processes.

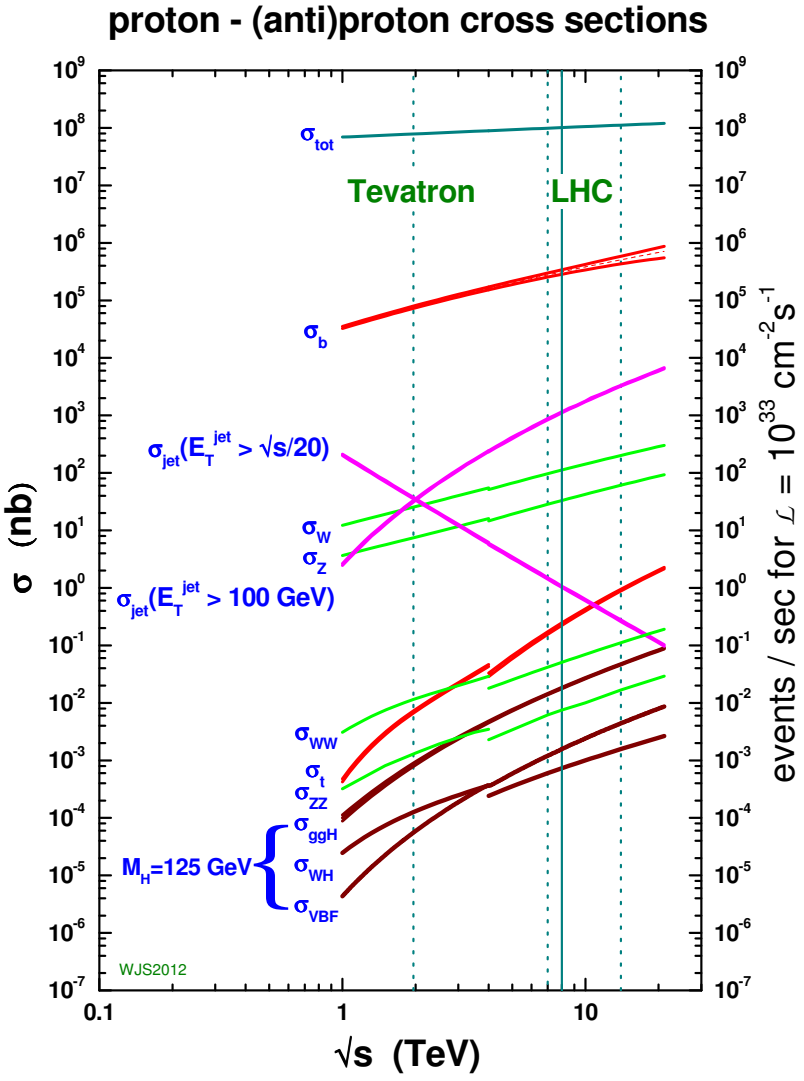


Figure 1.4.: Standard Model prediction of production cross sections for various processes [22] assuming a Higgs boson mass of 125 GeV. The discontinuity at a centre-of-mass energy of around 4 TeV is due to the change from proton-anti-proton to proton-proton collisions in the prediction. σ_t denotes the top quark pair production cross section.

With the electroweak interaction, top quarks are produced individually in contrast to quark-anti-quark pairs. Single top quark production in association with a W boson allows a measurement of the CKM matrix element $|V_{tb}|$. The current world average of this matrix element is $|V_{tb}| = 0.89 \pm 0.07$ [2]. A good understanding of top quark pair production is also necessary for this measurement as top quark pairs are one of the dominant background processes for single top quark production. The corresponding Feynman diagrams are presented in Figure 1.5.

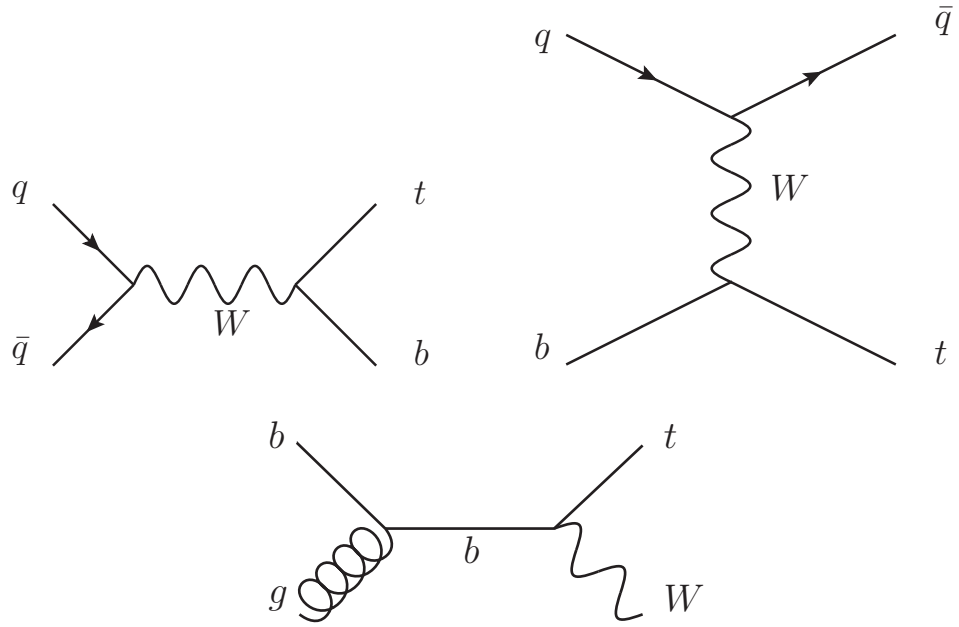


Figure 1.5.: Feynman diagrams single top quark production, taken from [3].

1.7.2. Decay of Top Quark Pairs

The top quark decays almost exclusively into a b quark and a W bosons due to the large value of the CKM matrix element $|V_{tb}|$. As a result, the decay of top quark pairs is classified in terms of the decay of the W boson. If both of the W bosons decay hadronically into light quarks one speaks of the fully-hadronic decay with the signature

$$t\bar{t} \rightarrow W^+W^-b\bar{b} \rightarrow q\bar{q}\,q\bar{q}\,b\bar{b}.$$

If both of the W bosons decay leptonically, one refers to the event topology as the di-leptonic decay channel with the signature

$$t\bar{t} \rightarrow W^+W^-b\bar{b} \rightarrow \ell\nu\,\ell\nu\,b\bar{b}.$$

In events with one leptonically decaying W boson and one hadronically decaying W boson one speaks of the semi-leptonic decay with the signature

$$t\bar{t} \rightarrow W^+W^-b\bar{b} \rightarrow \ell\nu\, q\bar{q}\, b\bar{b}.$$

Each of these decay channels has advantages and disadvantages for the measurement of the properties of top quark pairs. Taking QCD corrections into account 45.7% of all top quark pairs decay fully hadronically, 43.8% decay semi-leptonically and 10.5% decay di-leptonically.

This means that the channel with the largest abundance shows a decay signature of six quarks which hadronise to jets. This makes it challenging to separate top quark events from the large amounts of multijet background events observed at a hadron collider. The di-leptonic channel has a very clean signature with two leptons and additional jets but its branching fraction in the top quark decay is the smallest of all decay channels and the existence of two neutrinos in the event which cannot be measured directly in the detector makes it difficult to reconstruct the top quark kinematics. The semi-leptonic event channel combines the advantages of the fully-hadronic and fully-leptonic decay signature as the branching fraction is relatively large, it has a clear signature due to the charged lepton in the event and only one neutrino results from a W boson decay. This makes it an interesting decay channel to perform measurements on. The decay signature is illustrated in Figure 1.6.

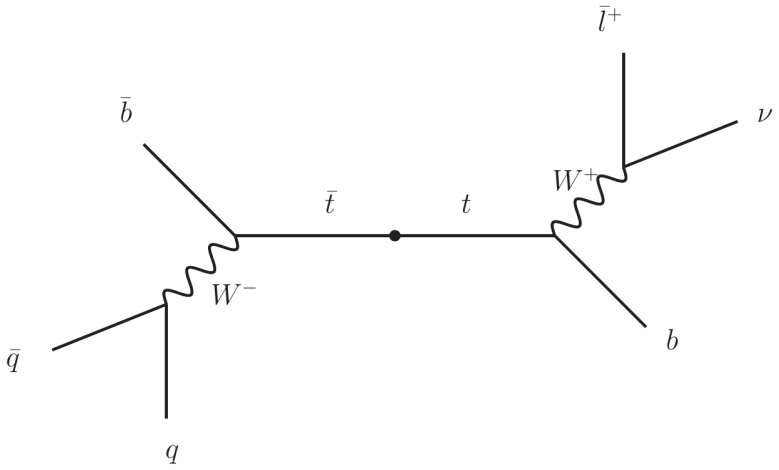


Figure 1.6.: Top quark pair decay in semi-leptonic channel.

2. The Large Hadron Collider and the Compact Muon Solenoid Experiment

The Large Hadron Collider (LHC) is a particle accelerator situated up to 150 metres underground on French and Swiss territory near Geneva, Switzerland. In the LHC proton-proton collisions take place at a centre-of-mass energy of 7 TeV in 2010 and 2011 and 8 TeV were reached in 2012. In the future, at the design energy of 7 TeV per beam, collisions will take place at a centre-of-mass energy of 14 TeV and an instantaneous luminosity of up to $10^{34}\,\mathrm{cm}^{-1}\mathrm{s}^{-1}$. Beside its capability as proton-proton collider, the LHC can also be employed as a heavy-ion (Pb-Pb) and proton-ion (p-Pb) collider. In this chapter the LHC is introduced including its major experiments. The CMS detector is described based on the CMS Physics Technical Design Report [18] (TDR). An explanation of the reconstruction of physics objects in the CMS detector is given.

2.1. The Large Hadron Collider

The acceleration of proton bunches feeding the LHC is performed in several steps as shown in Figure 2.1. In the *Linac2* [23] a proton stream is accelerated up to an energy of 50 MeV and injected to the *Proton Synchrotron* [24] (PS). In the PS proton bunches are formed, accelerated up to an energy of 26 GeV and then injected into the *Super Proton Synchrotron* [24]. There the proton bunches are accelerated up to an energy of 450 GeV. In the final step the proton bunches are then injected into the LHC where they are accelerated up to a kinematic energy of currently 4 TeV.

In the LHC proton bunches are moving inside two separated vacuum tubes in both directions, clockwise and counter-clockwise as depicted in Figure 2.2. The LHC is a 27 km long circular collider and is divided into octants with eight arcs and eight straight sections. Superconducting NbTi dipole magnets with fields up to $B = 8.33\,\mathrm{T}$ keep the beam on its circular trajectory along the LHC and superconducting quadrupole magnets focus the beam to prevent its dispersion. The beam is accelerated in four locations in superconducting radio frequency (RF) cavities. At four collision points the accelerated proton bunches are crossed and collisions between the protons take place as described in Section 1.6. A beam dump insertion allows the extraction of the beam with horizontally deflecting fast-pulsed magnets [24].

Up to 1331 bunches of protons with a bunch spacing of 50 ns were formed per beam with $1.4 \cdot 10^{11}$ protons each in the LHC operation in 2011. The centre-of-

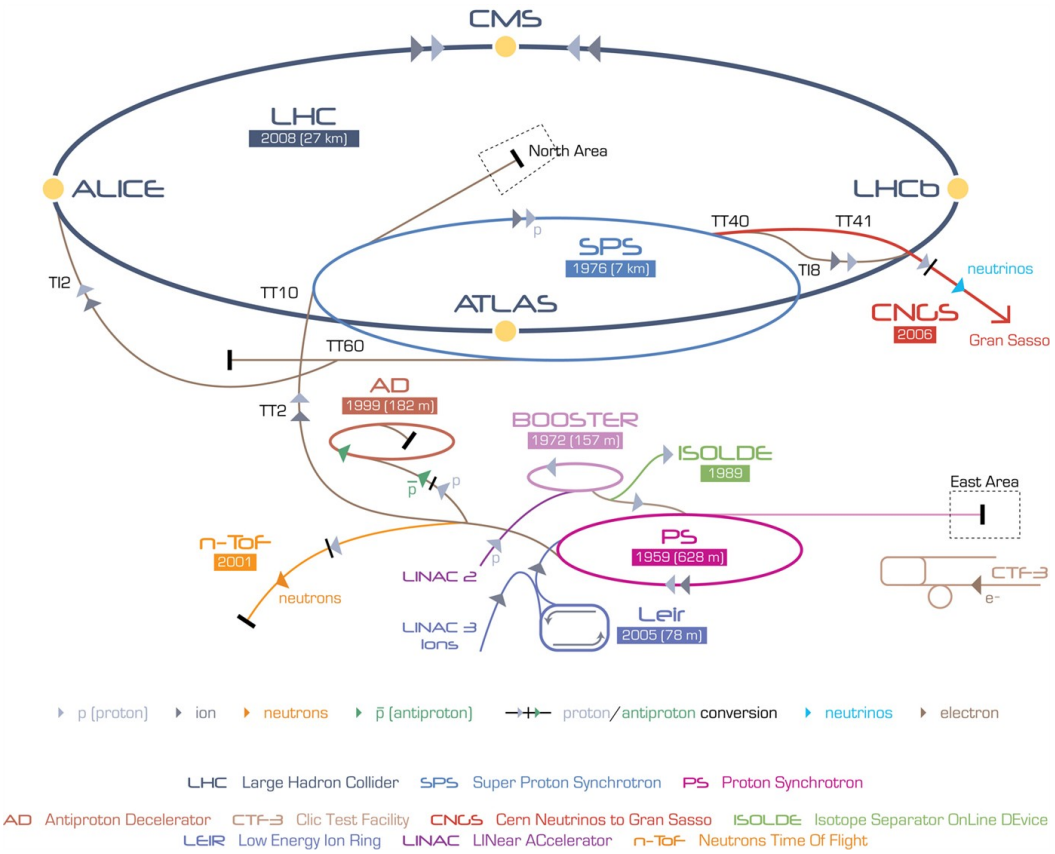


Figure 2.1.: Acceleration chain for protons. Protons undergo subsequent acceleration steps, first, in the Linac2, then in the Proton Synchrotron where bunches are formed and afterwards in the Super Proton Synchrotron from where they are injected into the LHC ring. In the LHC ring the proton bunches are accelerated to the nominal beam energy of currently up to 4 TeV. Figure taken from [25].

mass energy in the collisions was $\sqrt{s} = 7\,\text{TeV}$ [26]. A peak-luminosity of $3.65 \cdot 10^{33}\,\text{cm}^{-2}s^{-1}$ was reached. At the design centre-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$ the number of bunches per beam will be doubled to 2808 with a bunch spacing of 25 ns and $1.15 \cdot 10^{11}$ protons per bunch. The design luminosity is $10^{34}\,\text{cm}^{-1}\text{s}^{-1}$ [24].

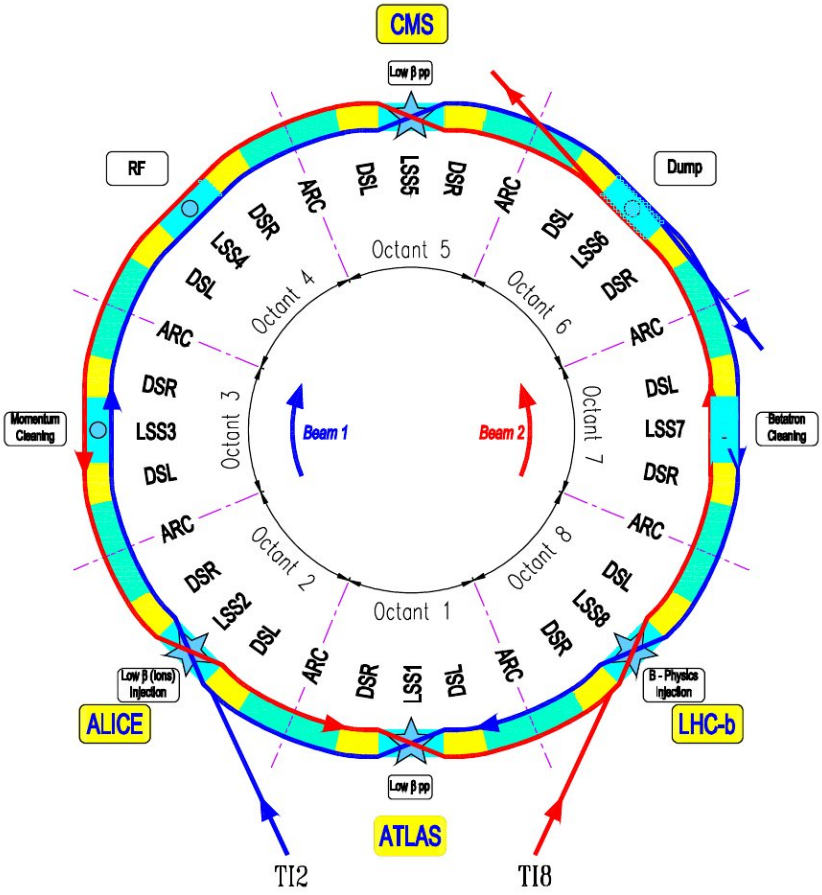


Figure 2.2.: Layout of the Large Hadron Collider. In eight arcs and straight sections protons are accelerated, focused and kept on their trajectory along the LHC beam line. Proton-proton collisions take place at four interaction points. Figure taken from [24].

2.1.1. The Experiments at the LHC

There are several large-scale experiments located at the LHC. The two multi-purpose detectors CMS and ATLAS [27] were designed to be able to perform a large variety of physics measurements and searches, for example the search for physics beyond the Standard Model and the discovery of the Higgs boson. The latter was discovered in July 2012 [28, 29]. The ALICE experiment [30] was designed

to study heavy-ion collisions and the properties of quark-gluon plasma. The fourth large experiment is the LHCb experiment [31], specialised in the precise study of *bottom*-quark physics and CP violation.

In addition, the LHCf experiment [32] performs measurements of particle showers similar to those in astroparticle physics but under laboratory conditions. It is situated between the two LHC beampipes 140 m down the beamline from the ATLAS detector. The TOTEM experiment [33] measures the forward charged particle pseudorapidity density in proton-proton collisions at the LHC centre-of-mass energy.

2.2. The Compact Muon Solenoid Experiment

As a general-purpose detector the Compact Muon Solenoid (CMS) experiment targets a large variety of physics analyses. It is 21.6 m long, with a diameter of 14.6 m and a weight of 12.5 thousand tons. The detector is divided in a central barrel region which is completed by two endcaps, one on each end of the detector. A schematic view of the CMS detector is given in Figure 2.3.

There were several main design goals. First, a very precise muon system which allows the measurement of muon momenta up to 1 TeV with high resolution. Second, an electromagnetic calorimeter with very low response times for high luminosity conditions with a high energy resolution and short radiation length. Furthermore, a precise tracking system for efficient vertex reconstruction and impact parameter determination. A motivation from the analysis point of view for these characteristics is the search for the Higgs boson in the $H \to \gamma\gamma$, $H \to WW$ and $H \to ZZ$ channels.

The CMS detector features four concentrically layered major subdetector systems as well as the magnetic coil for the generation of a magnetic field. The inner tracker and the electromagnetic and the hadronic calorimeters are placed inside the magnetic coil while the muon system is placed outside the coil.

2.2.1. The CMS Coordinate System

For the discussion of the detector components and of physics events the coordinate convention used by the CMS collaboration is explained. The centre of the system lies in the nominal collision point of the CMS detector. The y-direction in the CMS coordinate system is on the vertical axis of the detector towards the top of the CMS detector. The x-axis is on the horizontal axis of the detector pointing towards the centre of the LHC accelerator. The z-axis lies counter-clockwise along the LHC beam line direction. The azimuthal angle ϕ is defined in the x-y plane of the detector. The polar angle θ is measured from the z axis. ρ is the radial transverse distance of a point from the beam axis.

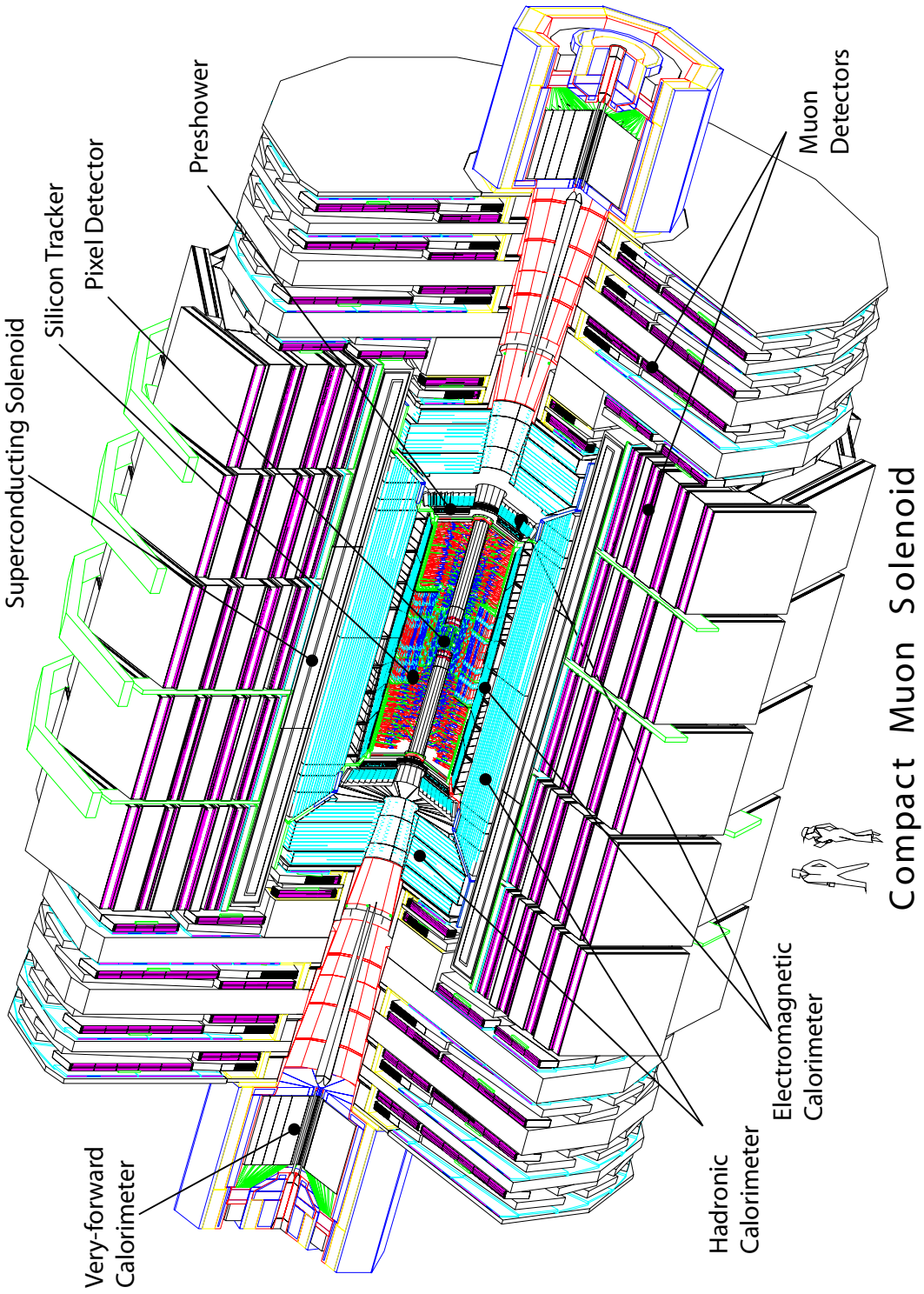


Figure 2.3.: Schematic view of the CMS detector. Figure taken from [18].

2.2.2. The Superconducting Solenoid Magnet

A high magnetic field makes it possible to measure precisely the momentum of charged particles, in particular the momentum of muons with a high transverse momentum. CMS features a 13 m long superconducting solenoid magnet with an inner diameter of 5.9 m and a design magnetic field of 4 T. Up to now, the magnetic field was kept at a value of 3.8 T for safety reasons.

2.2.3. The Inner Tracking System

The inner tracking system is the subdetector located the closest to the interaction point in the very centre of the CMS detector. The cylindrical tracking volume is 5.8 m in length and 2.4 m in diameter.

The pixel detector is located the closest to the interaction vertex with a radial distance of the detector of $\rho < 20\,\text{cm}$ from the centre. The pixel detector features 66 million pixels which are arranged in three layers in the barrel region and in two layers in the endcap region. The size of an individual pixel is $100 \times 150\,\mu\text{m}^2$. Although this is the region with the highest particle flux, the average occupancy per LHC crossing is only 10^{-4} thanks to the small size of the pixels.

The silicon microstrip detector is located in the next concentric layer around the pixel detector. With ten layers and 9.6 million strips, it features in the radial distance $20 < r < 55\,\text{cm}$ a minimum cell size of $10\,\text{cm} \times 80\mu\text{m}$ which leads to an average occupancy of $2 - 3\%$ per bunch crossing. In the region of $\rho > 55\,\text{cm}$ the microstrip detector consists of strips with a maximum cell size of $25\,\text{cm} \times 180\mu\text{m}$ with an average occupancy of 1%. The layout of the tracker modules in the CMS detector is shown in Figure 2.4.

The capability of high precision measurements with the inner tracking system is important for measurements of the impact parameter and transverse momentum of charged particle tracks as well as for the reconstruction of secondary vertices used in *bottom*-flavoured jet identification techniques.

The expected tracker resolution from test beams is $30-40\,\mu\text{m}$ [35]. The measured resolutions on 7 TeV collision data are presented in Figure 2.5. The resolutions vary for different modules between 15 and $45\,\mu\text{m}$.

The signal-over-noise and efficiency performance of the CMS tracker is presented in Figure 2.6. The average signal-over-noise ratio is around a value of 20. The tracker efficiency for enabled modules in the measurement is close to 100%.

2.2.4. The Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is a hermetic and homogeneous calorimeter build from 61200 lead tungstate ($PbWO_4$) crystals in the barrel and 7324 lead tungstate crystals in the endcap. It covers the pseudorapidity range of $0 < |\eta| < 3.0$ as can be seen in Figure 2.7. Lead tungstate features radiation hardness, a fast response as 80% of the light is emitted within 25 ns and an radiation length of $X_0 = 0.89\,\text{cm}$ with a Molière radius of 2.2 cm.

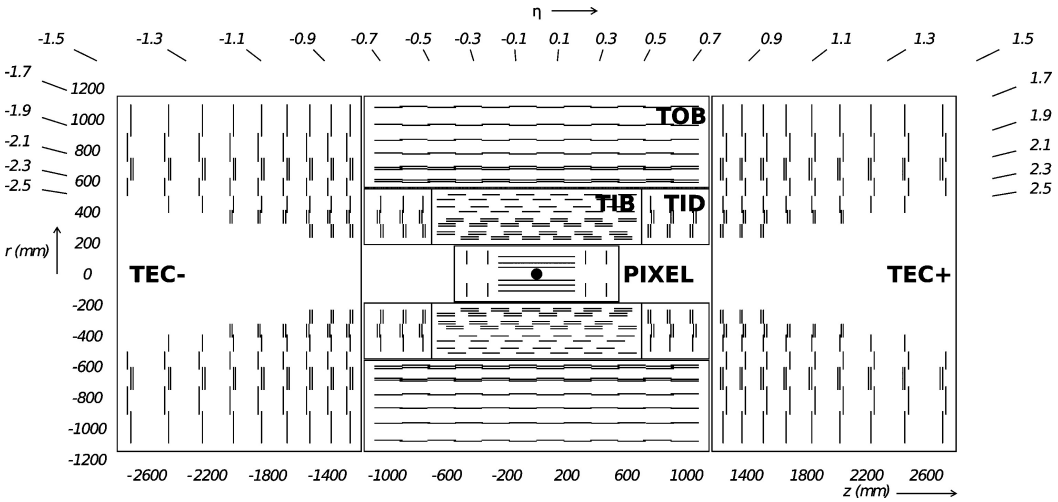


Figure 2.4.: Layout of the modules in the CMS inner silicon tracker in the ρ-z plane. The vertical axis shows the radius from the collision point in millimetres. The horizontal axis shows the distance towards the collision point along the beam line in millimetres. Figure taken from [34].

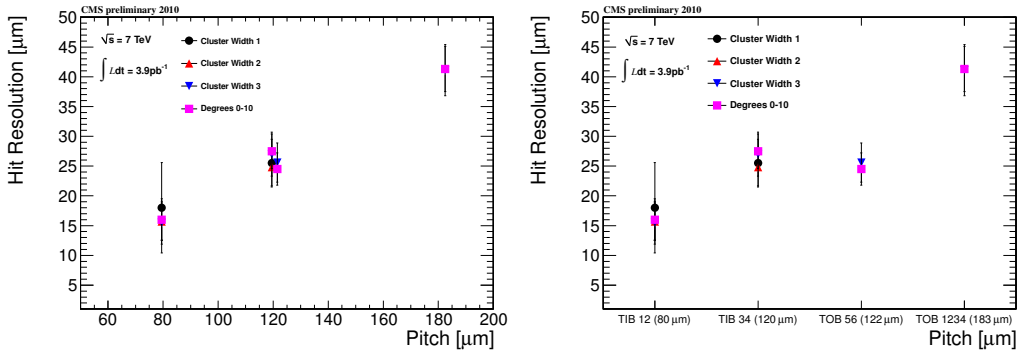


Figure 2.5.: Overall tracker resolution and resolution for individual modules in 7 TeV collision data. Figures taken from [36].

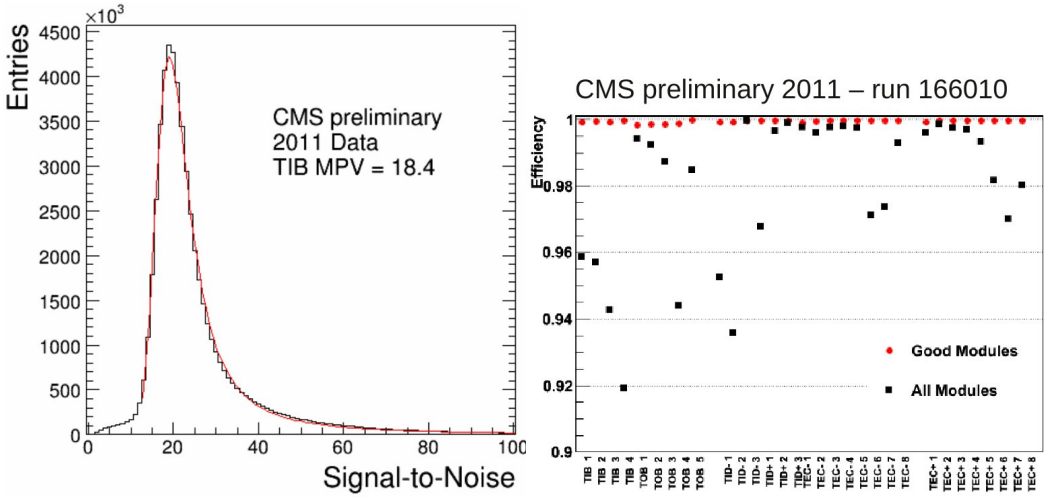


Figure 2.6.: Tracker performance in 2011 data. The signal-over-noise ratio is in average close to a value of 20. In an average run the track efficiency in good modules is close to 100 % while the efficiency is larger than 92 % including also modules with known problems. Figures taken from [36].

Photodetectors with an intrinsic gain of signal amplitude are employed due to the low gain of the scintillating light in lead tungstate. In the barrel region of the detector silicon avalanche photodiodes (APDs) are used while in the endcap region with its larger proton-beam induced backgrounds vacuumphototriodes (VPTs) are used.

The barrel region is subdivided into 36 supermodules in the region $0 < |\eta| < 1.479$. Each supermodule consists of 20×85 crystals in $\phi \times \eta$ and each crystal covers $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$. Each crystal has a front face cross section of $22 \times 22\,\mathrm{mm}^2$ and a length of $230\,\mathrm{mm}^2$ which corresponds to 25.8 radiation lengths.

In the endcap region a preshower detector is installed in front of the electromagnetic calorimeter in order to enhance π^0 rejection. The active region is for the electromagnetic calorimeter in the pseudorapidity range of $1.479 < |\eta| < 3.0$. The front face crystal cross section is $28.6 \times 28.6\,\mathrm{mm}^2$ with a crystal length of 270 mm which corresponds to 24.7 radiation lengths.

The energy resolution σ_E/E of the electromagnetic calorimeter has been measured in electron test beams and is very good approaching $\sigma_E/E = 0.39 \pm 0.01\%$ for electrons with a kinetic energy of 120 GeV [18]. The electromagnetic calorimeter was designed for and used in the search of the Higgs boson, especially in the $H \to \gamma\gamma$ channel. Measured energy resolutions on 2011 data can be found in Figure 2.8. The measurements on 2011 data contain electrons from an energy range which explains the worse resolution observed. The reconstructed di-electron mass spectrum in 7 TeV collision data is presented in Figure 2.9.

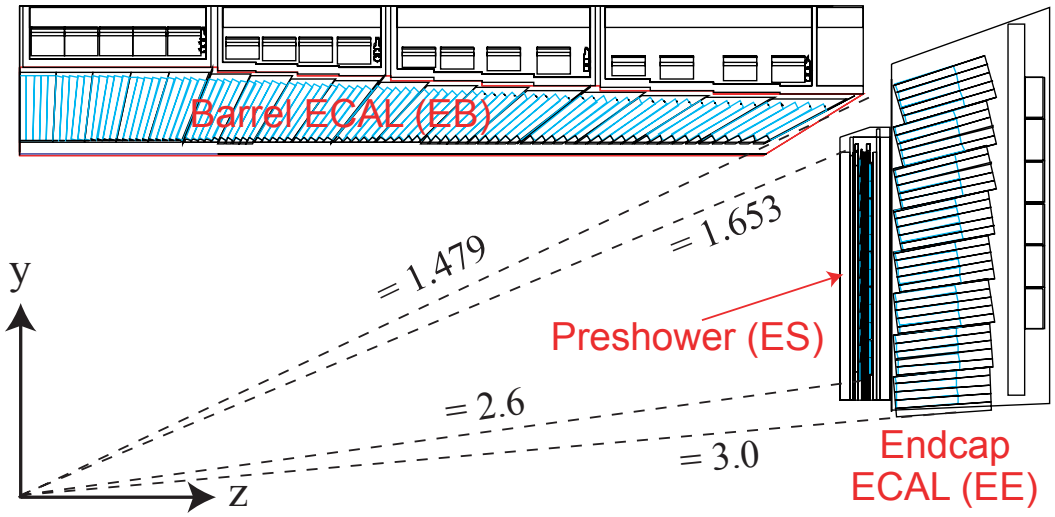


Figure 2.7.: Schematic view of the CMS electromagnetic calorimeter and the preshower detector. Figure taken from [18].

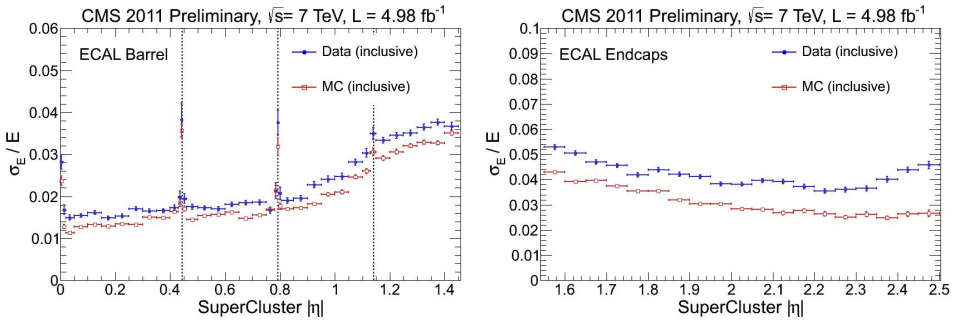


Figure 2.8.: Energy resolution of the electromagnetic calorimeter in the barrel and endcap region for the 2011 data taking period [37].

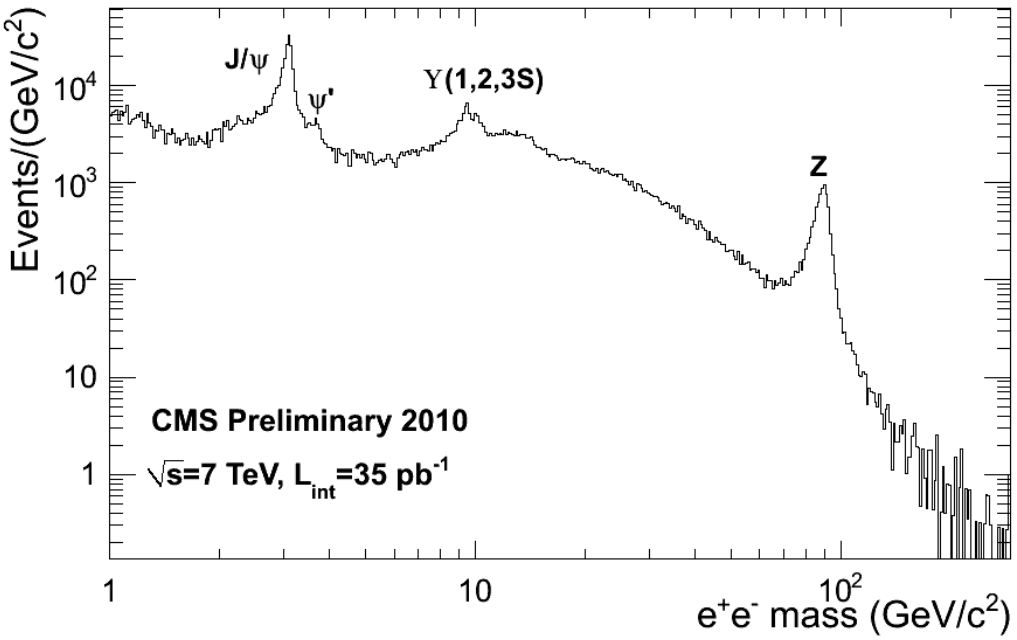


Figure 2.9.: The high resolution of the CMS electron reconstruction becomes clear from the reconstructed dielectron mass spectrum in the 2010 data. Figure taken from [38].

2.2.5. The Hadronic Calorimeter

The hadronic calorimeter (HCAL) is a brass/scintillator sampling calorimeter in the pseudrorapidity range $0 < |\eta| < 3.0$. The scintillation light is conducted through wavelength-shifting fibres which are embedded in the plastic scintillator tiles towards multi-channel hybrid photodiodes (HPDs). Brass as an absorber has the advantage of short interaction lengths and being non-magnetic. In the barrel region $|\eta| < 1.4$ the hadronic calorimeter features 2304 calorimeter towers with a dimension of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. An additional layer of the hadronic calorimeter is placed outside of the magnetic coil in the pseudorapidity range $0 < |\eta| < 1.26$ to sample energy from hadron showers leaking through the rear of the calorimeters. Together with the outer hadronic calorimeter the effective thickness in the barrel region is twenty interaction lengths.

In the endcap in the pseudorapidity region $|\eta| < 5.0$ the hadronic calorimeter is composed of iron and quartz-fibre arranged in 2304 towers. The Čerenkov light emitted in the quartz when hit by particles is measured by photomultipliers. A map of the hadronic calorimeter towers inside the CMS detector is shown in Figure 2.10.

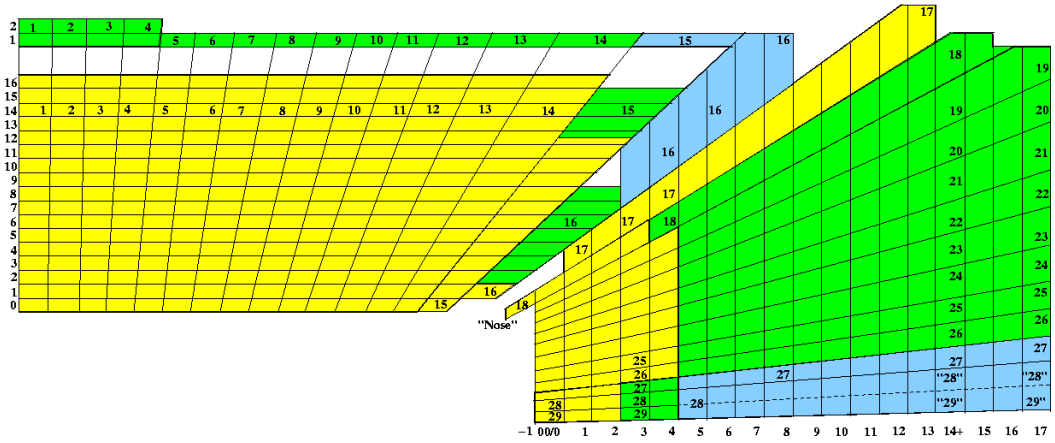


Figure 2.10.: Map of the towers of the hadronic calorimeter in the ρ-z plane [18].

The energy response of the hadronic calorimeter as a function of the transverse momentum of tracker tracks is shown in Figure 2.11 for different regions of the CMS detector. In all regions, an uncalibrated response of larger than 0.8 is observed for tracker tracks exceeding a transverse momentum of 20 GeV/c. Reconstructed jets are calibrated as described in Section 2.4.6.

The most important features of the hadronic calorimeter are the ability to minimise non-Gaussian tails in the energy resolution by using an absorber material with short interaction lengths and to provide good containment and hermiticity for the measurement of missing transverse energy E_T^{miss}. E_T^{miss} is defined as the negative vectorial sum of all particle candidates. The performance of E_T^{miss} reconstruction in 2011 data is presented in Figure 2.12 for $Z \to \mu\mu$ events with no intrinsic source

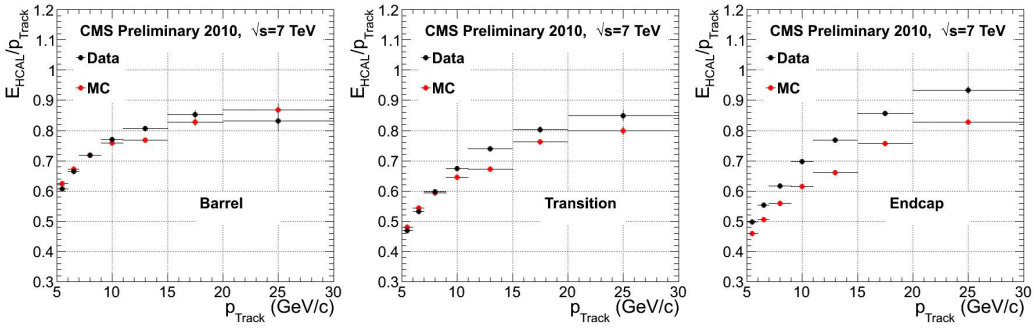


Figure 2.11.: Energy response of the hadronic calorimeter relative to the transverse momentum of tracker tracks as a function of the tracker track in 7 TeV collision data. Figure taken from [39].

of E_T^{miss}. In the early data with lower numbers of pile-up interactions in average a value of $E_T^{\text{miss}} = 14\,\text{GeV}$ was reconstructed. In the data taking period with higher numbers of pile-up interactions the average amount of reconstructed E_T^{miss} increases to $E_T^{\text{miss}} = 18\,\text{GeV}$. The agreement between the collision data and the simulation is very good.

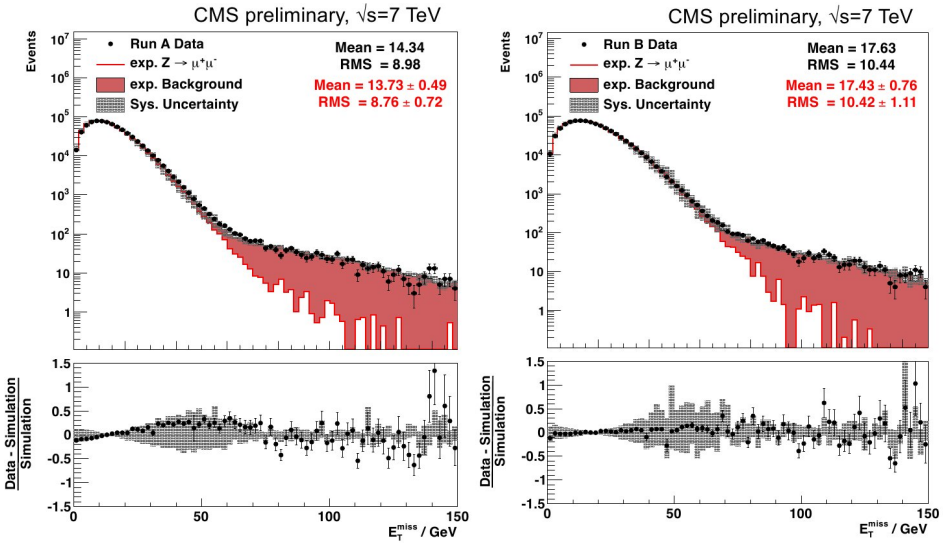


Figure 2.12.: E_T^{miss} resolution in the 2011A (left) and 2011B (right) data taking period in $Z \rightarrow \mu\mu$ events with no intrinsic source of E_T^{miss}. Figures taken from [40].

2.2.6. The Muon System

The muon system consists of three types of detectors. In the barrel region within the pseudorapidity range $0 < |\eta| < 1.2$ where the neutron-induced background is small and the muon rate is low, aluminium drift tubes (DTs) are employed. The size of the drift tube chambers varies between $1990 \times 290 \times 2536\,\text{mm}$ and $5966 \times 290 \times 2536\,\text{mm}$ [41]. In the endcap region, up to a pseudorapidity of $|\eta| < 2.4$ with a high neutron flux and magnetic field cathode strip chambers (CSCs) are used. Cathode strip chambers are up to 2 m long and 25 cm thick. Up to a pseudorapidity of $|\eta| < 1.6$ resistive plate chambers (RPCs) are installed in addition to the aforementioned detectors. They provide a fast response with a good time resolution but feature a coarser position measurement than DTs and CSCs. There are 480 resistive plate chambers installed with a strip length of either 85 cm or 130 cm. A schematic view of the CMS muon system is presented in Figure 2.13.

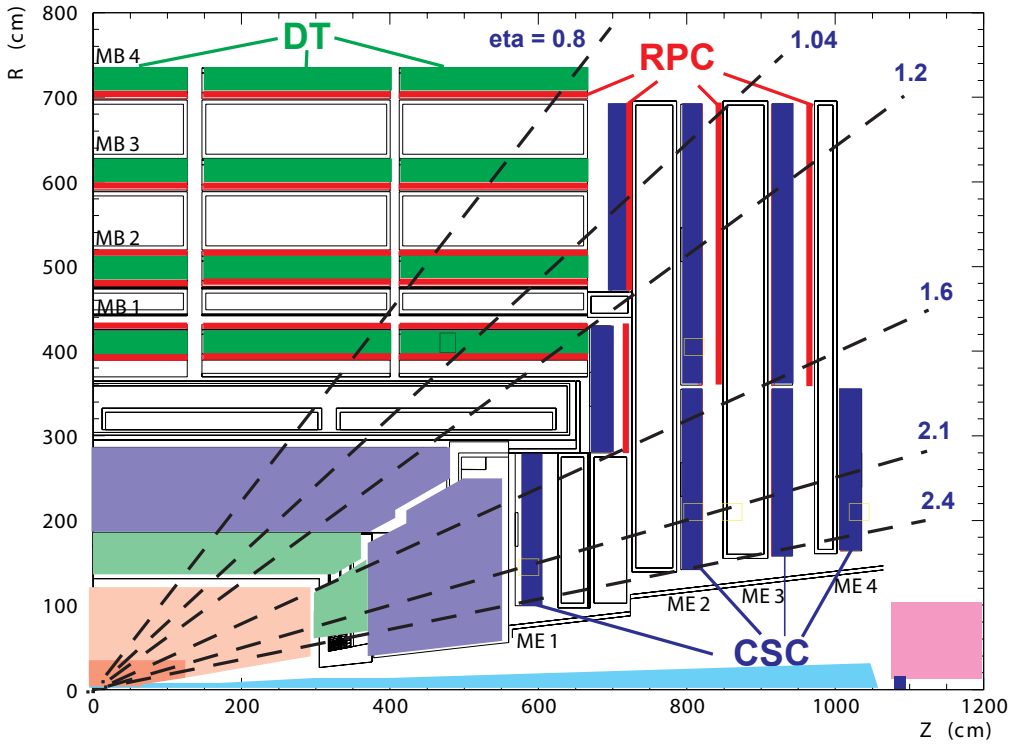


Figure 2.13.: Schematic view of the CMS muon system. In the barrel region drift tubes are installed while in the endcap region cathod strip chambers are used. In both regions resistive plate chambers are installed. Figure taken from [18].

Muons are measured in the inner tracker, in the calorimeter as minimum ionising particles and in the outer muon system. The inner tracker provides high precision for the measurement of the momentum of muons with low transverse momentum

while the muon system dominates the resolution of the momentum measurement for high transverse momentum muons up to $p_T = 1\,\text{TeV}$. Resolutions of better than $\Delta p_T/p_T = 10^{-2}$ can be reached for muons with a transverse momentum below $50\,\text{GeV}/c$ [18].

Figures 2.14 and 2.15 show the muon reconstruction performance in 7 TeV collision data. The average reconstruction efficiency is higher than 90 % in all subdetectors. The high resolution of the CMS muon system becomes clear from the reconstruction quality of the dimuon mass spectrum shown in Figure 2.16.

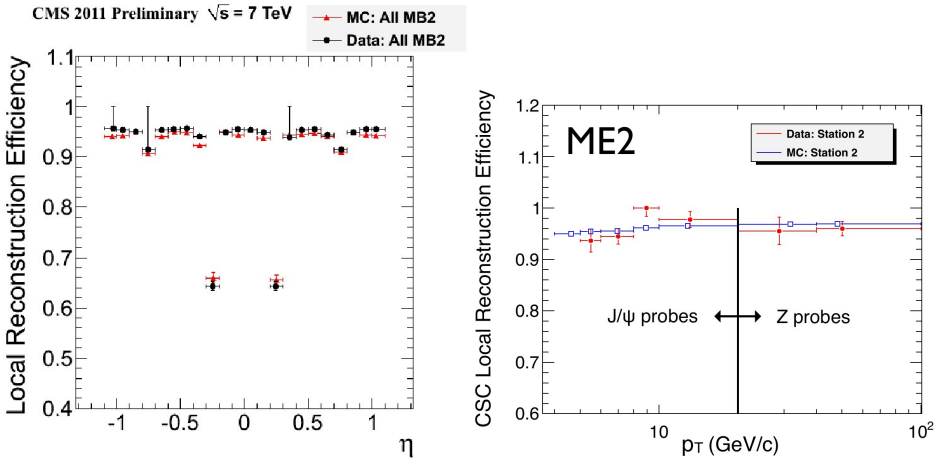


Figure 2.14.: Local reconstruction efficiency of drift tube chambers in the 2011 data and cathode strip chambers in $\sqrt{s} = 7\,\text{TeV}$ collision data. Figures taken from [42, 43].

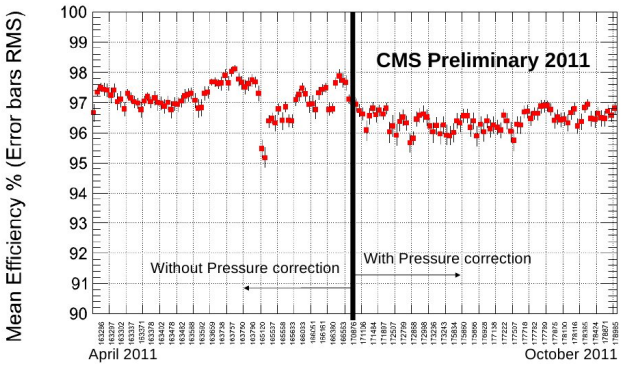


Figure 2.15.: Reconstruction efficiency of resistive plate chambers in the 2011 data as a function of time. Figures taken from [44].

Muons offer an excellent handle to identify events as they have a very low background contamination due to the large amount of absorber material between the muon chambers and the core of the detector.

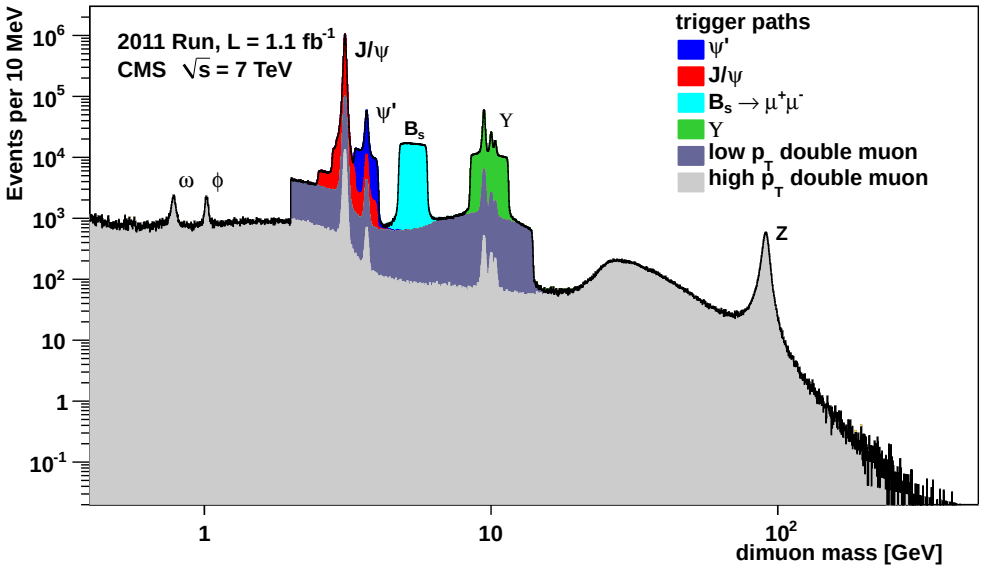


Figure 2.16.: The high resolution of the CMS muon system becomes clear from the reconstructed dimuon mass spectrum in the 2011 data. Figure taken from [45].

2.2.7. The Trigger System

At the LHC bunch crossings of the proton bunches happen at a rate of 40 MHz at design luminosity, while it is half that, 20 MHz, for 2011 data. With the computing resources available today, it is not possible to write out the collision information for all of these events; only the storage of events at a rate of approximately 300 Hz is possible. To achieve this reduction in data rate the CMS *trigger system* has been implemented. During data taking this system identifies events which are of interest for further analysis and triggers the write-out of the event.

In the CMS detector the trigger system is implemented in different levels. First, the hardware based Level-1 trigger system processes physics events; the software based High-Level Trigger then conducts a more advanced analysis of the events accepted by the Level-1 trigger before they are written out.

In the Level-1 trigger system the event rate is reduced from 40 MHz to 100 kHz. To reduce this rate further to 300 Hz the more advanced reconstruction algorithms of the High-Level Trigger system are employed.

The Level-1 Trigger System

The Level-1 trigger system [46] is a hardware-based triggering system. Custom-made integrated circuits and electronics provide low-level event and object reconstruction mechanisms. Muons are reconstructed from hits in the muon system

while electrons and jets are reconstructed from energy depositions in the calorimeter above a certain threshold.

The Level-1 trigger system is very fast because it evaluates the detector information only locally. There are no correlations between the subdetectors and the time-consuming track reconstruction in the inner tracking system is not used.

The High-Level Trigger System

At a rate of 100 kHz the full detector information for a given event can be transferred through the readout system to a local computing farm. There, the events are processed by the High-Level Trigger system [47]. All sub-detector channels are read out for the High-Level Trigger and the information between the sub-detectors is correlated. The time consuming tracking which is not used in the Level-1 trigger is available and used in the High-Level Trigger.

The event reconstruction algorithms that are used by the High-Level Trigger are more sophisticated than those of the Level-1 Trigger and close to the offline event reconstruction.

For a given run a certain list of *trigger paths* is defined which specify different sets of software modules for the event reconstruction and filters which decide based on pre-defined conditions if an event is written out to the storage system or is discarded. An example for a top quark physics specific trigger path is discussed in Section 5.1.1. At an event rate of 300 Hz events are selected by the High-Level Trigger and then transferred for full event reconstruction and permanent storage to the CERN Tier-0 centre, a computing infrastructure which is described in Section 3.5. A more detailed explanation of the CMS trigger system with respect to the trigger paths used for the analysis described in this thesis can be found in Section 2.3 and Section 5.1.

2.3. Online Object Reconstruction in the CMS Trigger System and General Trigger System Setup

2.3.1. General Setup of the CMS Trigger System

Trigger Menus

As described in Section 2.2.7 there is a maximum rate at which data can be recorded for permanent storage. The overall rate consists of the contribution of various triggers. A certain set of triggers in a given version is represented by a *trigger menu*.

The rate of events accepted by the menu must not exceed the overall available trigger rate. There is a need to adapt the trigger menu when necessary as the LHC beam conditions changed significantly throughout the 2011 data taking period in terms of an increase in instantaneous luminosity and number of pile-up interactions. Those changes make it necessary to drop triggers from the menu but also to

Table 2.1.: Trigger menus of the 2011 data taking period and run ranges for which the given trigger menu was used. The recorded luminosity per trigger menu is quoted.

Menu	Run range	Recorded Luminosity (fb^{-1})
5e32	160404 – 163869	0.2
1e33	165088 – 167913	1.0
2e33	170249 – 173198	0.8
3e33	173236 – 178380	2.1
5e33	178381 – 180252	0.9
sum		5.0

include new triggers. The trigger menus considered for the data of this analysis are summarised with the according run ranges in Table 2.1. The name of the trigger menu originates from the instantaneous luminosity in a given data taking period for which the trigger menu is optimised.

Groups of triggers are written out in different data streams called *primary datasets*. This makes the access for the user easier and reduces the processing time of the relevant data for the analysis. Primary datasets share triggers with a large event overlap to reduce duplicate data transfer and storage of events. Examples for primary datasets are the *SingleElectron* dataset to which all events are written by single electron triggers, or the *ElectronHad* dataset which is used to store events with at least one trigger electron and additional jets or other significant energy deposits in the hadronic calorimeter.

Prescaling of Triggers

If a trigger has to be kept in a given trigger menu and, together with the remaining triggers of the menu, the rate of this trigger is too high for the overall maximum allowed rate of the trigger system, it is possible to apply a *prescale* to the trigger at Level-1 and High-Level Trigger stage. This means that a certain fraction of the overall events that would have been written out by this trigger is randomly discarded which reduces the trigger rate but also the effective luminosity recorded by this trigger.

In a given trigger menu there are several fixed prescale values for all triggers available called *prescale columns*. These make it possible to react to changes of beam-conditions during data taking, for example the reduction of the instantaneous luminosity during a long run, which leads to a decreasing collision rate throughout the run.

This means it is possible that triggers are prescaled in one part of a given run but not in other parts. The triggers used in this thesis were at all times unprescaled in

all prescale columns to maximise the amount of recorded events of interest.

2.3.2. Electron and Jet Reconstruction in the CMS Trigger System

The calorimetric Level-1 Trigger system of the CMS detector is based on the definition of trigger towers. A trigger tower in the electromagnetic calorimeter is defined as an array of 5×5 lead tungstate crystals. The size of these trigger towers matches the size of the physical towers in the hadronic calorimeter and there is a one-to-one correspondence between the trigger towers in the hadronic and electromagnetic calorimeters [48].

Electron Reconstruction in the Level-1 Trigger System

In the Level-1 trigger system [46] there are two separate reconstruction modes available, both based on a 3×3 trigger tower sliding window technique as depicted in Figure 2.17. The trigger algorithm described below is executed within each 3×3 trigger tower window. The window is consecutively centred on each trigger tower of the CMS calorimeters. One Level-1 trigger electron reconstruction mode is for non-isolated electrons, the second one for isolated electrons.

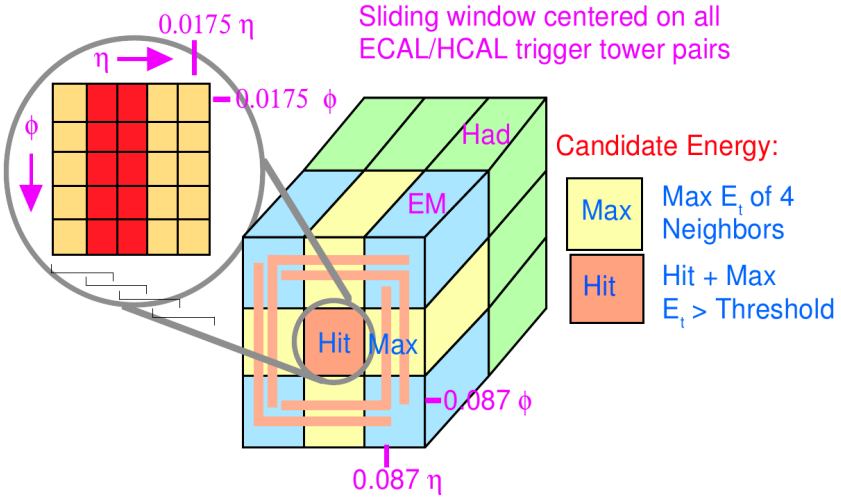


Figure 2.17.: Sliding window technique used for electron reconstruction in the Level-1 trigger [46].

In the non-isolated mode the electron reconstruction is based on a large energy deposit in one or two adjacent trigger cells of the electromagnetic calorimeter. Furthermore, the fine grain energy spread in the central cell is used as lateral shower profile. The fine grain bit veto is defined in a way that the highest energy adjacent strip pair has less than a certain fraction R of the tower energy. A typical value for R is 90 %.

An upper limit on the ratio of the transverse energy E_{T} of the deposits in the hadronic and the electromagnetic central cells, H/E, as longitudinal shower profile is applied. E is the energy measured in the electromagnetic, H is the energy deposited in the hadronic calorimeter. A typical upper threshold on H/E is 5 %.

For isolated electrons all conditions that hold for non-isolated electrons are also applied. In addition, there is a requirement on the electromagnetic isolation, defined as the transverse energy deposited in the electromagnetic calorimeter in towers surrounding the 3×3 window, as well as on the hadronic isolation, defined as fine grain bit veto and H/E veto on all eight nearest neighbours of the 3×3 window. Furthermore, at least one quiet corner is required which means one of the four five-tower corners has all towers below a certrain threshold. A typical value for this threshold is 1.5 GeV [49].

The transverse energy of the electron is estimated as the sum of the transverse energies in the hit tower added to the four highest transverse energy broad side neighbours.

For each calorimeter region the Level-1 electron candidates are sorted for their transverse energy E_{T} and the four highest isolated as well as the four highest non-isolated electron candidates are used for the trigger decision and are forwarded as Level-1 seed for the electron reconstruction in the High Level Trigger. The two collections of isolated and non-isolated electron candidates are mutually exclusive.

Electron Reconstruction in the High-Level Trigger System

In the High-Level Trigger clusters and superclusters [47] are reconstructed and a geometrical matching of superclusters above a certain threshold to Level-1 electron candidates is required [50]. The Island algorithm [18] describes the building of clusters and clusters of clusters, the latter are referred to as supercluster. Crystals above a certain threshold are used as seed. Adjacent crystals are examined and added to the cluster if there is a positive energy deposition registered in the crystal, the crystal has not been assigned to another cluster already and the previous crystal added in the same direction has higher energy. The same algorithm is applied on clusters instead of crystals to combine clusters to superclusters [51, 18, 47].

The supercluster is required to exceed a certain transverse energy threshold. Tracks in the pixel layers of the tracker that are compatible with a match to the calorimeter supercluster are used to distinguish electron candidates from photon candidates.

Quality criteria on the isolation and identification of the electron are applied to the reconstructed candidates and depend on the specific trigger. An example is discussed in detail in Section 5.1.1.

Jet Reconstruction in the Level-1 Trigger System

Similar to the Level-1 electron trigger system, jets are found in the Level-1 trigger employing a 3×3 calorimeter region sliding window method as shown in Figure 2.18.

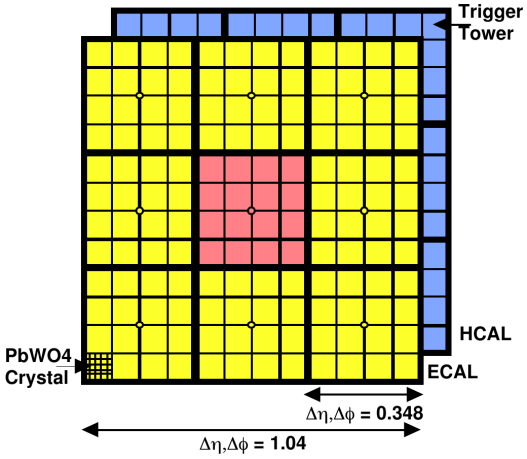


Figure 2.18.: Sliding window technique used for jet reconstruction in the Level-1 trigger [46].

Each calorimeter region consists of 4×4 trigger towers. The jet is characterised by the transverse energy E_{T} measured in this 3×3 calorimeter region. It is required that the transverse energy in the central region of the window is higher than in the eight neighbouring regions. The four central jets in the pseudorapidity region $|\eta| < 2.6$ and four forward jets in the region $|\eta| > 2.6$ with the highest transverse energy are selected for further processing. A third category of jets, τ jets is also reconstructed but not used in this thesis.

Jet Reconstruction in the High-Level Trigger System

Jets in the High-Level Trigger are reconstructed using the *anti-k_{T}* algorithm [52]. This algorithm is described in Section 2.4.6. In the CMS trigger system it is executed on calorimeter towers throughout most of the 2011 data taking period. Towards the end of the 2011 data taking in the "5e33" menu, the algorithm was also executed on *particle flow* objects. The offline particle flow event model is described in the following section. Online Particle Flow object reconstruction is a simplified version of the offline Particle Flow reconstruction, where only localised detector information is used to reduce the processing time in the trigger system.

2.4. Physics Objects Reconstruction within the Particle Flow Model

The *Particle Flow* (PF) model is used to reconstruct physics objects. This technique combines all the available detector information and yields improved resolution over classical methods which rely solely on individual subdetectors. Figure 2.19 shows a slice through the CMS detector and which subsystems are used to

detect which particles. Information coming from all the subdetector measurements is combined to reconstruct electrons, muons, photons, charged and neutral hadrons. A summary of the Particle Flow event reconstruction is given here based on the official CMS procedure [53].

All major subsystems of the CMS detector are involved in the Particle Flow scheme. In the silicon tracker charged particles can be measured down to 150 MeV/c and up to a pseudorapidity of $|\eta| < 2.6$. In the electromagnetic calorimeter photons can be reconstructed up to a pseudorapidity of $|\eta| < 3.0$. Furthermore, combining the hadronic and electromagnetic calorimeter, charged and neutral hadrons can be measured with an energy resolution of about 10% at 100 GeV. The energy measurement of charged hadrons benefits from the superior resolution of the silicon tracker [55] with $\Delta p_{\mathrm{T}}/p_{\mathrm{T}} = 2\%$ at 100 GeV/c up to a pseudo-rapidity of $|\eta| = 1.6$.

Electrons are reconstructed from a tracker track as well as multiple energy deposits in the electromagnetic calorimeter from the electron and from bremsstrahlung photons emitted by the electron. Muons are reconstructed from the information of the track in the muon system, the track measured in the inner silicon tracker and as minimum ionising particle in the calorimeter.

Missing transverse energy E_T^{miss} in the Particle Flow event reconstruction is defined as the modulus of the vector sum of the transverse momenta of all reconstructed Particle Flow objects in the event.

A big challenge in the event reconstruction at high luminosity is the background due to *pile-up* interactions. That means several physics collisions that take place in the detector or so close to each other in time that the detector response is not fast enough to resolve the two collisions.

The number of primary vertices in an event is a good indicative criterion on the amount of pile-up interactions that were registered in the detector in addition to the hard scattering event. The comparison of the true number of pile-up interactions in an event and the number of reconstructed primary vertices is presented in Figure 2.20 based on top quark simulated samples. The average number of primary vertices in the 2011 data is around seven as can be seen from Figure 2.21.

2.4.1. Primary Vertex Reconstruction in CMS

The first step in the event identification is the reconstruction of primary interaction vertices. These methods are independent of the Particle Flow event reconstruction.

Reconstructed tracks in the inner tracker are used to measure proton-proton interaction vertices. First, tracker tracks are selected, based on a criterion defining a maximum impact parameter significance for the tracks. In a next step, tracks which belong to the same interaction vertex are clustered together. A deterministic annealing algorithm [56] is used to assign tracks to a given interaction vertex by minimising a χ^2 and finding the most likely distributions of track-to-vertex assignments following the principle of maximum entropy. Finally, an adaptive vertex fit [57] is performed to determine the position of the interaction vertex. An adaptive vertex fit applies a fit to the vertex based on iteratively re-weighted Kalman

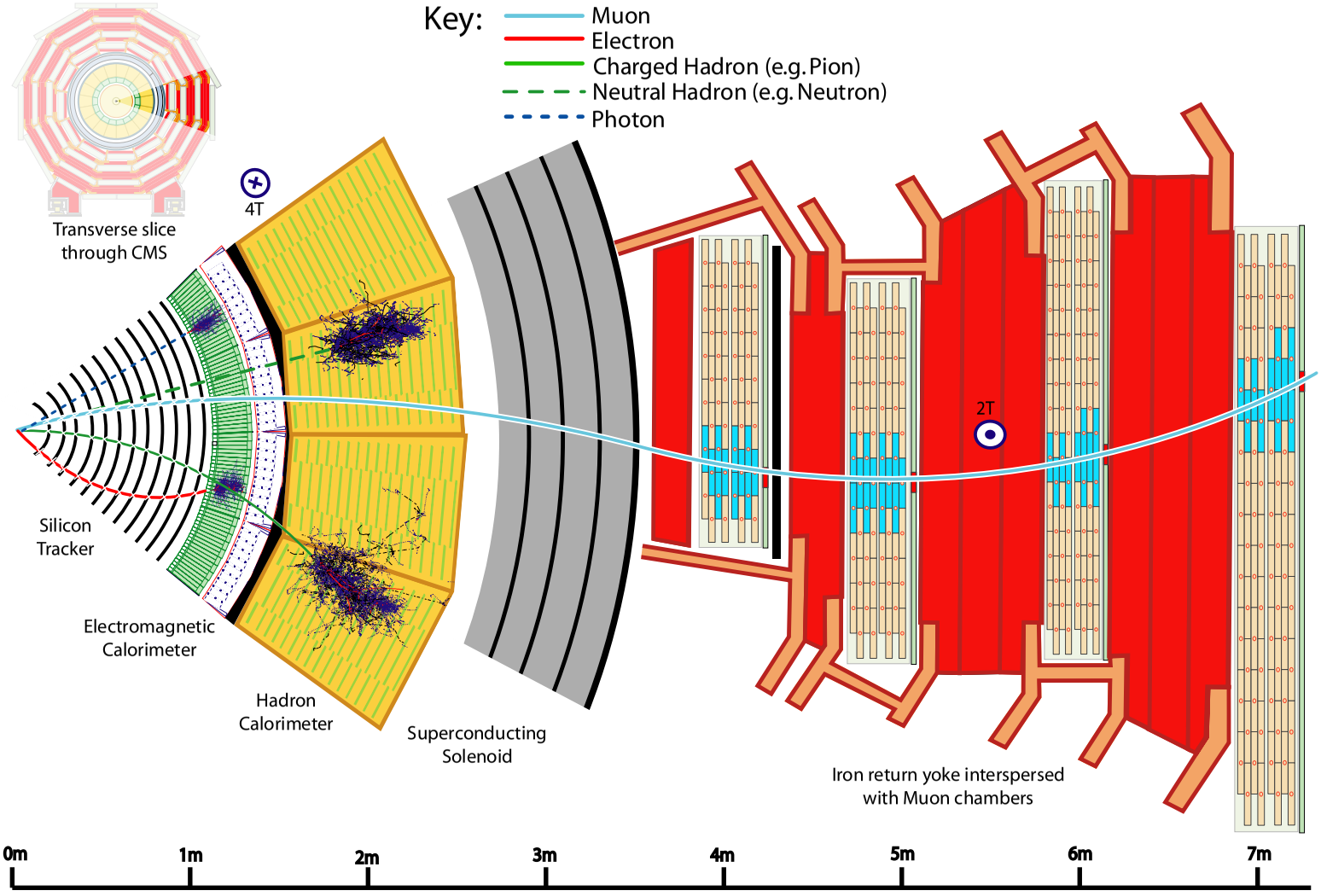


Figure 2.19.: Schematic view of a slice through the CMS detector in its barrel region. Electrons are reconstructed in the tracker and the electromagnetic calorimeter, photons in the electromagnetic calorimeter, neutral hadrons in the hadronic and electromagnetic calorimeter and charged hadrons in the calorimeters and in the tracker. Muons are reconstructed in the tracker, in the calorimeters as minimal ionising particles and in the muon system. Figure taken from [54].

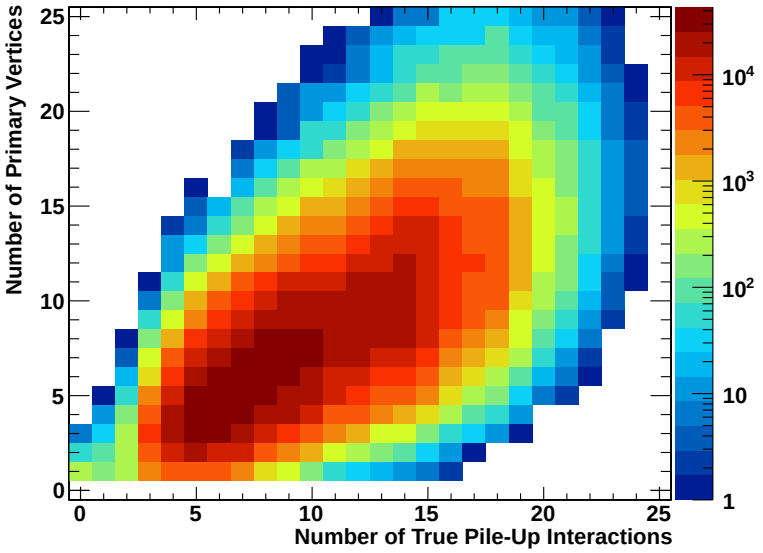


Figure 2.20.: Number of reconstructed primary vertices as a function of the number of pile-up interactions in an event based on top quark pair simulated samples.

filters [58]. The Kalman filter provides an efficient solution of the least-squares method and uses a form of feedback control. It processes the state at a given time to provide an estimate on given variables and on their uncertainties. In iterative steps the measurement is repeated using the estimate from previous states in the estimate of the current state [59]. The vertices in the collection are sorted according to the sum of p_{T}^2 of all tracks of the vertex.

Quality criteria on the fit to the vertex are applied, for example on the χ^2 value of the vertex fit or the z coordinate at the point of closest approach to the beamspot.

The number of primary vertices as indicative criterion for the amount of pile-up interactions for the 2011 data taking period is shown in Figure 2.21.

2.4.2. Tracking in Particle Flow

One of the steps in the Particle Flow event reconstruction is the reconstruction and handling of tracks in the silicon tracker. It is crucial for the identification of charged particles and allows the momentum as well as angular measurement of the particle at the production vertex before any deviation caused by the magnetic field.

The Particle Flow track reconstruction uses an iterative tracking procedure. The goal is a high track identification efficiency with a low fake rate. In a first iteration, very tight track selection criteria are applied. This leads to a moderate track reconstruction efficiency but with a negligible rate of fake tracks. The hits associated with the identified tracks are removed and in a next iteration tracks with

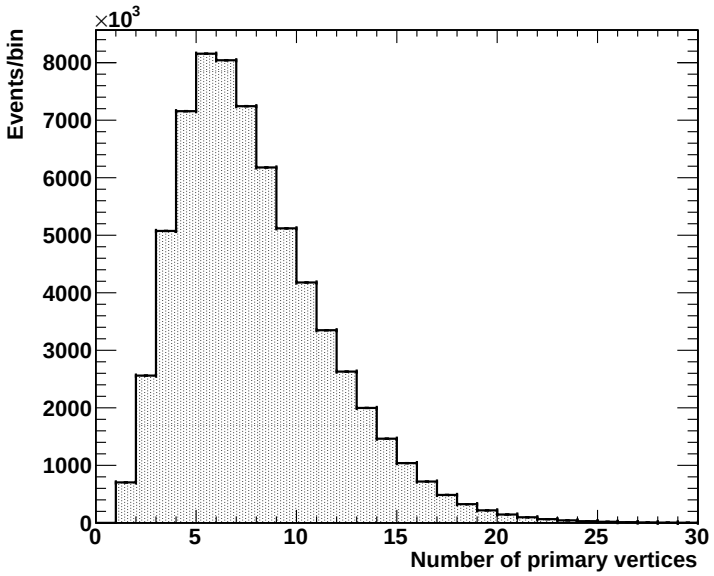


Figure 2.21.: Distribution of the number of primary vertices in an event as indicator of the number of pile-up interactions that were recorded in addition to the nominal proton-proton collision in the 2011 data.

looser selection criteria are reconstructed. This increases the track reconstruction efficiency. At the same time, the lowered amount of combinatoric reconstruction possibilities due to the removed hits from high quality tracks guarantees a low fake rate. After three iterations a track reconstruction efficiency of more than 90% for charged hadrons in jets with a negligible fake rate is achieved.

2.4.3. Calorimeter Clustering in Particle Flow

Calorimeter clustering in Particle Flow is used for the measurement of the energy and the direction of neutral hadrons and photons. The separation of charged particles and neutral particles is based on the measured energy. Furthermore, calorimetric measurements are crucial for the reconstruction of electrons and the accompanying bremsstrahlung photons. For charged particles with low quality tracks or tracks resulting from particles with a high transverse momentum the calorimetric energy measurement helps to improve the energy resolution.

In the process of calorimeter clustering in the first step *cluster seeds* are identified. That means a local calorimeter cell is observed with a measured energy above a certain threshold. In a second step *topological clusters* are constructed where cells are combined that have a measured energy above a certain threshold and share one side with a reconstructed cluster.

For cells that are common among several clusters the energy of the cell is shared according to the cell–cluster distance.

2.4.4. Linking of Particle Flow Object Constituents and Reconstruction of Physics Objects

Particle Flow objects are reconstructed by linking the different constituents that are described in the previous sections. The linking removes double counting from the various subdetectors.

Following the CMS *global muons* [60] definition, Particle Flow muons are reconstructed by linking tracks from the inner tracker to tracks in the outer muon system. A global fit in both systems is performed and muons with a normalised χ^2 value of the fit below the threshold of $\chi^2/\text{n.d.f.} < 10$ are accepted. The muon tracks as well as an estimate for the muon energy deposition in the calorimeter are removed from the list of object constituents.

In a next step, particle flow electrons are reconstructed as described in Section 2.4.5.

For the remaining tracks tighter quality criteria are applied. Together with the remaining calorimetric energy depositions they are used to decide if charged hardrons, neutral hadrons or photons are reconstructed. To reconstruct charged hadrons the energy deposit in a calorimeter cluster is compared to the transverse momentum of the assigned charged particle measured in the tracker. If they are compatible a charged hadron is reconstructed. If they are not compatible either a neutral hadron or a photon is reconstructed. If the dominant fraction of the energy of the particle is deposited in the electromagnetic calorimeter a photon is reconstructed, otherwise a neutral hadron is reconstructed.

2.4.5. Electron Reconstruction

Particle Flow electrons are a subset of so-called *Gaussian Sum Filter (GSF) Electrons*. GSF electrons are reconstructed based on two different seeding procedures, tracker driven seeding and electromagnetic calorimeter based seeding. The tracker driven seeding is used for the reconstruction of electrons with low transverse momentum and electrons within jets while the calorimeter driven seeding is optimised for electrons with a transverse momentum larger than 5 GeV.

In the calorimeter driven seeding, superclusters with a transverse energy $E_\text{T} > 4\,\text{GeV}$ are reconstructed as described in Section 2.3.2. The energy and position measured in the superclusters is back-propagated through the magnetic field to the inner layers of the tracker. There, the extrapolation is matched to pairs and triplets of tracker hits in the inner tracker layers, called track seeds. The trajectories of the electron candidates are reconstructed by a fit to the electron track with a Gaussian-Sum Filter [61] taking energy losses into account. This improves the reconstruction considering the short track length in the inner tracker and the energy loss caused by the emittance of bremsstrahlung in the tracker layers. A Gaussian-Sum Filter is a non-linear generalisation of the Kalman filter discussed in Section 2.4.1. In the Gaussian-Sum Filter all state vectors are weighted sums of Gaussian distributions, each Gaussian distribution modelling different degrees of hardness of

the bremsstrahlung in the tracker layer under consideration [61].

Calorimeter seed driven electrons are selected by applying a filter criterion in η and ϕ on the matching between tracker track and supercluster. In the tracker driven seeding approach, track seeds are selected using a multivariate technique. Electron candidates are kept in cases of calorimeter driven electrons where the geometric matching criteria are not fulfilled but the tracker track passes the multivariate tracker seed selection.

The Particle Flow electron reconstruction links calorimeter clusters measuring bremsstrahlung photons emitted by the electron to the electron itself. To achieve this, tangents from the inner tracker tracks to electromagnetic calorimeter clusters are calculated. A linking of calorimetric energy depositions in the preshower detector, the electromagnetic calorimeter and the hadronic calorimeter is done based on the condition that the cluster position in the more granular calorimeter is within the envelope of the less granular calorimeter. The tracker tracks and calorimeter energy depositions assigned to electrons are removed from the list of particle flow object constituents.

2.4.6. Jet Reconstruction in the CMS Experiment

In the CMS collaboration the default algorithm for the reconstruction of jets is the $anti{-}k_{\mathrm{T}}$ algorithm [52]. As part of the offline event reconstruction the algorithm is executed on offline reconstructed Particle Flow objects. In general, all Particle Flow objects in the event are considered for jet reconstruction. Prompt leptons that are used in the analysis have to be removed from the list of particles for the jet reconstruction to avoid double counting of objects. Quality criteria are defined to separate prompt leptons from leptons that are produced as decay-in-flight leptons within a jet. In this thesis, the prompt leptons must be isolated as discussed in Chapter 4. These leptons are then used for the definition of leptons used in the event selection as well as looser lepton veto definitions.

Charged particles that are assigned to a different production vertex and not to the primary vertex considered in a given physics event are removed from the Particle Flow objects collection in a procedure called *charged hadron subtraction* to compensate for pile-up effects. On the remaining Particle Flow objects the anti-k_{T} jet algorithm is executed to cluster jets.

The anti-k_{T} Jet Clustering Algorithm

The anti-k_{T} algorithm is a sequential recombination jet algorithm [52]. It is a fast, infrared and collinear safe jet algorithm, this means that the emission of virtual soft particles does not affect the reconstructed jets [9]. In previous infrared and collinear safe algorithms such as the `SISCone`, the k_{T} and the `Cambridge/Aachen` algorithms soft radiation can provoke irregularities in the boundaries of the reconstructed jets. On the other hand, it is desirable to work with soft-resilient jets as the shape knowledge can simplify theoretical calculations and the determination of jet

energy corrections to correct from the energy measured in the calorimeter to the energy of the originating jet on particle level. Furthermore, part of the momentum resolution loss caused by the underlying event and pile-up contamination can be eliminated [52].

Jet clustering algorithms are based on two distance measures,

$$d_{ij} = \min(k_{\mathrm{T}i}^{2p}, k_{\mathrm{T}j}^{2p})\Delta_{ij}^2/r^2 \text{ and}$$
$$d_{iB} = k_{\mathrm{T}i}^2,$$

where d_{ij} is the distance between two particles or pseudojets and d_{iB} is the distance between a particle or pseudojet and the beam B. $k_{\mathrm{T}i}$ is the transverse momentum and r the radius parameter of the jet clustering algorithm. The parameter p is called the power of the jet clustering algorithm and is $p = 1$ for the k_{T} algorithm, $p = 0$ for the `Cambridge/Aachen` algorithm and $p = -1$ for the anti-k_{T} algorithm. Δ_{ij}^2 is defined as

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

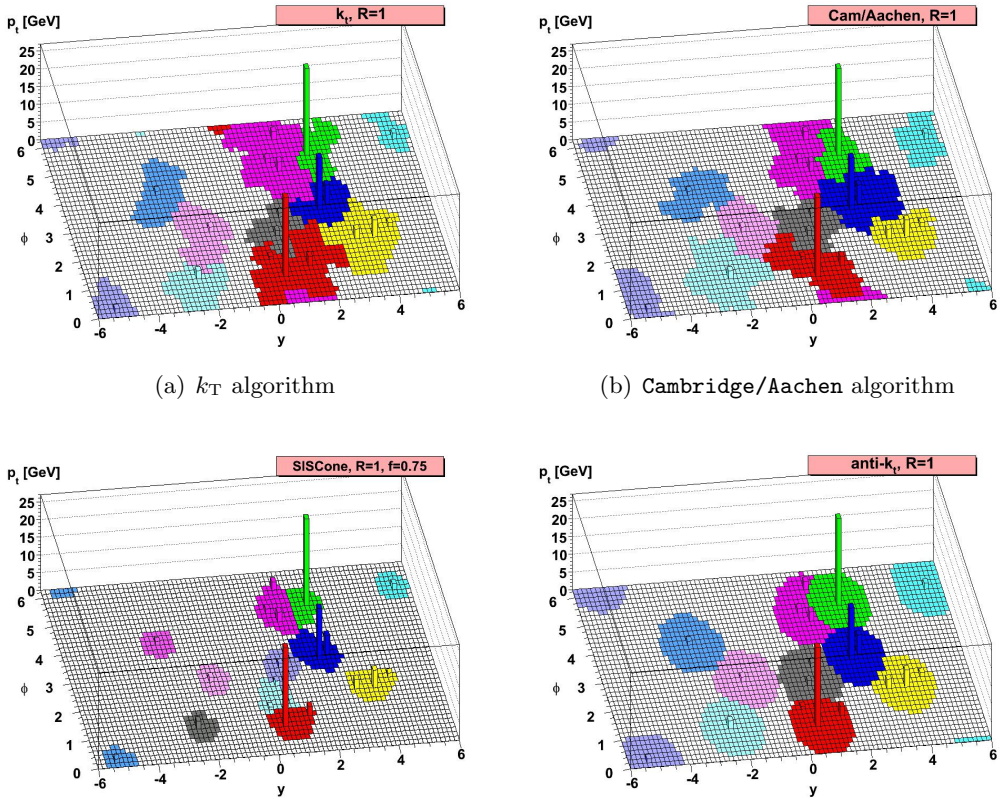
where y_i is the rapidity and ϕ_i is the azimuth of the ith particle.

Jet clustering is an iterative procedure. The smallest distances d_{ij} and d_{iB} are identified. If the smallest d_{ij} is smaller than the smallest d_{iB} the objects i and j are recombined by summing their four-momenta. Otherwise object i is considered a jet and removed from the list of objects. This procedure is repeated until no more objects are left.

If a hard particle does not have any hard neighbours within a distance of $2 \cdot r$ all soft particles around the hard particles will be clustered into the jet and the resulting jet shape is perfectly conical. Two hard particles within the distance $r < \Delta_{12} < 2 \cdot r$ result in two hard jets of which at most one is perfectly conical and the second is partly conical, whereas two hard particles within the distance $\Delta_{12} < r$ result in one single jet which is in general not perfectly conical due to the combination of two jet cones. The jet boundaries are not affected by soft particles. A comparison of jet shapes as a result from different jet reconstruction algorithms can be found in Figure 2.22. This analysis uses a radius parameter of $r = 0.5$. This choice of the parameter is large enough so that it reduces the amount of particles that are not clustered into a jet which result from the hadronisation of a given quark but on the other hand it is not too large that two separate jets are merged.

Jet Energy Calibration

In the CMS detector reconstructed jets have not the full energy of the initiating parton or particle jet. Energy is lost in jets due to charged hadrons of which the trajectory is bent and which do not reach the calorimeter, electromagnetic noise, threshold effects of the calorimeter cells, the non-linear response of the calorimeter as well as energy deposited in the absorber material of the calorimeter [9].



(a) k_T algorithm

(b) `Cambridge/Aachen` algorithm

(c) `SISCone` algorithm

(d) anti-k_T algorithm

Figure 2.22.: Comparison of jet shapes in different jet clustering algorithms [52]. The clustering algorithms k_T in Figure 2.22(a) and `Cambridge/Aachen` in Figure 2.22(b) show very non-uniform jet shapes. The `SISCone` algorithm in Figure 2.22(c) returns rather uniform jet shapes but is unattractive due to the large time comsumption during its execution. The anti-k_T algorithm in Figure 2.22(d) combines the advantages of cone-based algorithms with those of clustering algorithms.

Figure 2.23.: Factorised scheme of jet corrections of the CMS experiment. The first three levels of jet corrections are mandatory for every analysis and perform an offset correction, the flatening of the jet response over the whole pseudorapidity range and the absolute correction of the transverse momentum. Higher order corrections can be applied to correct back to the parton level kinematics.

A factorised jet-energy-correction scheme [62] is employed in the CMS experiment; it consists of several independent levels of jet calibration steps as can be seen in Figure 2.23. The jet energy calibration steps from Level-1 to Level-3 are mandatory for every analysis. Higher level corrections are only necessary or advantageous for certain analyses, for example if the jet energy must be calibrated back to parton level. As this is not necessary for the measurement of the top quark production cross section, only the calibration levels one to three are considered. The Level-1 calibration step performs a constant shift of the jet energies. This becomes necessary to take pile-up effects in the event into account. Level-2 relative jet calibrations are applied to make the jet response flat over the full pseudorapidity range up to $|\eta| < 5.0$ to match the response in the barrel region of $|\eta| < 1.3$. In the third level of calibration absolute corrections to the jet transverse momentum are applied to correct the jet energy back to particle jet level, that means to the energy of the jet reconstructed from the individual hadrons. This kind of correction is derived from Z + jets and γ + jets events.

There are several sources of systematic uncertainties on the derivation of the jet energy calibrations [65]. The dominant uncertainties for jets with low transverse momentum are the pile-up estimation uncertainty, for medium transverse momentum jets the uncertainty due to the jet flavour mix in multijet events and for high transverse momentum jets the uncertainty of the extrapolation of the derived corrections in the medium energy regime to high transverse momenta. Figure 2.24 shows the uncertainties as a function of the jet transverse momentum in the 2011 data taking period. The jet energy resolution has been measured on 7 TeV data and is also shown in Figure 2.24.

Pile-Up Subtraction in Jets

As described at the beginning of Section 2.4.6 charged hadrons are removed to compensate for pile-up effects in a procedure called *charged hadron subtraction*. To compensate also for the neutral component of pile-up interactions, for each event the k_{T} jet algorithm with a cone size of $r = 0.6$ is executed. It can be shown that these jets tend to organise a uniform background of soft particles into

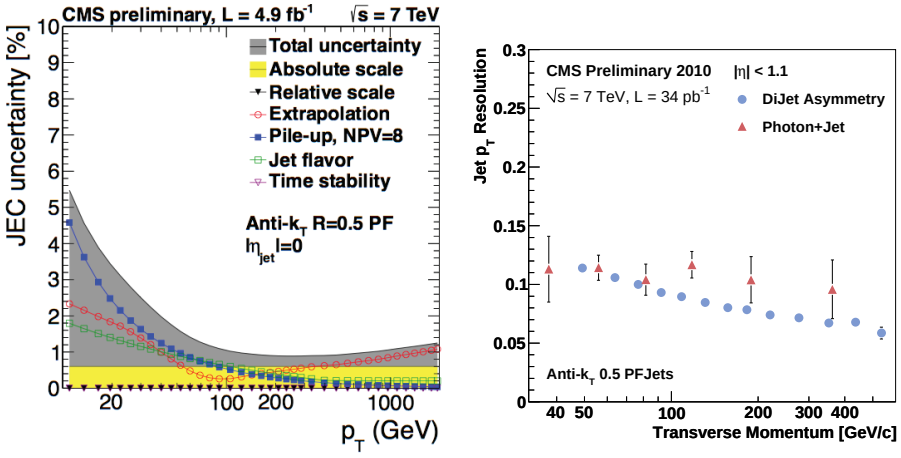


Figure 2.24.: Jet energy scale uncertainties in 2011 data [63] and jet energy resolution at $\sqrt{s} = 7\,\text{TeV}$ [64].

structures [66, 67]. If one divides the transverse momentum p_T of such a jet by its *area* A, one receives a measure of the average diffuse noise ρ in the event. The jet area is measured on the rapidity y and azimuth ϕ circle where particles are clustered into a given jet and can be approximated by $\sim \pi r^2$, where r is the jet distance parameter. To get an averaged value of ρ for the entire event one averages the ρ values of each individual diffuse jet as

$$\rho = \text{median}\left[\left\{\frac{p_\text{T}^j}{A_j}\right\}\right],$$

for each k_T jet j in the event. This value ρ is used to subtract the pile-up energy contribution to the jet transverse momentum as part of the Level-1 jet energy corrections.

2.4.7. Photon Conversion Rejection

In the CMS detector prompt electrons suffer from a background by electrons from converted photons. There are several methods available [50] to reject this type of background. The impact parameter of positron-electron pair tracks from converted photons is on average greater than that of prompt electrons. In the same context, prompt electron trajectories start at the beam line, therefore hits are recorded in the innermost pixel layers. If there are one or more missing inner tracker hits this is a sign of a possible photon conversion that took place inside the tracker. A third, more complex method is based on the search for a conversion partner-track.

Tracks from the resulting electron–positron pair from a photon conversion are parallel to each other at the decay point and remain so in the r–z plane. All tracks

within a cone of $\Delta R < 0.3$ around the electron track are selected. For each of these track the quantities $\Delta\cot(\theta) = \cot(\theta_{\text{sel. track}}) - \cot(\theta_{\text{el. track}})$ and *dist* as the two-dimensional distance in the x–y plane between the two tracks extrapolated to the point where the two tracks would be parallel are calculated. A selection of electrons with $|\Delta\cot(\theta)| < 0.02$ and $|dist| < 0.02\,\text{cm}$ rejects largely electrons from photon conversions with a rejection of up to 49.3% of all electrons from photon conversions and with only a moderate prompt electrons rejection of less than 5%. The prompt electron rejection is studied in detail in Section 5.2.2.

2.5. Luminosity Measurement in CMS

It is crucial to have a precise measurement of the amount of data recorded by the CMS experiment to measure the cross-section of a certain process. The currently most precise method for the luminosity determination employed in the CMS experiment is the *pixel cluster counting method* [68].

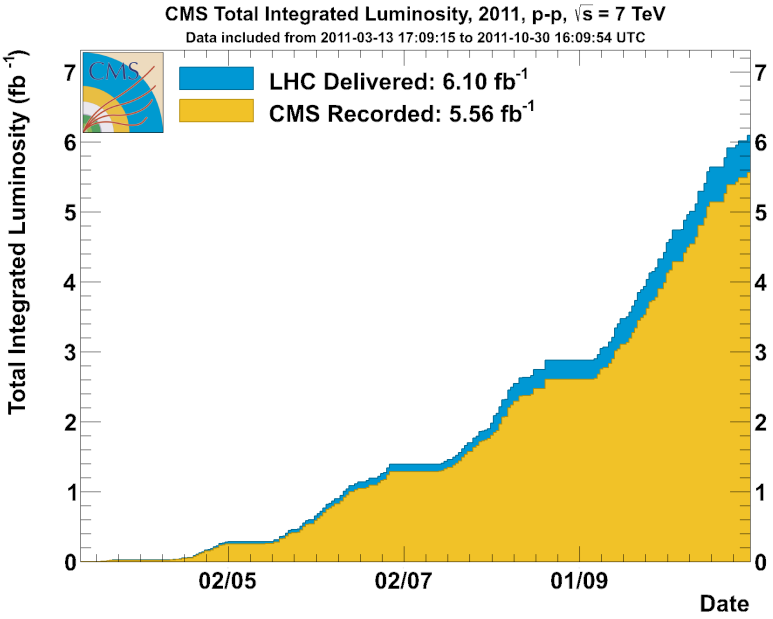


Figure 2.25.: Amount of recorded data as a function of time for the year 2011 [69].

As explained in Section 2.2.3 the CMS detector features 66 million pixels in the pixel detector. The number of pixel cluster hits measured in the pixel tracker is approximated as a linear function of the number of interactions per bunch crossing. As there are per bunch crossing always some proton-proton interactions which lead to a certain amount of pixel-cluster hits this is a good measure for the instantaneous luminosity in a given time interval. This time interval is chosen as the length of one *luminosity section* which equals 23.3 seconds.

To measure the luminosity a *zero-bias trigger* is employed, that means a trigger with the only requirement that the proton bunches pass through one another. The integrated luminosity is measured based on the formula

$$\langle N_{\text{cluster}} \rangle = \frac{\sigma_{\text{cluster}}}{f} \cdot L,$$

where $\langle N_{\text{cluster}} \rangle$ is the mean number of pixel cluster hits per trigger, f is the LHC orbital frequency of $f = 11,246\,\text{Hz}$, L is the instantaneous luminosity and σ_{cluster} is the effective pixel cross section. σ_{cluster} is determined in *van der Meer scans* [70]; the sweeping of beams transversely across each other and counting the interactions. With this knowledge the equation is solved to measure the luminosity per time frame. The systematic uncertainty on the luminosity determination is 2.2 % [68]. The dominant sources of systematic uncertainty are scan-to-scan variations in the measured cross section, the uncertainty on the afterglow correction, which removes in average 2.8 % of the integrated luminosity per luminosity section, and the evolution of the beam width over time. Late-arriving particles and energy originating from activated detector material is referred to as afterglow and has to be corrected for in the pixel detector.

The amount of recorded data as a function of time for the year 2011 is shown in Figure 2.25. The LHC delivered more than $6\,\text{fb}^{-1}$ of data. $5.6\,\text{fb}^{-1}$ of data was recorded with the CMS detector during operational periods of the detector. $5\,\text{fb}^{-1}$ of this data has been certified for physics analysis and is of highest quality. For this analysis $4.6\,\text{fb}^{-1}$ have been used. $0.4\,\text{fb}^{-1}$ are not used because a part of the data was recorded with a trigger menu very different to the ones used in the remainder of 2011 and due to processing inefficiencies on the computing resources.

3. Software Tools

With the enormous amount of data of more than 1 Gb per minute recorded by each of the LHC experiments new software based processing tools had to be developed for the analysis and simulation of events. In this chapter the process of Monte Carlo event generation will be introduced. After an overview of the general software tools widely used in high energy physics, a summary of the CMS software framework (CMSSW) will be given. This includes the CMS detector simulation used to describe the propagation of particles through the CMS detector. The chapter concludes with the introduction of the LHC computing grid, a world wide network of computing centres that has been developed and advanced in the scope of the LHC experiments.

3.1. Monte Carlo Event Generation

The purpose of Monte Carlo event generators is to simulate physics events as detailed as could be observed by a perfect detector [71].

The event simulation is factorised into several steps. A schematic view of a typical parton-parton interaction in high energy physics is shown in Figure 3.1. Two partons, that means quarks or gluons, as constituents of the colliding protons interact with a certain probability for a given momentum transfer. This is modelled by parton distribution functions which are measured from data. One of the commonly used PDF libraries at production stage in CMS is the CTEQ 6.L library [72]. In a next step the matrix element for the hard process is calculated analytically, typically with a tree-level matrix element generator like MADGRAPH as described in Section 3.1.1 or a higher order event generator like POWHEG [73, 74, 75] or MC@NLO [76]. In the following hadronisation process the coloured partons form jets of colourless hadrons and through subsequent decays also photons and leptons. The fragmentation is usually modelled by a showering and hadronisation event generator (SHG) like PYTHIA [71, 77] as described in Section 3.1.2 or HERWIG++ [78]. The decay of the final state particles is described either by the same SHG or by a specialised package like for example Tauola [79] which simulates the production and decay of tau leptons.

The events that are to be simulated show fluctuations according to the quantum mechanics nature of the Standard Model of particle physics. One of the big challenges in the event generation is the multiparticle production with on the order of 10 to 100 particles per simulated event, most of them as part of jets as described in Section 1.3.2. They arise from higher-order corrections to the tree-level matrix element of the hard process. There are corrections of the bremsstrahlung type of

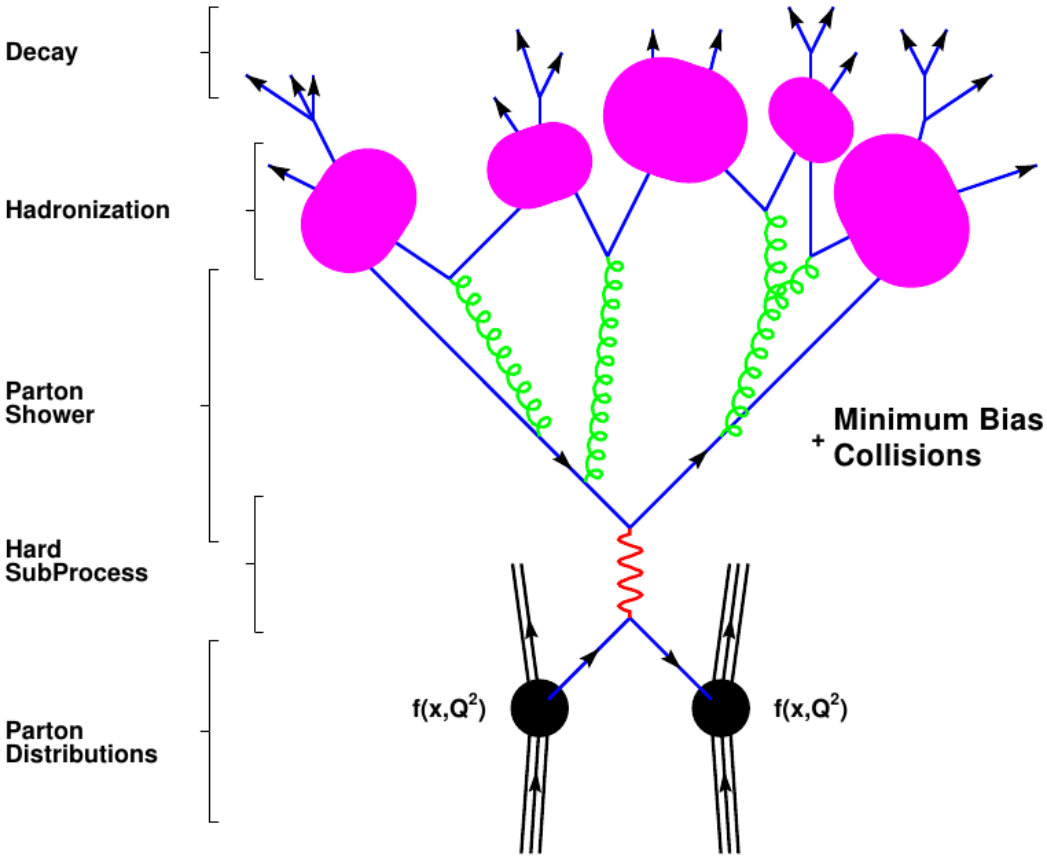


Figure 3.1.: Schematic view of a typical process in high energy physics. Two partons interact based on a certain probability for a given momentum transfer as described by parton distribution functions. As a result, the hard subprocess of the interacting proton consituents takes place. This can lead to the emittance of quarks and gluons which results through the radiation of gluons in a parton shower. Finally, the partons of the parton shower form hadrons which decay.

radiations with the emission of final-state partons. Those corrections are large relative to photon radiation due to the comparatively large value of the strong coupling constant α_s and the existence of gluon self-interactions. Divergencies cancel when virtual and soft corrections in combination are taken into account. Finally, the quark-gluon confinement leads to the hadronisation process which is modelled on a probabilistic basis [71].

The addition of corrections means a convolution of probability density functions which is the equivalent of a high-dimensional integration. An analytical solution to this problem is extremely difficult and time intense. The Monte Carlo method [80] introduced by John von Neumann and Stan Ulam for the description of the behaviour of neutrons inside a nuclear reactor provides a remedy to this problem. Random numbers are generated according to a given probability density function and can be interpreted as simulated measurements of a process following the probabilistic behaviour under study. From this, the probability that a certain process yields a result in a given region of phase space is estimated which is the equivalent to an integration of the probability density function. Hence, the problem of a high-dimensional integration is reduced to the generation of random numbers following a certain high-dimensional function. It can be shown that when using the Monte Carlo method the statistical precision of the estimate for the integral reduces with $1/\sqrt{N}$ where N is the number of generated random points in the phase space described by the probability density function [81].

3.1.1. MadGraph

MadGraph 5 [82] is an open source matrix element generator for particle physics processes written in Python [83]. It generates all tree-level Feynman diagrams for a process, for example decays or $2 \to n$ scattering processes, following user specific requests in terms of initial and final state particles and given constraints. Matrix elements at the tree-level can be generated for any Lagrangian based model implemented in FeynRules [84]. MadGraph outputs computer code in C++ [85], Fortran [86] or Python which can be used for the evaluation of helicity wavefunctions and amplitudes at a given phase space point to calculate cross sections or decay widths and to generate simulated events, for example by means of the MadEvent [87] package.

In addition, MadGraph allows to produce the output of the Feynman diagrams of a given process as images. In this thesis, simulated samples in MadGraph are used for the processes of top quark pair events, W + jets and Z + jets events.

3.1.2. Pythia

Pythia 6.4 [71] is a showering and hadronisation event generator based on the JetSet package [88] for hadron shower simulation and the Fortran based Pythia event generator merged under the label of Pythia. JetSet has been developed in the Lund group in the late 1970s for studies of the hadronisation process. In this

context, the *Lund string fragmentation model* for the parton shower description had been developed which was very successful in the description of processes at e^+e^- colliders. In the Lund fragmentation model colour flux tubes connect divergent partons and the potential energy between the partons increases until the creation of new partons is energetically advantageous. This creation process of new partons is iterated until a Q^2 cut-off is reached.

The PYTHIA Monte Carlo generator features over 300 hard processes, mainly with two partons in the initial state and one or two partons in the final state. Currently, a new version of PYTHIA is developped in C++ under the label of PYTHIA 8 [77]. As the amount of processes available in PYTHIA 8 is still smaller than in its FORTRAN counterpart, PYTHIA 6.4 is still the most often used version of this package.

3.1.3. PowHeg

The POWHEG [73, 74, 75] method is a prescription for interfacing next-to-leading-order calculations with parton shower generators and is used in the CMS collaboration to generate single top quark events. POWHEG uses exact next-to-leading-order matrix elements for the event generation. It can be interfaced with any showering and hadronisation event generator. In the CMS collaboration POWHEG is interfaced with PYTHIA. POWHEG is a generalised framework used to implement the POWHEG method. In this thesis, single top quark samples simulated by POWHEG are used.

3.2. Geant

GEANT 4 [89] is a toolkit for the simulation of the propagation of particles through matter and the interaction between the particles and matter. It provides a wide range of features, including tracking of particle trajectories, the possibility to design user-specific geometries, the application of various physics models and the simulation of particle hits in different subdetectors. The particle hits are then propagated to a simulated read-out system of the detector in a process called digitisation. GEANT 4 is used in the CMS collaboration for the full detector simulation as described in detail in Section 3.4.3.

3.3. Root

The software data analysis framework ROOT [90] has been developed in 1995 by René Brun and Fons Rademaker in the context of the NA49 project [91] at CERN. It is the successor of the PAW [92] framework which is a FORTRAN based analysis framework that has been widely used in particle physics in the past. With the NA49 experiment the limits of PAW were reached because it was not longer suitable for the large amounts of data to be analysed. Today, ROOT is the predominantly used framework in particle and nuclear physics.

ROOT is an object oriented framework developed in C++ and hence it is very easy to reuse code and extend it transparently within the framework. The CINT [93] command line interpreter and script processor, developed by Masa Goto in Japan, is incorporated in the ROOT framework and allows for the analysis user interactive development and analysis of physics data.

ROOT provides all features necessary for routine physics analyses. Multidimensional histograms are managed, functional fits are performed and input/output from and to mass storage is handled efficiently along with data compression features. For large scale computing power intense analyses also threaded programming or fully parallelised usage of computing clusters becomes possible with the PROOF [94] part of the framework. Furthermore, advanced features like socket and network communication are available.

ROOT features also the ROOFIT library [95]. This library provides extensive tools for the statistical modelling of physics processes and a framework to perform likelihood fits, statistical tests and generate pseudo-experiments.

3.4. The CMS Software Framework

The CMS software framework referred to as CMSSW is a modularised collection of software [18]. It contains an event data model [96] (EDM), services for the detector simulation using GEANT 4 [89], calibration, alignment of the detector and the reconstruction of physics objects. Furthermore, it provides for analysis users an extensible environment for analysis code development and application. The same software is used to process online data and simulated events.

The event data model is centred around the concept of an *event* and provides C++ object containers for RAW as well as reconstructed (RECO) data. An *event* is formed during the data acquisition step for all events that pass the trigger selection as described in Section 2.2.7. The *event* container holds all detector data for a given event as well as products of this data, for example reconstructed physics objects such as electrons or jets. In addition metadata is stored, for example conditions and calibrations used for data taking. The latter are accessed from a central database which uses as identification key the CMS *global tag* [97].

CMSSW is configured with configuration files in PYTHON format. For each module that is used, a set of parameters is specified to control the behaviour of the given module. Furthermore, the control flow through the individual modules in a sequence is specified. There are various predefined types of modules in CMSSW.

3.4.1. Modules in the CMS Software Framework

Six different types of modules are currently foreseen in CMSSW. They take different roles in the data processing and managing; an overview is given below.

Source Modules

Source modules are responsible for the read-in of data of various sources. A source is for example a Monte Carlo simulation generator as described in Section 3.1 or a ROOT file on a hard disk. The data is read-in for further processing in other modules.

EDProducer Modules

If data must be added to a given *event* an `EDProducer` module is employed. `EDProducer` modules are the most commonly used modules in the CMSSW framework to write out additional information to the event content. An example is a jet algorithm producing a jet collection for further processing.

EDFilter Modules

`EDFilter` modules test an event for certain conditions and return a boolean decision if an event passes given criteria. If an event does not pass the criteria, the processing of the module sequence is stopped for the event under study.

EDAnalyser Modules

The most important module for the analysis user is the `EDAnalyser` module. Users can implement analysis code inside this type of module, for example the filling of histograms or the generation of lightweight output files in the ROOT file format.

EDLooper Modules

`EDLooper` modules allow multi-pass looping over input data from a given source. An example for the application of this type of module are track-based alignment studies.

Output Modules

After the processing of an event is finished, `Output` modules allow the write-out of the generated information to a mass storage device. The output is in the standardised CMSSW specific ROOT file format and can be read-in again with CMSSW or ROOT directly.

3.4.2. Workflow in the CMSSW Framework

The workflow within the CMSSW framework is illustrated in Figure 3.2. Events are input to the framework from a source, for example from the detector with a positive High-Level trigger decision or from a Monte Carlo generator and a RAW event is created. In the next step, for simulated events the read-out from the detector in the read-out electronics is simulated, the digitisation. For data measured in the

experiment this is what is read-out from the data acquisition system. In the next step physics objects are reconstructed and necessary calibrations are applied. The event is then stored to a mass storage device in the CMSSW ROOT file format.

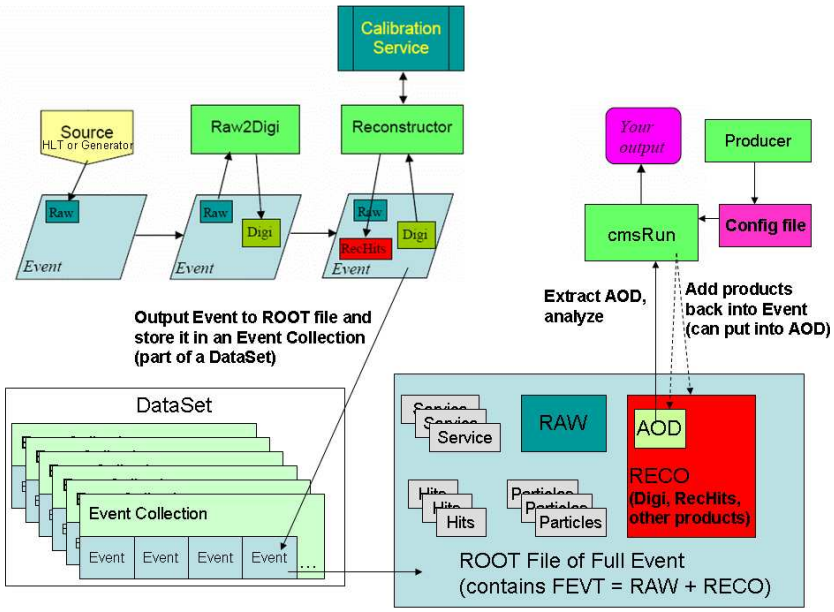


Figure 3.2.: Workflow in the CMSSW framework [98]. The step *Raw2Digi* is performed only in simulation and *Digi* is the equivalent of the recorded detector output.

The composition of the event content is illustrated on the lower right of Figure 3.2. Depending on the definition, the event contains the RAW information from the detector, all reconstructed objects RECO or a subset thereof, the AOD information, which is the collection of all reconstructed objects that are necessary for the physics analysis.

Analysis users read-in this event content, as shown on the upper right of Figure 3.2, and process it as described in Section 3.4.1.

3.4.3. The Full CMS Detector Simulation

The CMS detector simulation [99, 100] is an object oriented framework based on GEANT 4 [89, 101]. A precise simulation of the CMS detector, including all subdetectors and the read-out thereof is crucial for design optimisations of the detector – also with respect to possible upgrades –, calibration and alignment of the detector, object identification and physics analyses.

The full CMS detector simulation features a simulation of all CMS subdetectors, including forward detectors. It simulates the sensitive behaviour of the detectors, creates hit collections and performs the digitisation of the event.

The result of the simulation is stored as various `EDProducts` as part of the CMSSW *event*. The CMS detector simulation is extensible and configurable to adapt to new detector conditions. Important aspects in the development have been the reproducibility of events, hence an automated recording of the provenance information as well as the immutability of events, and the extensive validation of the simulation. Furthermore, tuning of the simulated detector response has been done based on physics collisions and test beam data.

The event processing chain is divided into several steps, allowing it to be interrupted at a given step and the current state can be saved. This allows for example a subsequent event mixing to account for different pile-up conditions. The pile-up events are simulated separately from the physics events and the simulation outputs are then merged in an additional step.

Events generated with a Monte Carlo event generator as described in Section 3.1 are converted to the GEANT 4 event format. The CMS simulation uses the CMS Detector Description services and *DDD/XML* description files from the CMS geometry package in GEANT to simulate the various CMS subdetectors. Multiple sets of detector *conditions* are available to account for the detector setup, state and alignment at a given time. The magnetic field is simulated independently by the CMS magnetic field simulation which allows for the choice of the field type and the configuration of the propagation of particles in the magnetic field. In a last step, the digitisation of the detector is performed. For each particle hit simulated in the detector the digitisation is performed by simulating the electronics read-out, taking into account energy loss, diffusion, noise and the coupling between different channels.

The output of the simulation is stored in `EDProducts` which are then used for physics objects reconstruction as described in Section 2.4. A considerably faster simulation method [102] which is less accurate and uses approximations and parametrisations, instead of detailed particle trajectories and showering with GEANT 4, is also available but is not employed in this analysis.

3.5. The Worldwide LHC Computing Grid

The amount of data recorded with the LHC and the necessary amount of simulated samples can no longer be handled by a single or only a few central computing centres. This point is valid in terms of storage capacity and in terms of computing power necessary to produce simulated samples, to reconstruct simulated events and collision data events and to process events for analyses by users.

As a solution to this problem, the LHC Computing Grid [96] (LCG) has been developed. The LCG is a distributed, hierarchical system of computing centres around the world. The workload and storage usage is distributed over those centres.

3.5.1. Organisational Structure of the Grid

The LCG is layered in different hierarchies and the participating computer centres are classified as Tier-0, Tier-1, Tier-2 and Tier-3 centres. The importance of an individual centre for the overall work of the LCG decreases with increasing Tier status. On the other hand, the number of centres increases with their Tier status as can be seen in Figure 3.3. There is one Tier-0 centre, there are seven Tier-1 centres, more than 50 Tier-2 centres and more than 60 Tier-3 centres that support the CMS software environment.

The most crucial computing centre is the Tier-0 centre located at CERN. The Tier-0 centre is used to store data recorded by the experiments located at the LHC. Furthermore, a prompt reconstruction of recorded data is performed at the Tier-0 centre and a copy of the reconstructed data is then distributed via high speed links to Tier-1 centres. Also the central coordination of the production of simulated samples is coordinated by the Tier-0 centre.

There are seven regional Tier-1 centres as part of the LCG that support the CMS environment and span the globe. The purpose of Tier-1 centres is to provide large amounts of storage capacity for simulated samples and data recorded in the experiments. Tier-1 centres are also employed for the centrally coordinated production of simulated samples, data skimming, reprocessing and selection activities.

Tier-2 and Tier-3 centres are mainly dedicated to user activity and the performance of analysis tasks.

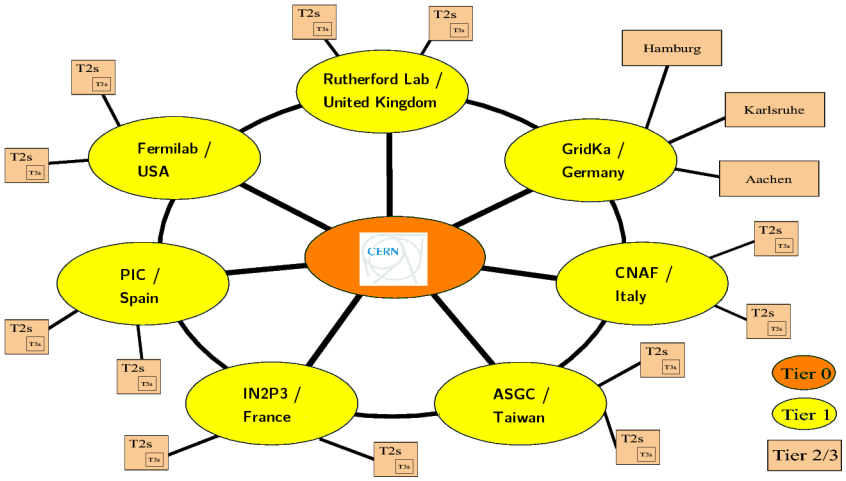


Figure 3.3.: Grid centre hierarchy. The central Tier-0 centre at CERN is connected with high speed links to the Tier-1 centres.

3.5.2. Grid Workflow

A job must pass through each of the steps in Figure 3.4 while being processed on the LCG. On the user interface an analyst develops a program. He creates there a

so-called Grid proxy which allows the authenticated access to LCG resources. The job can then be submitted via the Grid middleware to the *Workload Manager*. The latter searches for compatible computing resources on the grid by accessing the *Data Aggregation Service* (DAS) database [103] and submits the job for processing to the *Computing Element* of a local site. After the job finishes, output files can be transferred to an external *Storage Element*. To transfer status information to and from the job on the local site the *input* and *output sandbox* can be used.

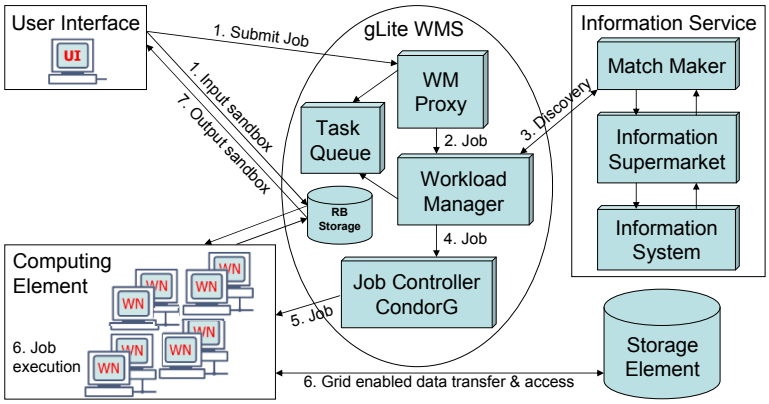


Figure 3.4.: The workflow of a typical Grid job. Taken from [104].

3.5.3. The Tier-2 Centre T2_BE_IIHE

One of the two regional Tier-2 centres for Belgium is the site T2_BE_IIHE. It is associated to the French Tier-1 centre T1_FR_CCIN2P3 [105]. Top quark physics analysis users associated to Belgian institutions have direct access to the computing environment of T2_BE_IIHE. The site features eight portal machines to provide access to the analysis users. Grid jobs can be submitted from the portal machines. The processing of all simulated and data samples relevant to this analysis takes approximately one week. As part of these jobs only relevant information for the analysis is extracted from the files centrally provided by CMS and smaller files in the ROOT format are produced. These smaller ROOT files can be processed on a local computing site.

At T2_BE_IIHE, a direct submission of computing jobs to the local Grid cluster is possible through the PBS job submission system [106]. 1800 CPU cores are available for job processing. One petabyte of storage capacity is available in terms of dCache storage pools [107]. All computing resources are operated under the CENTOS operating system [108].

For each analysis user it is possible to submit arbitrary amounts of jobs to the computing cluster while at most 400 jobs will be processed at the same time. The processing of 250 million physics events, the equivalent of the nominal simulated

samples relevant for top quark physics and all collision data recorded in the 2011 data taking period, takes with a customised ROOT framework less than 60 minutes outside of peak usage times.

Storage capacity and CPU resources are sufficient to allow for fast turn-around times. Performance improvements could be achieved by providing high-performance MYRINET [109] or INFINIBAND [110] network connections between local file servers and computing nodes. Currently, a *network file system* [111] based approach is chosen which limits the processing speed of events.

Part II.

Event Selection and Selection Efficiencies

4. Event Selection for Semi-Leptonic Top Quark Pair Events in the Electron + Jets Channel

An event selection is applied which reduces the background contribution as much as possible while retaining a signal selection efficiency as high as possible to measure the semi-leptonic top quark pair production cross section most precisely. One optimises the selection in a way that the overall uncertainty of the measurement is minimised.

The event selection is applied in several steps. First, events are selected by a certain trigger algorithm as described in Section 2.2.7. In a second step, offline selection criteria are applied to select individual objects as part of the event and events themselves.

The data considered for this thesis includes all data recorded by CMS after the first technical stop in spring 2011 in the CMS run range 165088 to 180252 which equals an integrated luminosity of $4.6 \pm 0.1\,\mathrm{fb}^{-1}$. Earlier runs from the 2011 data taking period are not used because the trigger setup was very different compared to the rest of the data taking in 2011. This leads to a loss of $0.2\,\mathrm{fb}^{-1}$ of integrated luminosity. The final results are not affected by this because the precision on the measurement is not limited by statistical uncertainties but instead by systematic uncertainties.

The offline selection criteria are explained in Section 4.1 including the determination of the lepton selection and isolation efficiency in data and simulation. In Section 4.2 the application of event weights to simulated events is discussed to tune the simulation to measured collision data.

4.1. Offline Event Selection of Electron and Jets Events

There are several physics processes that mimic top quark pair events in the semi-leptonic decay channel in the detector. The main background contribution is W + jets production where jets are produced in association with a W boson. A similar background contribution consists of Z + jets events where one of the two leptons of the Z boson decay is not reconstructed or does not fulfil certain quality criteria. Multijet events can be misidentified as top quark pair events if one jet fakes a lepton or a lepton is produced in the decay of one of the hadrons as part of the jet. In addition, single top quark events with additional jets have the same signature as top quark pair events. The goal is to find a selection that separates those background

Table 4.1.: Standard Model predictions for top quark pair production cross section and relevant background processes [113, 114, 115, 116, 117]. Multijet events are simulated in two orthoganal sets of exclusive regions of momentum transfer in the hard interaction $\hat{p}_{\mathrm{T}}$. At production time it is ensured with filters that a significant energy fraction is registered in the electromagnetic calorimeter or that the event contains leptons from decay-in-flight of *bottom* or *charm* flavoured hadrons. The quoted cross sections include the filter efficiencies from the generation process.

Process	Cross Section (pb)
$t\bar{t}$	165 ± 10
W + jets	31314
Z + jets	3048
single top, s-Channel	4.6
single top, t-Channel	64.6
single top, tW-Channel	15.7
multijet, $b/c \to e$, $20\,\mathrm{GeV}/c < \hat{p}_{\mathrm{T}} < 30\,\mathrm{GeV}/c$	132160
multijet, $b/c \to e$, $30\,\mathrm{GeV}/c < \hat{p}_{\mathrm{T}} < 80\,\mathrm{GeV}/c$	136804
multijet, $b/c \to e$, $80\,\mathrm{GeV}/c < \hat{p}_{\mathrm{T}} < 170\,\mathrm{GeV}/c$	9360
multijet, e/m enriched, $20\,\mathrm{GeV}/c < \hat{p}_{\mathrm{T}} < 30\,\mathrm{GeV}/c$	2454400
multijet, e/m enriched, $30\,\mathrm{GeV}/c < \hat{p}_{\mathrm{T}} < 80\,\mathrm{GeV}/c$	3866200
multijet, e/m enriched, $80\,\mathrm{GeV}/c < \hat{p}_{\mathrm{T}} < 170\,\mathrm{GeV}/c$	139500

processes from the top quark pair events under study to increase the efficiency and purity of the selection [112].

The theory predictions of the cross section for top quark pair production and the relevant background processes are summarised in Table 4.1.

4.1.1. Selection Requirements for Electron and Jets Events

In Section 1.7.2 it is explained that top quark pairs in the semi-leptonic decay channel have the decay signature

$$t\bar{t} \to W^+W^-b\bar{b} \to \ell\nu\, q\bar{q}\, b\bar{b}.$$

As a consequence, signal events feature two light flavoured jets from the hadronisation of the two light quarks from the hadronic W decay, two b-jets from the hadronisation of the two *bottom* quarks from the top quark pair decay and a charged lepton as well as a neutrino. The latter is detected as missing transverse energy E_T^{miss} in the detector. This analysis focuses on the decay channel where the charged lepton is an electron or positron. The event display for a candidate for such an event is shown in Figure 4.1.

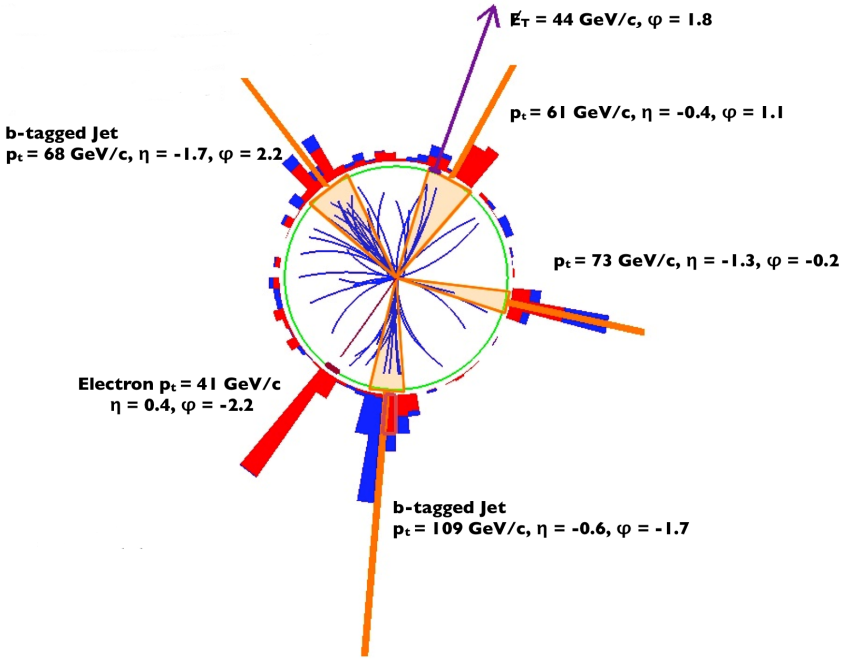


Figure 4.1.: Event display for semi-leptonic electron + jets candidate event [118].

This analysis does not use missing transverse energy as the gain in signal to background separation is not significant and the corresponding systematic uncertainties are large. Furthermore, no *bottom*-jet identification is applied, avoiding the systematic uncertainties introduced by this technique. Hence, the signature for this analysis is represented by at least four well reconstructed jets and exactly one well reconstructed electron.

Online Event Selection

All events that are recorded in CMS must pass a certain trigger path as described in Section 2.2.7. For electron and jets events the path with the loosest selection criteria while still being unprescaled is the electron and three jets trigger path `HLT_Ele25_TriCentralJet30`[1]. This means the trigger algorithm is based on the presence of an online electron in the pseudorapidity region $|\eta| < 2.5$ with a transverse momentum of at least 25 GeV/c and three jets in the pseudorapidity region $|\eta| < 2.6$ with a transverse momentum larger than 30 GeV/c. A detailed study of this trigger algorithm together with its selection efficiency measurement is presented in Section 5.1.

[1] Short-hand for `HLT_Ele25_CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT_TriCentral(PF)Jet30`.

Technical Filters

Two technical filters are applied on collision data. These filters discard events which are possibly affected by technical problems during the data taking process. The filters are not used on the simulated samples because the technical problems they address are not modelled in the simulation.

The so-called *HBHO noise filter* removes events with a large noise level in the hadronic calorimeter. The *beam scraping veto filter* discards events where there was not actually a collision taking place but rather beam remnants caused energy deposits in the detector.

Primary Vertex Selection

Selected events are required to contain at least one good vertex. Only the first vertex in the collection of reconstructed vertices is considered for the selection – this is a common agreement in the CMS Top Physics Analysis Group originating in requirements of certain *bottom*-jet identification algorithms. The primary vertex must not be identified as fake primary vertex, that means the beamspot position is used to approximate the vertex position, the number of degrees of freedom of the fit to the tracks of the primary vertex is required to be larger than four, the absolute value of the z-coordinate of the primary vertex is chosen to be smaller than 24 cm and the value of the spherical coordinate ρ must be smaller than 2 cm. The value of 24 cm is larger than the bunch length of 7.55 cm [26].

Jet Selection

Jets are reconstructed from *Particle Flow* objects as described in Section 2.4 with the anti$-k_T$ jet clustering algorithm [52] with a distance parameter $r = 0.5$.

Several jet quality criteria are applied [119]. For the following analysis steps a jet is considered if the number of constituents is larger than one, the charged electromagnetic fraction is smaller than 0.99, the neutral hadronic fraction is smaller than 0.99, the neutral electromagnetic fraction is smaller than 0.99, the charged hadronic fraction is larger than zero and the number of charged hadrons that it contains is larger than zero. This selection ensures that the jet is reconstructed involving several subdetectors which leads to higher quality and reduced misidentification.

Four jets are expected as a result of the decay of the top quark pair. However, additional jets can be reconstructed due to the underlying event, soft radiation or jet splitting. Therefore, an event is accepted if it contains at least four jets within the pseudorapidity region $|\eta| < 2.4$. The two leading jets in transverse momentum are required to have a transverse momentum of $p_{\mathrm{T}} > 50\,\mathrm{GeV}/c$, the transverse momentum of the third jet has to be larger than $40\,\mathrm{GeV}/c$ and for the fourth jet must hold $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$. This selection is motivated by the transverse momentum threshold of the hadronic part of the trigger that is used in the event selection. A detailed explanation is provided in Section 5.1.6.

The jet requirement leads to a large reduction of multijet production and of the W + jets and Z + jets background processes. The latter are significantly reduced because the cross section for the W/Z boson production with associated jet production decreases rapidly with an increasing number of additional jets.

Electron Selection

An electron is required to pass a list of quality criteria as part of the *Cuts in Categories* (CiC) electron ID [120]. The *Cuts in Categories* electron ID is optimised for simplicity, efficiency and robustness in the 2011 data taking period. It allows to understand better the different influences of the detector systems on the electron identification than a multivariate technique and still shows comparable performance [120]. The selection is based on the ratio of energy in the hadronic calorimeter behind the super cluster to the energy of the supercluster, the distance in pseudorapidity $\Delta\eta$ and polar angle $\Delta\phi$ between the position of the supercluster and the track direction at the vertex extrapolated to the electromagnetic calorimeter – assuming no radiation, and the cluster shape covariance $\sigma_{i\eta i\eta}$ [121]. The selection thresholds vary within different *categories* that are assigned to electrons. The nine categories are high *bremsstrahlung* electrons, low *bremsstrahlung* electrons, electrons with a low quality of the reconstructed track, and electrons that are reconstructed in the gap between barrel and endcap. For each type of electron in the list there are two categories defined, one for electrons in the barrel region and one for electrons in the endcap region which leads to eight categories in total. The ninth category contains pure tracker-driven electrons. A full list of the multi-dimensional selection threshold matrix can be found in [120].

In order to suppress electrons arising from hadronic activity, the electrons are required to be isolated. Relative Particle Flow isolation is defined as the ratio

$$\text{rel. PF Isolation} = (\text{photon iso.} + \text{neutral hadron iso.} + \text{charged hadron iso.})/p_{\mathrm{T}}^{\ell},$$

where p_{T}^{ℓ} is the transverse momentum of the lepton. The isolation value defines the amount of energy carried by Particle Flow objects – photons, neutral hadrons or charged hadrons – in an isolation cone around the lepton axis as illustrated in Figure 4.2.

In the alternative definition of relative isolation, as used in the online selection, the energy of charged tracks and the calorimeter deposits in an isolation cone around the lepton are summed. In this definition, a veto-cone is used to subtract the energy of the particle under study.

In the Particle Flow based isolation, in order to account for additional neutral energy deposits due to pile-up interactions, the isolation quantity is corrected using an effective area approach

$$\mathrm{Iso}_{\,\mathrm{rel.}}(\mathrm{e}) = \frac{1}{p_{\mathrm{T}}^{e}} \cdot \Big(\sum_{\Delta R<0.3} \mathrm{p_T(CH)} + \max\big(0, \sum_{\Delta R<0.3} \mathrm{p_T(NH)} + \sum_{\Delta R<0.3} \mathrm{p_T(Ph)} - \rho \cdot A_{\mathrm{eff}}\big)\Big),$$

where ρ is the average energy density in the event as described in Section 2.4.6, A_{eff} is the approximate effective area of electrons in different regions of the detector [122], and CH, NH and Ph denote the charged hadrons, neutral hadrons and photons inside the isolation cone $\Delta R < 0.3$. The effective area varies between values of 0.2 in the barrel region of the detector up to values of 0.5 in the endcap region.

The effective area is derived in two steps. A straight line is fit to the distribution of the average electron isolation as a function of the number of primary vertices in the event. In addition, a second straight line is fit to the average energy density ρ as a function of the number of primary vertices. The effective area is defined as the ratio of the slopes s of the two straight lines

$$A_{\text{eff}} = s_{\text{iso}}/s_{\rho}.$$

A selection requirement of $\text{Iso}_{\text{rel.}}(e) < 0.1$ is imposed on selected electrons.

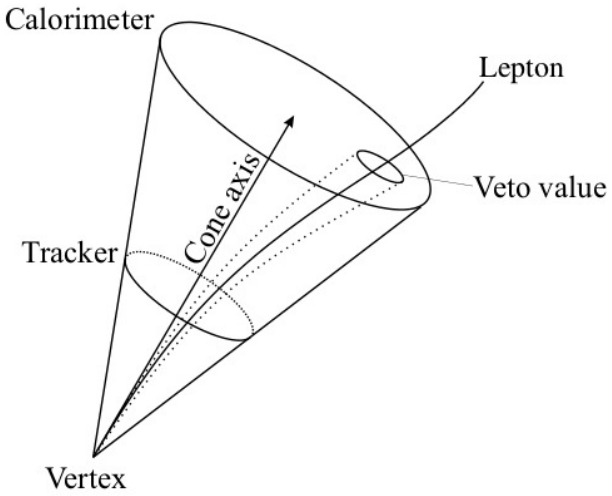


Figure 4.2.: Definition of lepton isolation. The isolation value represents the energy carried by Particle Flow objects in a cone in the η-ϕ-plane around the axis of the lepton. For alternative isolation definitions as they are used in the online event selection, an additional veto-cone allows the subtraction of the energy of the lepton under consideration. Figure based on [18].

In order to reduce the rate of fake electrons and electrons from pile-up interactions, the difference in the z-coordinate of the electron track and the primary vertex is required to be smaller than 1 cm and the transverse impact parameter with respect to the primary vertex of the electron is required to be less than 0.02 cm. The impact parameter is defined as the shortest distance of the linearised electron track to the primary vertex. The distribution is presented in Figure 6.10(b) in Chapter 6.

It is required that exactly one electron with a transverse energy of $E_{\mathrm{T}} > 30\,\mathrm{GeV}$ is found in the event within the pseudorapidity region $|\eta| < 1.5$. The restriction to the barrel region is justified as the reconstruction efficiency for electrons is significantly lower in the endcap region of the detector. In addition, due to the heavy mass of the top quark, top quark pair events tend to be produced more centrally than multijet events and W + jets events. Hence, only electrons from the barrel region of the detector are considered for this analysis. To avoid electrons with an impaired reconstruction quality due to the gap in the transition region between the barrel and endcap part of the electromagnetic calorimeter, electrons are rejected if the assigned supercluster lies in the region of $1.4442 < |\eta_{\mathrm{SC}}| < 1.5660$.

The background contribution from photon conversion is specific to every electron analysis. Section 2.4.7 gives an overview of the methods that are available in CMS to reject electrons from photon conversions. In this analysis it is required that there is a measurement of the electron track in the innermost layer of the pixel tracker. Furthermore, a partner track veto is applied, using the rejection criteria for electrons described in Section 2.4.7.

Veto for Additional Leptons

Events are required to contain neither an additional electron nor a muon. This requirement separates top quark pair events in the semi-leptonic electron decay channel from those in the di-leptonic decay channel where both W bosons decay to a charged lepton and a neutrino. In addition, the rejection of a second electron in the event suppresses Z + jets events where the Z boson decays into two electrons $Z \to ee$.

The definition of the second electron in the event is less strict than that of the primary lepton to remove also di-lepton events with a badly reconstructed second lepton. Events are rejected with a second electron with a transverse energy of $E_{\mathrm{T}} > 20\,\mathrm{GeV}$ within the pseudorapidity region $|\eta| < 2.5$, again excluding the overlap region between barrel and endcap. Veto electrons have to pass at least the CiC electron ID at a loose working point and are required to have a maximum value of relative Particle Flow isolation with $\mathrm{Iso}_{\mathrm{rel.}}(e) < 0.2$.

Events are also rejected if a muon is identified with a transverse momentum of $p_{\mathrm{T}} > 10\,\mathrm{GeV}/c$. Furthermore, the relative Particle Flow isolation may not exceed 0.2. Unlike for electrons, corrections for pile-up effects are based on an approach which includes only nearby charged particles not originating from the primary vertex following the official CMS recommendations. The isolation is defined as

$$\begin{aligned}\mathrm{Iso}_{\mathrm{rel.}}(\mu) \;=\; & \frac{1}{p_{\mathrm{T}}^{\mu}} \cdot \Big(\sum_{\Delta R<0.4} p_{\mathrm{T}}(\mathrm{CH}) + \max\big(0, \sum_{\Delta R<0.4} p_{\mathrm{T}}(\mathrm{NH}) + \sum_{\Delta R<0.4} p_{\mathrm{T}}(\mathrm{Ph}) \\ & - 0.5 \cdot \sum_{\Delta R<0.4} p_{\mathrm{T}}(\mathrm{CP})\big)\Big),\end{aligned}$$

where CH, NH and Ph denote the charged hadrons, neutral hadrons and photons

inside the isolation cone and CP the charged particles in the cone of interest but with particles not originating from the primary vertex. The factor 0.5 corresponds to an average ratio of neutral to charged particles measured in jets [123].

4.1.2. Summary of the Event Selection

The process of the event selection is illustrated in Table 4.2 and Table 4.3. Table 4.2 shows the theory prediction for the number of selected events at each selection step. In the first row, the theory prediction for data corresponding to an integrated luminosity of $4.6\,\mathrm{fb}^{-1}$ at a centre-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$ is shown. Subsequently the number of events is listed after the trigger requirement and requiring exactly one well isolated electron, vetoing muons, vetoing additional electrons with a looser definition and requiring at least one to four jets. The uncertainties are purely statistical and based on the number of simulated events. The raw number of simulated events for each sample and each selection step can be found in Table A.7 in Appendix A.2.3.

Table 4.3 shows the sum of all predicted simulated samples from Table 4.2 and compares it to the number of observed collision events. Furthermore, Table 4.3 quotes for each selection step the signal selection efficiency

$$\epsilon_{\mathrm{sig.}} = \frac{N_{\mathrm{sig.}}^{\mathrm{sel.}}}{N_{\mathrm{sig.}}^{\mathrm{prod.}}}$$

for top quark pair events in the semi-leptonic electron + jets decay channel, where $N_{\mathrm{sig.}}^{\mathrm{sel.}}$ is the number of selected electron + jets signal events and $N_{\mathrm{sig.}}^{\mathrm{prod.}}$ is the number of produced electron + jets signal events. Other top quark pair decay channels are considered as background process. The purity of the selection is defined as

$$\pi = \frac{N_{\mathrm{sig.}}^{\mathrm{sel.}}}{N_{\mathrm{sig.}}^{\mathrm{sel.}} + N_{\mathrm{bkg.}}^{\mathrm{sel.}}}$$

where $N_{\mathrm{sig.}}^{\mathrm{sel.}}$ is the number of selected electron + jets signal events and $N_{\mathrm{bkg.}}^{\mathrm{sel.}}$ is the number of selected background events. In addition, the table lists the product of selection efficiency and purity as a figure of merit for the expected statistical uncertainty of the measurement [112].

The rejection capability for the various background processes becomes clear from Table 4.2 and Table 4.3. The trigger requirement of three trigger jets and a trigger electron performs already a preselection of top-quark-like events. After this step the signal efficiency is 52%. The selection is dominated by W + jets, Z + jets and multijet events.

The number of multijet events is in general difficult to estimate. In the selection table, an inaccuracy appears for multijet events in the early selection steps. The considered multijet events are produced at generation level enriched in decay-in-flight electrons or electromagnetic energy deposits in the calorimeter. Furthermore,

Table 4.2.: Theory prediction of data composition at each selection step. The comparison of the sum of all simulated samples to the observed collision data can be found in Table 4.3. The simulation is tuned to the observed collision data as described in Section 4.2. The uncertainties are purely statistical and based on the raw number of simulated events.

Selection	$t\bar{t}$ (e + jets)	$t\bar{t}$ (other)	W + jets	Z + jets	multijet	single top
Produced	112151	651484	144923916	14106409	31186012515	392924
Trigger	58835 ± 23	24386 ± 15	432306 ± 832	203997 ± 270	860399 ± 14104	10569 ± 21
1 iso. e	40584 ± 19	14040 ± 11	220855 ± 595	92973 ± 182	73513 ± 4174	6752 ± 17
μ veto	40568 ± 19	11013.2 ± 9.8	220840 ± 595	92839 ± 182	73513 ± 4174	6615 ± 17
loose e veto	40489 ± 19	9247.8 ± 9.0	220603 ± 595	44786 ± 126	73513 ± 4174	6505 ± 16
≥ 1 jets	39495 ± 19	8729.6 ± 8.8	120870 ± 440	24820 ± 94	43264 ± 2979	5918 ± 16
≥ 2 jets	32681 ± 17	6418.1 ± 7.5	58303 ± 306	9237 ± 57	13082 ± 1295	3933 ± 13
≥ 3 jets	25309 ± 15	4123.9 ± 6.0	26591 ± 206	3503 ± 35	5935 ± 844	2275 ± 10
≥ 4 jets	16812 ± 12	2343.3 ± 4.5	8746 ± 118	1201 ± 21	1919 ± 409	1011 ± 6

Table 4.3.: Theory prediction of number of events at each selection step compared to collision data. The signal selection efficiency $\epsilon_{\text{sig.}}$, the selection purity π and the product of efficiency and purity is quoted for each step. The numbers in the third column are the sum of the corresponding rows in Table 4.2. The discrepancies between data and simulation in the first four steps of the selection process is attributed to phase-space restrictions in the simulated samples.

Selection	Data	Simulation	$\epsilon_{\text{sig.}}$	π	$\epsilon_{\text{sig.}} \cdot \pi$
Trigger	3953966 ± 1988	1590491 ± 14131	0.525	0.037	0.019
1 iso. e	320755 ± 566	448718 ± 4220	0.362	0.090	0.033
μ veto	317475 ± 563	445388 ± 4220	0.362	0.091	0.033
loose e veto	279720 ± 529	395142 ± 4218	0.361	0.103	0.037
≥ 1 jets	200927 ± 448	243096 ± 3012	0.352	0.163	0.057
≥ 2 jets	115186 ± 339	123654 ± 1331	0.291	0.264	0.077
≥ 3 jets	66489 ± 258	67736 ± 869	0.226	0.374	0.084
≥ 4 jets	30612 ± 175	32032 ± 426	0.150	0.525	0.079

for technical reasons the considered multijet events are enriched by signal-like events by requiring at least one electron with transverse energy $E_{\text{T}} > 15\,\text{GeV}$ without further quality criteria imposed and three jets with transverse momentum exceeding $15\,\text{GeV}/c$ without further quality requirements. In later selection steps the prediction of multijet events from simulation suffers from low statistical precision due to the difficulty of simulating multijet events. This problem will be addressed in more detail in Section 6.2.3. This explains also discrepancies in the early selection steps when comparing the theory prediction to the observed collision data events in Table 4.3.

In Table 4.2, requiring exactly one well isolated electron in the event rejects largely multijet events, but also reduces the fraction of Z + jets and W + jets events by 50% compared to the previous step. Vetoing muons in the event leaves the amount of selected signal events basically unchanged while rejecting more than 15% of the events from other top quark pair decay channels compared to the previous step. This is due to the rejection of di-leptonically decaying top quark pairs. The veto of events with an additional electron with looser selection criteria applied has only very little influence on the number of selected signal events but helps to reduce the number of other top quark decay channels by approximately 5%, mainly rejecting events from the di-electron channel, and reduces the number of Z + jets events by more than 50% relative to the previous selection step. Requiring at least one to at least four jets reduces the background from W + jets and Z + jets events significantly with increasing jet number because the cross section of the associated jet production in these processes decreases rapidly. Overall, due to the small production cross section as listed in Table 4.1, the background contribution

from single top processes is small compared to the other backgrounds.

Table 4.3 shows that the selection efficiency for top quark events decreases most significantly after the trigger requirement, the selection of one isolated electron and the subsequent two, three and four jet requirements. On the other hand, the selection purity increases throughout the selection steps up to the fourth jet selection step due to the larger background rejection. The product of efficiency and purity increases up to the requirement of at least three jets which guarantees a minimal statistical uncertainty when extracting the top quark production cross section with a counting method as described in Section 6.1. However, to reduce the systematic uncertainties on the measurement it is advantageous to require at least four jets in the event which reduces the overall uncertainty on the measurement. This is due to an increased separation in the template fit technique used to measure the cross section and explained in Chapter 6.

The selection efficiency for top quark pair events in the semi-leptonic electron + jets decay channel is $\epsilon_{e+\text{jets}} = 0.15$. The selection efficiency for any type of top quark pair events, either in the semi-leptonic electron + jets channel or from any of the other decay channels including the branching fractions, is

$$\epsilon_{t\bar{t}} = 0.02502 \pm 0.00002.$$

The quoted statistical uncertainty is the binomial uncertainty on the efficiency. The low value is explained by the low selection efficiency for top quark pair decays in other than the electron + jets decay channel and the branching fraction of approximately 15 % of the electron + jets decay channel.

4.2. Application of Event Weights for Tuning Simulation to Collision Data

When simulated data samples are produced there are always inaccuracies observed in comparison with real data. This is caused by imperfect modelling of physics processes, for example in the case of the precision of a given theoretical calculation in terms of order of perturbation theory. On the other hand, the detector simulation reflects the state of the detector at the time the event samples are simulated. It can take more than six months to generate a consistent set of simulated event samples because of the large time requirements of the simulation. In the meantime, the setup of the detector changes to adapt to data taking conditions or due to technical problems that are resolved and does not reflect any longer the original behaviour of the detector at the time of the production of the simulated samples. Furthermore, ageing effects of the detector have to be taken into account, for example the loss of transparency of the crystals of the electromagnetic calorimeter over time. Also changes in the beam conditions and hence pile-up conditions have to be taken into account.

Tuning of the simulation to collision data by applying event weights is performed

to account for these effects and to allow most precise measurements. An event weight based on one or more event quantities is assigned to every simulated event to adjust the simulated physics and detector behaviour to the one actually observed in data.

4.2.1. Trigger and Lepton Identification Event Weights

An event weight is assigned to every simulated event passing the event selection described in Section 4.1 to compensate for different lepton and trigger selection efficiencies in collision data and simulation. The applied weights are the scale factors, that means efficiency ratios between collision data and simulation, listed in Table 5.3 and in Table 5.4. The derivation of the scale factors is discussed in detail in Chapter 5.

4.2.2. Pile-up Interaction Weights

At simulation production stage a certain pile-up distribution is chosen which covers the expected full range of pile-up interactions during the data taking period. In collision data the pile-up distribution tends to shift towards higher values of pile-up interactions throughout the data taking period because the LHC is continuously tuned towards higher values of instantaneous luminosity which increases the probability to observe pile-up interactions. The normalised distribution of pile-up interactions in the 2011 data compared to the distribution used for the simulation is shown in Figure 4.3. The applied correction factors are calculated as a ratio in terms of the number of true pile-up interactions [124].

4.2.3. Correction of $W \to \ell\nu$ to Next-To-Leading-Order

In the simulation of top quark pair events with MadGraph the branching fraction for the different decay channels of the W boson is chosen to be the result of the calculation at leading order of perturbation theory with $\mathrm{BF}(W \to \ell\nu) = 1/9$. All top quark pair events have been reweighted to match the next-to-leading order prediction of the branching fraction of 0.1080 ± 0.0009 [2]. The weights applied for the different decay channels are listed in Table 4.4.

4.2.4. Parton Distribution Function Reweighting

The parton distribution functions that are widely used in the CMS collaboration to produce simulated events in 2011 are based on the Cteq library in version 6.l. There are no uncertainties provided for these PDF libraries, but uncertainties are available for the more recent version 6.6 of the Cteq PDF libraries. Therefore, all simulated events are reweighted from the central value of Cteq in version 6.l to that of version 6.6. The reweighting is based on the ratio of the two parton distribution functions at the value of momentum transfer simulated in the event.

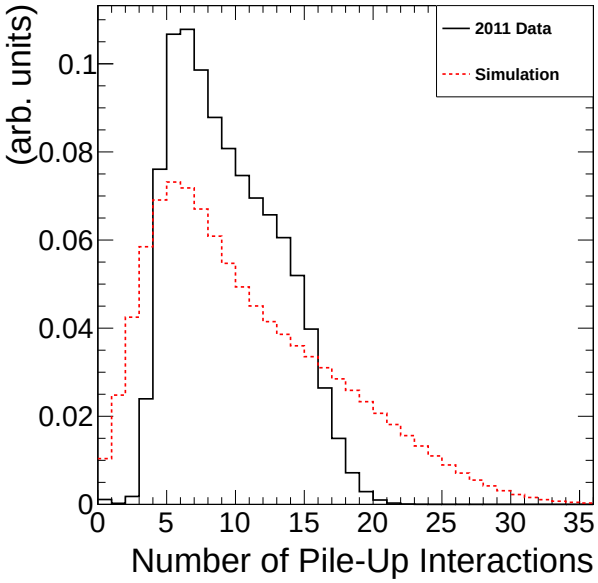


Figure 4.3.: Normalised distribution of number of pile-up interactions in 2011 data and simulation.

Table 4.4.: Event weights applied to top quark pair events to correct from the branching fraction value on leading order of perturbation theory in the simulation to the next-to-leading order value.

Decay Channel	Event Weight
$t\bar{t} \rightarrow \ell\nu b\, \ell\nu b$	0.945
$t\bar{t} \rightarrow \ell\nu b\, qqb$	0.986
$t\bar{t} \rightarrow qqb\, qqb$	1.038

Table 4.5.: Jet energy resolution scale factors and uncertainties in different regions of jet pseudorapidity η.

η Region	Resolution Scale Factor
0.0-0.5	1.052 ± 0.073
0.5-1.1	1.057 ± 0.067
1.1-1.7	1.096 ± 0.079
1.7-2.3	1.134 ± 0.120
2.3-5.0	1.288 ± 0.280

The quoted uncertainties in subsequent studies are based on the uncertainty values provided for CTEQ version 6.6.

4.2.5. Jet Energy Resolution

The jet energy resolution in collision data is in average 10% worse than the one predicted by the simulation. This is corrected for by artificially increasing the jet resolution in the simulation. The transverse momentum of a jet in the simulation is scaled by the product of the transverse momentum difference of the particle level jet and the reconstructed Particle Flow jet and the ratio of the resolution on data and on simulation

$$p_{\mathrm{T}} \rightarrow \max[0, p_{\mathrm{T}}^{\mathrm{gen}} + c \cdot (p_{\mathrm{T}}^{\mathrm{PF}} - p_{\mathrm{T}}^{\mathrm{gen}})].$$

$p_{\mathrm{T}}^{\mathrm{gen}}$ is the transverse momentum of the generator level jet, $p_{\mathrm{T}}^{\mathrm{PF}}$ is the transverse momentum of the Particle Flow jet and c is the core resolution scale factor, defined as the ratio of the average resolution on data and on simulation. The energy resolution is dependent on the pseudorapidity of the reconstructed jet. Table 4.5 lists the core resolution scale factors and their uncertainties [125, 126] in regions of pseudorapitity η.

5. Lepton and Trigger Efficiency Measurement in the CMS Experiment

A precise determination of the signal selection efficiency of the overall selection is crucial as it directly affects the measurement of the production cross section $\sigma \propto \frac{1}{\epsilon}$ where σ is the production cross section and ϵ is the signal selection efficiency. One needs a thorough knowledge of the trigger and lepton selection efficiencies and their uncertainties as these have a strong influence on the overall event selection efficiency.

In Section 5.1 the online event selection performed by the CMS trigger system is described as well as the selection efficiency determination of this system. In Section 5.2 the measurement of lepton identification and isolation efficiencies is described. For both types of efficiencies, scale factors are derived that correct the simulated efficiencies to the efficiencies measured in the collision data.

5.1. Online Event Selection

In the case of semi-leptonic electron + jets events, a cross-object trigger is employed, which requires one isolated lepton in addition to three reconstructed jets. The reconstruction of these objects is described in Section 2.2.7.

In the following section the electron-and-three-jets trigger is described as well as methods that are used to measure the trigger efficiencies for electron + jets events in data and in simulated samples. A measurement of the corresponding efficiencies is performed and a resulting scale factor is derived.

5.1.1. Electron & Jets Triggers in the CMS Experiment

Section 2.2.7 of Chapter 2 explains that in CMS the trigger is divided into different levels. To record electron + jets events a trigger is used that requires at least one trigger electron and three trigger jets in a given event for the event to be written out to storage. Electron and jet reconstruction in the trigger are two independent processes which are described in Section 2.3.2. They are combined in an additional step to result in the described electron-and-three-jets trigger. While the actual trigger menu changes during the data taking in 2011, only one menu is used in the simulation, namely the “3e33” trigger menu.

Table 5.1.: Identification and isolation criteria for trigger electrons as used in electron-and-three-jets triggers as well as updated requirements for the "5e33" trigger menu.

Name	Requirements	Updated Requirements "5e33"
CaloIdVT	$H/E < 0.05$ $\sigma_{i\eta i\eta} < 0.011$	$H/E < 0.05$ $\sigma_{i\eta i\eta} < 0.031$
CaloIsoT	ECALIso$/E_T < 0.125$ HCALIso$/E_T < 0.125$	ECALIso$/E_T < 0.075$ HCALIso$/E_T < 0.075$
TrkIdT	$\Delta\eta < 0.008$ $\Delta\phi < 0.07$	$\Delta\eta < 0.008$ $\Delta\phi < 0.05$
TrkIsoT	TrackIso$/E_T < 0.125$	TrackIso$/E_T < 0.075$

The Electron & Three Jets Trigger in the CMS Experiment

In the CMS Top Physics Analysis Group (PAG) a dedicated trigger has been developed for top quark pair events in the semi-leptonic electron + jets decay channel. This trigger requires that the event contains at least one trigger electron passing specific quality criteria. Variables that are sensitive to the quality of an electron are the ratio of energy in the hadronic calorimeter behind the supercluster H to the energy of the supercluster E, H/E, the pseudorapidity difference $\Delta\eta$ between the position of the supercluster and the track direction at the vertex extrapolated to the electromagnetic calorimeter assuming no radiation, the polar difference $\Delta\phi$ between the position of the supercluster and the track direction at the vertex extrapolated to the electromagnetic calorimeter assuming no radiation as well as the cluster shape variance $\sigma_{i\eta i\eta}$. The electron identification and isolation requirements used in the trigger employed by this analysis are listed in Table 5.1. HCALIso and ECALIso refers to the energy deposited in a cone of $\Delta R < 0.3$ around the electron in the hadronic and electromagnetic calorimeter respectively. TrackIso refers to the sum of the transverse momentum of the tracks in a cone of $\Delta R < 0.3$ around the electron trajectory in the tracker.

In the High-Level Trigger several modules are executed in sequence as can be seen in Table 5.2 for the trigger `HLT_Ele25_TriCentralJet30`. Every trigger begins with the module `HLTBeginSequence` in step 1 which starts the trigger process. The seed *L1SingleEG20* from the Level-1 trigger has at least one electromagnetic object, that means an electron or photon. Electrons and photons are treated in combination in the Level-1 trigger due to the unavailability of tracker information at this step. The seed is examined at the High-Level Trigger in step 2. The transverse momentum of the electromagnetic object is $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$. The High-Level Trigger prescale can be applied in step 3. In step 4 the trigger electron reconstruction is executed and the presence of a trigger electron in the event is required. Trigger jets are reconstructed

Table 5.2.: Trigger sequence for `HLT_Ele25_TriCentralJet30` in the "3e33" trigger menu. The modules relevant for the trigger efficiency determination are highlighted. In step 2 the Level-1 trigger information becomes available, in step 4 trigger electrons are reconstructed and selected and in step 7 trigger jets are selected.

Step	Module Name
1	HLTBeginSequence
2	**hltL1sL1SingleEG20**
3	hltPreEle25CaloIdVTCaloIsoTTrkIdTTrkIsoTTriCentralJet30
4	**HLTEle25CaloIdVTCaloIsoTTrkIdTTrkIsoTSequence**
5	HLTRecoJetSequenceAK5Corrected
6	hltCleanEle25CaloIdVTCaloIsoTTrkIdT-TrkIsoTFromAK5CorrJetsTriCentralJet30
7	**hltEle25CaloIdVTCaloIsoTTrkIdTTrkIsoT-CentralTriJet30EleCleaned**
8	HLTEndSequence

with the anti-k_T jet algorithm in step 5. In step 6 only jets are selected with a transverse momentum of $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$ in the central pseudorapidity region $|\eta| < 2.6$. In the same step an jet–electron cleaning is performed, jets within a distance of $\Delta R < 0.3$ of the first trigger electron in the event are removed from the trigger jet collection. Step 7 guarantees that at least three jets were selected in step 6. The trigger sequence concludes with the `HLTEndSequence` module which marks the end of processing.

If an event does not pass the requirements of any of the modules in a trigger sequence, the trigger sequence is aborted.

5.1.2. Measurement of the Efficiency of Electron and Jets Triggers

Based on the different reconstruction methods of electrons and jets in the CMS trigger system and the quality cuts applied to the trigger objects it is assumed that the triggering of electrons and jets is uncorrelated. The independence of the electron trigger efficiency of the number of trigger jets in the event is verified in Section 5.1.4. A geometric jet-electron cleaning using a requirement of discarding the jet if $\Delta R(\text{jet, electron}) < 0.3$ for the leading trigger electron is used as part of the trigger algorithm. In Section 5.1.5 offline selection requirements are discussed that result in a fully efficient jet part of the trigger. The efficiency of the jet part of the trigger could only increase without the jet-electron cleaning because more jets would be reconstructed. Since the trigger is already 100% efficient with the criteria from Section 5.1.5 this means that the electron part of the trigger does not influence the efficiency of the jet part of the trigger.

The trigger efficiency of the electron-and-jet trigger is assumed to factorise as

$$\epsilon_{e+3\text{jets}} = \epsilon_e \cdot \epsilon_{\text{jets}},$$

where $\epsilon_{e+3\text{jets}}$ is the overall trigger efficiency, ϵ_e is the efficiency of the electron part of the trigger and ϵ_{jets} is the efficiency of the hadronic part of the trigger.

Hence, the efficiency of the leptonic part of the trigger and the efficiency of the hadronic part of the trigger can be measured separately. In Section 5.1.3 the *Tag & Probe* method is introduced which allows the measurement of the leptonic part of the trigger efficiency for electrons that pass the full offline electron selection. The results of this measurement can be found in Section 5.1.4. A *Cut & Count* method which is used to measure the hadronic contribution to the trigger efficiency is described in Section 5.1.5 and the results of the measurement are presented in Section 5.1.6.

5.1.3. The Tag and Probe Method for the Measurement of Lepton Efficiencies

The *Tag & Probe* method provides an unbiased way of measuring lepton efficiencies in high energy physics. Z boson events that decay into two charged leptons $Z \to \ell\ell$ are selected and employed to measure various lepton efficiencies. The decay signature of $Z \to ee$ is very clean and background processes are negligible. One of the two leptons from the Z boson, the *tag* lepton, has to fulfil tight quality criteria. The second lepton, the *probe* lepton, is defined with looser criteria than those for which one intends to measure the efficiency. It is assumed that the two leptons are uncorrelated in terms of the selection efficiency of each one of the leptons. Furthermore, to suppress background events, a Z boson mass constraint is applied, this means the invariant mass of the two leptons must not differ by more than a fixed amount from the nominal Z boson mass. In this analysis the difference between reconstructed invariant di-lepton mass and nominal Z boson mass must be less than $15\,\text{GeV}/c^2$

$$|M_{\ell\ell} - M_Z| < 15\,\text{GeV}/c^2.$$

In events where the probe lepton passes the tight selection criteria, also the probe lepton fulfils the definition of a tag lepton and has to be treated like a tag lepton. That means that there are three type of events. Events where both leptons pass the tag criteria TT, events where one lepton passes the tag criteria and the other lepton does not pass the criteria TP, and events where both leptons do not pass the tag criteria PP. The last type of events is inaccessible for the method and not further considered. The efficiency ϵ is measured as

$$\epsilon = P_{\text{pass}}/P_{\text{all}} = \frac{\text{TT}}{\text{TT} + \text{TP}},$$

where P_{pass} is the number of probe leptons passing the tight selection and P_{all} is the overall number of leptons. If the probe lepton passes the tag criteria in TT

events a double pairing in the same event is performed, that means the same event is considered twice with exchanged roles of tag and probe lepton.

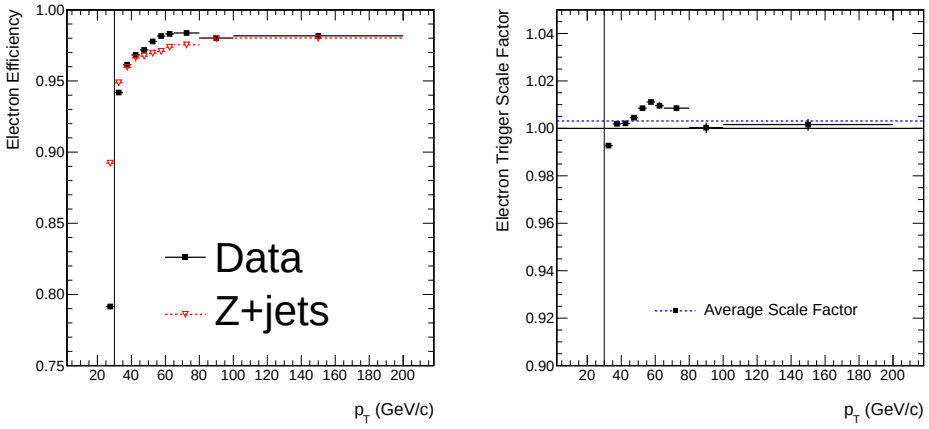
5.1.4. Measurement of the Electron Trigger Efficiency

Events are selected that contain two good offline electrons as defined in Section 4.1. Furthermore, the events must contain a well reconstructed primary vertex and pass the basic event quality filters as explained in Section 4.1. In addition, it is required that there is at least one trigger electron reconstructed according to the full electron selection criteria listed in Section 5.1.1. This trigger electron is geometrically matched within a distance of $\Delta R < 0.3$ to one of the two offline electrons, the tag electron. The trigger efficiency equals the efficiency of offline probe electrons to match a second trigger object

$$\epsilon_e^{\text{trig.}} = N_e^{\text{probe, matched}} / N_e^{\text{probe}},$$

where N_e^{probe} is the overall number of probe electrons and $N_e^{\text{probe, matched}}$ is the number of probe offline electrons that are matched to an online electron.

In this chapter, plots of a given variable are presented as so-called $N-1$ plots, that means a selection based on the given variable is removed, if the figure shows a measurement as a function of the variable under study. All remaining selection requirements are imposed. This means especially that all efficiencies are measured for electrons in the barrel region with a transverse momentum of $p_{\text{T}} > 30\,\text{GeV}/c$. The barrel region requirement is removed only in figures that display the behaviour as a function of the pseudorapidity.



(a) Electron trigger efficiency.

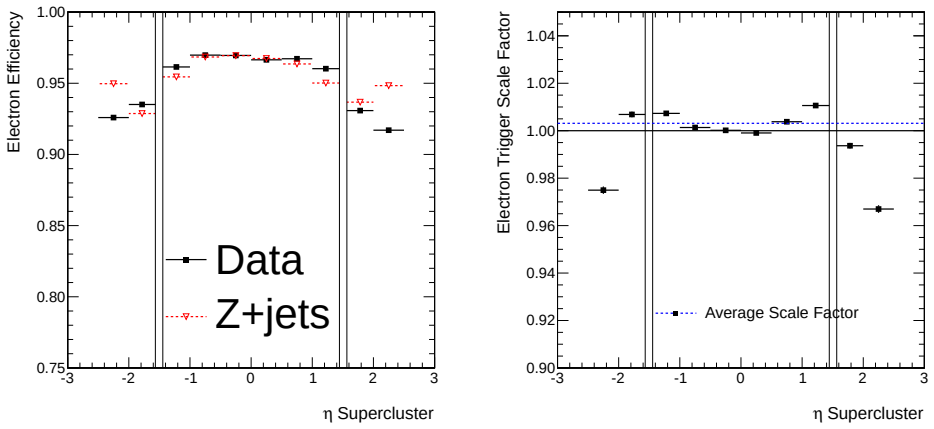
(b) Electron trigger scale factor.

Figure 5.1.: Electron trigger efficiency and scale factor as a function of p_{T} of the offline electron for electrons within $|\eta| < 1.5$. The black vertical line indicates the selection threshold for events.

Figure 5.1(a) shows the luminosity weighted mean of the trigger efficiency in the 2011 data taking period as a function of transverse momentum of the offline reconstructed electron. In addition, the figure illustrates the efficiency behaviour of the trigger in the simulation. One can see that the turn-on behaviour in the data in the transverse momentum range $40\,\mathrm{GeV}/c < p_\mathrm{T} < 80\,\mathrm{GeV}/c$ is steeper than that of the simulation. Nevertheless, the plateau efficiency in the transverse momentum region $p_\mathrm{T} > 80\,\mathrm{GeV}/c$ is identical in data and simulation with a selection efficiency of $\epsilon = 98\%$. Figure 5.1(b) shows the scale factor which corrects the simulation back to the efficiency observed in data. The scale factor is defined as

$$\text{scale factor} = \epsilon_\mathrm{Data}/\epsilon_\mathrm{Simulation},$$

where ϵ_Data is the efficiency measured in collision data, $\epsilon_\mathrm{Simulation}$ is the efficiency measured in the Z + jets simulation, both as a function of a given variable, for example the transverse momentum of the electron.



(a) Electron trigger efficiency.

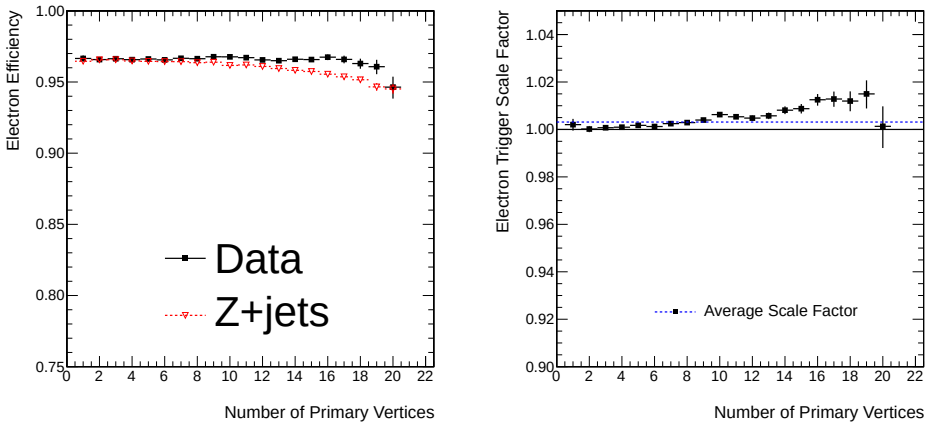
(b) Electron trigger scale factor.

Figure 5.2.: Electron trigger efficiency as a function of electron supercluster pseudorapidity η for electrons with transverse momentum $p_\mathrm{T} > 30\,\mathrm{GeV}/c$. The black vertical lines indicate the excluded transition region between detector barrel and endcap in electron supercluster pseudorapidity.

The behaviour of the electron trigger efficiency as a function of the pseudorapidity of the supercluster associated to the electron is shown in Figure 5.2(a). The trigger efficiency on data and simulation is approximately 97% in the barrel region of the detector. In the endcap regions the efficiency decreases in data down to a value of 92% whereas an efficiency of 95% is predicted by simulation. The scale factor plot in Figure 5.2(b) shows clearly that in the central detector region a scale factor between 1.0 and 1.01 is sufficient whereas in the endcap region a scale factor down to 0.97 would have to be applied. This thesis considers only electrons in the barrel

region of the detector.

Figure 5.3(a) shows the electron trigger efficiency as a function of the number of primary vertices in the event as an indicator for the number of pile-up interactions in the event. On collision data the efficiency is stable within 1%. In simulation, a pile-up dependence of up to 2% can be observed at high values of the number of primary vertices. The scale factor in Figure 5.3(b) shows a maximum deviation of 1% from the average scale factor over the full pile-up range.



(a) Electron trigger efficiency.

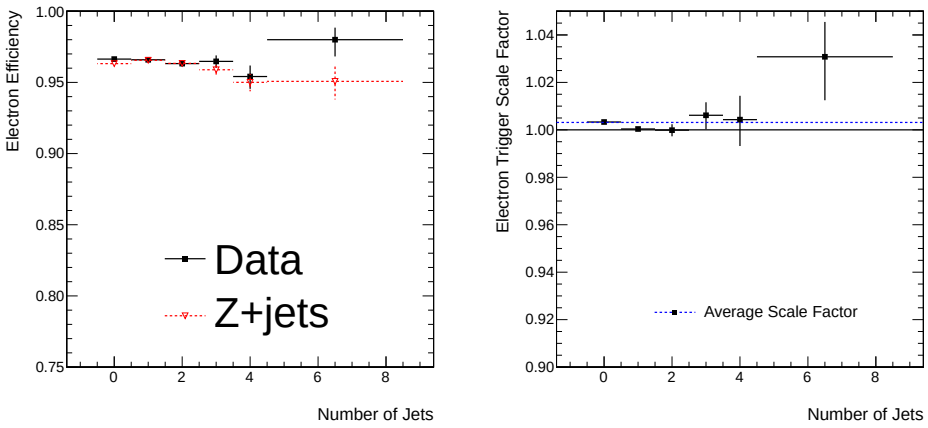
(b) Electron trigger scale factor.

Figure 5.3.: Electron trigger efficiency as a function of the number of primary vertices in the event for electrons within $|\eta| < 1.5$ and with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$.

Figure 5.4 shows the electron trigger efficiency and scale factor as a function of the number of offline jets in the event. The scale factor is flat within the statistical uncertainties.

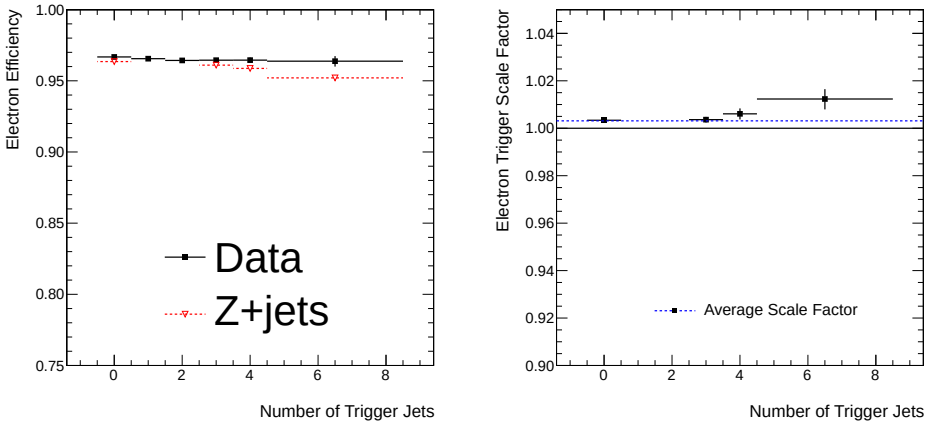
Figure 5.5 shows the electron trigger efficiency and scale factor as a function of the number of trigger jets in the event. The scale factor is flat within the statistical uncertainties. The efficiency on collision data is independent of the number of reconstructed trigger jets. This means that the jet part of the trigger does not influence the measured efficiency of the electron part. In the simulation a dependence on the order of 1% is observed between events without any trigger jets and events with up to eight trigger jets. This dependence is covered by the systematic uncertainty that is assigned to the trigger efficiency measurement.

In Figure 5.6 a two-dimensional scale factor is derived in pseudorapidity η_{SC} of the electron supercluster and transverse momentum p_{T}. This takes correlations between the two variables into account and allows for a pseudorapidity dependent parametrisation of the transverse momentum of the lepton. This two-dimensional scale factor is applied to the simulated samples to correct the simulated efficiencies to match those observed in collision data. A systematic uncertainty of 0.01 is applied in addition to cover the scale factor dependence as a function of the other



(a) Electron trigger efficiency. (b) Electron trigger scale factor.

Figure 5.4.: Electron trigger efficiency as a function of the number of offline jets in the event.



(a) Electron trigger efficiency. (b) Electron trigger scale factor.

Figure 5.5.: Electron trigger efficiency as a function of the number of trigger jets in the event.

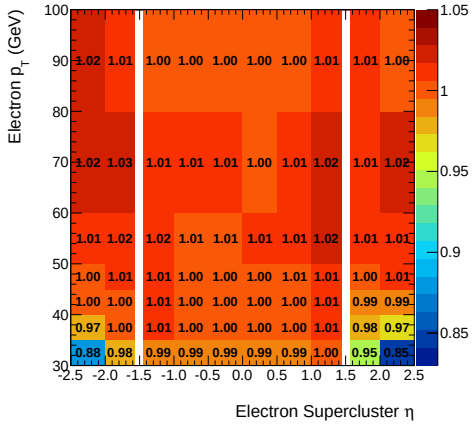


Figure 5.6.: Two dimensional scale factor for the electron leg of the jet trigger as a function of transverse momentum p_{T} and supercluster pseudorapidity η_{sc} of the offline electron in the event.

variables that have been studied.

5.1.5. The Inclusive Cut and Count Method for the Measurement of Jet Trigger Efficiencies

An inclusive *Cut & Count* method is employed to measure the trigger efficiency of the hadronic leg of the electron-and-three-jets trigger. Events that pass the full top quark offline event selection are selected from the single electron dataset. In addition, it is required that the events pass the electron module of the electron-and-three-jets trigger. This requirement guarantees that the leptonic part of the trigger is fully efficient and the efficiency measured with the method described relates only to the hadronic part of the trigger.

For all of the selected events it is checked if the event not only passes the selection mentioned above but in addition also the full electron-and-three-jets trigger. The efficiency is defined as

$$\epsilon_{\mathrm{jet}} = N_{\mathrm{trig}}/N_{\mathrm{tot}},$$

where N_{trig} is the number of events passing the full trigger path and N_{tot} is the number of the previously selected events. This results in a measure for the three jet reconstruction efficiency of the trigger.

It is desirable to keep the transverse momentum threshold for jets as low as possible in the event selection to increase the signal selection efficiency. The effect on the trigger efficiency of requiring at least three jets with a transverse momentum larger than $30\,\mathrm{GeV}/c$ and one additional jet with a transverse momentum of at least $20\,\mathrm{GeV}/c$ in an event is studied in the next section. The threshold of three jets with $30\,\mathrm{GeV}/c$ is the same threshold as in the trigger algorithm. Semi-leptonic

top quark events feature four jets, therefore an additional loose selection cut of $20\,\mathrm{GeV}/c$ for the fourth jet in the event is applied. In a later part of this section it will be demonstrated that this naïve approach yields large systematic uncertainties due to the influence of the jet energy scale uncertainty on the trigger efficiency. Stricter offline jet selection criteria will be adopted to reduce this uncertainty.

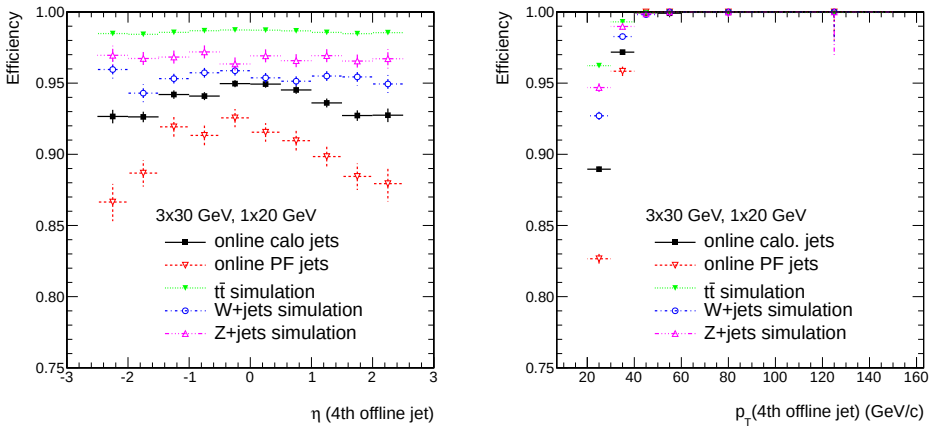
The efficiency is analysed as a function of the fourth offline jet in the event. As there are only three online jets required in the considered events, the fourth-leading offline jet in transverse momentum p_T is used as independent measure of the overall efficiency of the hadronic part of the trigger.

5.1.6. Measurement of the Jet Trigger Efficiency

In Section 2.2.7 it has been explained that during the 2011 data taking period several trigger menus have been deployed. The jet definition in the trigger was stable over most of the 2011 data taking period and based on trigger jets reconstructed from calorimeter towers. In the last weeks of proton-proton collision data taking of the 2011 period, in the run range $178381-180252$ the "5e33" trigger menu has been used as can be seen from Table 2.1. In this trigger menu the definition of jets for the triggers used for top quark analyses in CMS changed from calorimeter-tower based jets to Particle-Flow based jets. Those two types of jets will be studied separately. Jets labelled as *calorimeter jets* are taken from the data taking periods using the "1e33", "2e33" and "3e33" trigger menus. *Particle Flow trigger jets* represent data collected with the "5e33" menu.

In Figure 5.7(a) it becomes apparent that there is a difference in the trigger efficiency of trigger calorimeter jets compared to trigger Particle Flow jets. This is not only an absolute difference but also a difference as a function of jet pseudorapidity is observed. The trigger efficiency of Particle Flow trigger jets is 3% lower in the barrel region and the discrepancy increases up to 6% in the endcap region. In Figure 5.7(b) it becomes clear that the absolute shift can be explained by the turn-on point shifted to higher values of transverse momentum of the Particle Flow jet based trigger. In general, Particle Flow jets tend to be reconstructed with lower transverse momentum than calorimeter based jets. In the trigger no jet energy scale corrections are applied which explains why the turn-on of the Particle Flow trigger jets is shifted towards higher values of offline Particle Flow jet transverse momentum. In Figures 5.7(a) and 5.7(b) also the simulated behaviour of the trigger in three different simulated samples for top quark pair events, W + jets events and Z + jets events is shown. The trigger efficiency differs for all samples and hence also with the sample composition in data. Overall, this makes it difficult to model the jet trigger efficiency precisely in the simulation.

A parametrisation of the trigger efficiency can be done based on several simulated samples and the expected data composition. Nevertheless, the uncertainty on the offline jet energy scale leads to a change of the trigger efficiencies and hence the scale factor which has to be applied to the simulation. This leads to a significant systematic uncertainty that has to be propagated to the final cross section



(a) Efficiency as a function of pseudorapidity η of the fourth leading offline jet in the event.

(b) Efficiency as a function of transverse momentum p_{T} of the fourth leading offline jet in the event.

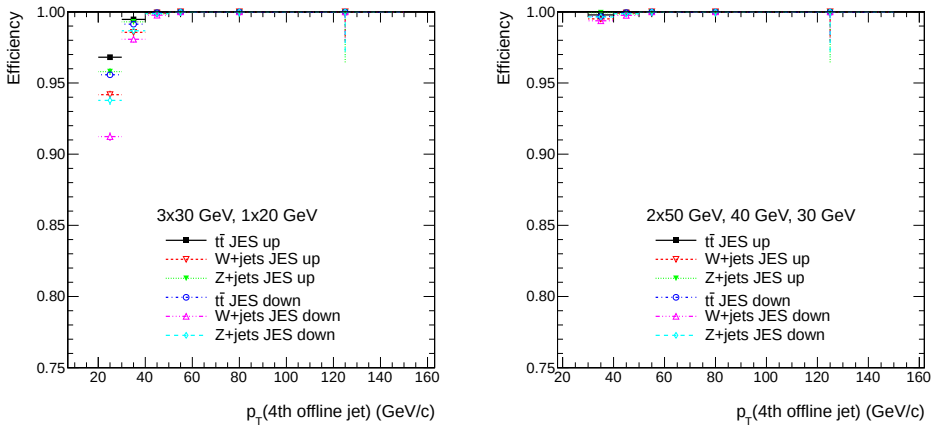
Figure 5.7.: Jet trigger efficiency as a function of pseudorapidity η and transverse momentum p_{T} of the fourth leading offline jet in p_{T} in the event for events with at least three jets above a transverse momentum threshold of $30\,\mathrm{GeV}/c$ and one additional jet above a momentum threshold of $20\,\mathrm{GeV}/c$.

measurement. The effect is illustrated in Figure 5.8(a).

As a remedy for both sample dependence and jet energy scale uncertainties, the offline jet selection requirements are adapted in a way that the trigger efficiency reaches its plateau and at the same time the signal efficiency loss is minimised. The desired behaviour is achieved when requiring as part of the offline selection at least two jets with a transverse momentum $p_{\mathrm{T}} > 50\,\mathrm{GeV}/c$, at least one additional jet with a transverse momentum $p_{\mathrm{T}} > 40\,\mathrm{GeV}/c$ and at least one further jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$. The signal selection efficiency is reduced by a relative decrease of 40% compared to the requirement of three jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$ and one jet with $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$ but only by a relative decrease of 15% compared to a selection requiring four jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$.

By increasing the offline transverse momentum threshold for the jets the large systematic uncertainty on the trigger efficiency measurement due to the jet energy scale vanishes and the measured efficiency becomes independent of the jet energy scale uncertainty within the uncertainties of the measurements as can be seen in Figure 5.8(b).

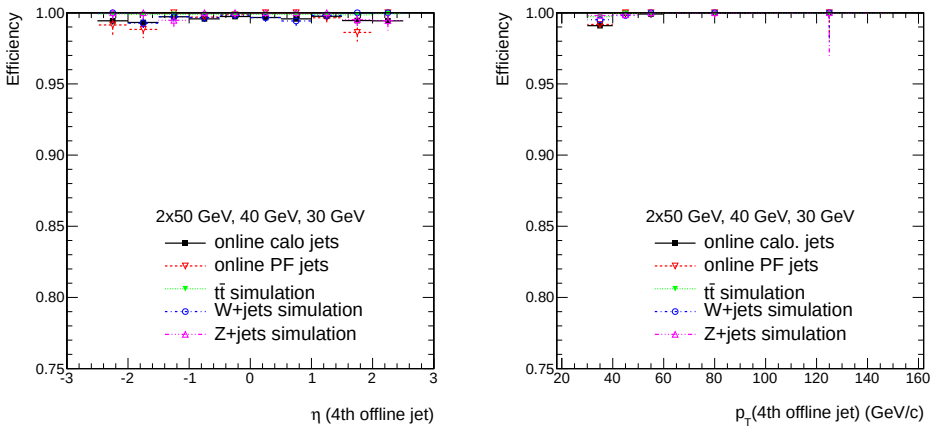
In Figure 5.9 it can be seen that the trigger is fully efficient within an uncertainty of 1% over the full pseudorapidity and transverse momentum range. The uncertainty is propagated to the top quark production cross section measurement and is quoted as systematic uncertainty thereof. Figure 5.10 shows the trigger efficiency



(a) Variation in trigger efficiency measured in simulated samples due to jet energy scale uncertainty for events with the p_{T} threshold of three times $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$ and once $p_{\mathrm{T}} > 20\,\mathrm{GeV}/c$.

(b) Variation in trigger efficiency measured in simulated samples due to jet energy scale uncertainty for events with the p_{T} threshold of twice $p_{\mathrm{T}} > 50\,\mathrm{GeV}/c$, once $p_{\mathrm{T}} > 40\,\mathrm{GeV}/c$ and once $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$.

Figure 5.8.: Variation of measured jet trigger efficiency in simulation due to jet energy scale uncertainty for different offline jet selection thresholds. *JES up* and *JES down* denote an upward or downward variation of the jet energy scale according to one standard deviation.



(a) Efficiency as a function of pseudorapidity η of the fourth leading offline jet in the event.

(b) Efficiency as a function of transverse momentum p_{T} of the fourth leading offline jet in the event.

Figure 5.9.: Jet trigger efficiency as a function of pseudorapidity η (5.9(a)) and transverse momentum p_{T} (5.9(b)) of the fourth leading offline jet in p_{T} in the event for events with at least two jets above a transverse momentum threshold of $p_{\mathrm{T}} > 50\,\mathrm{GeV}/c$, one additional jet above a momentum threshold of $p_{\mathrm{T}} > 40\,\mathrm{GeV}/c$ and one further jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$. The trigger efficiency does not show any dependence on pseudorapidity and transverse momentum and the trigger is fully efficient within an uncertainty of 1%.

as a function of the number of primary vertices in an event. The jet trigger efficiency is independent of the number of pile-up interactions within the uncertainty of 1%. A scale factor of 1.00 ± 0.01 is used for the final analysis.

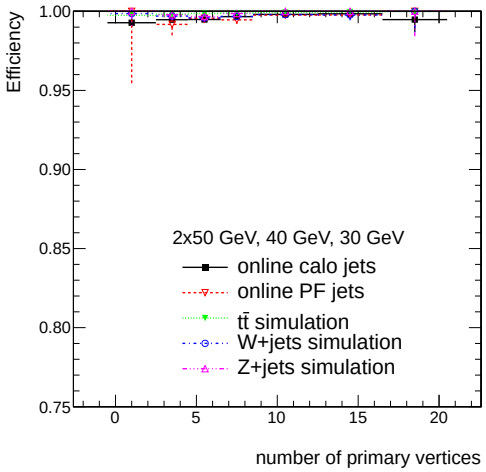


Figure 5.10.: Jet trigger efficiency as a function of number of primary vertices as a measure for the pile-up interactions in the event for events with at least two jets above a transverse momentum threshold of $p_{\mathrm{T}} > 50\,\mathrm{GeV}/c$, one additional jet above a momentum threshold of $p_{\mathrm{T}} > 40\,\mathrm{GeV}/c$ and one further jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$. The trigger efficiency does not show any pile-up dependence and the trigger is fully efficient within an uncertainty of 1%.

5.1.7. Summary on Trigger Efficiencies

The trigger efficiency of the electron and three jets trigger has been measured on simulation and collision data. The resulting scale factors are summarised in Table 5.3. For the electron part of the trigger a two-dimensional scale factor is used while for the jet part of the trigger a flat scale factor of 1.0 has been measured.

5.2. Lepton Isolation and Identification Efficiency Determination

The selection of top quark pair events in the semi-leptonic decay channel depends strongly on the identification of the lepton in the event. As the abundance of jets in the event is not a clear criterion of separation between top quark pair events and multijet events, the determination of the lepton selection efficiency is of crucial importance for the estimation of the number of produced top quark pair events and hence also the top quark pair production cross section.

Table 5.3.: Summary of the measured trigger scale factors. For the electron leg of the cross trigger a two-dimensional scale factor is applied as can be seen in Figure 5.6. The jet part of the trigger is fully efficient on data and simulation within an uncertainty of 1%.

Trigger Part	Scale Factor	Uncertainty
Electron	2-dim.	±0.01
Jet	1.00	±0.01

There are three parts of the lepton identification efficiency that are relevant, namely the efficiency of the geometrical electron identification at a tight working point, the efficiency of prompt electrons to pass the photon conversion rejection and the efficiency for the isolation requirement of $\mathrm{Iso}_{\mathrm{rel.}}(e) < 0.1$ using the effective area corrected relative Particle Flow isolation. The overall electron identification efficiency is the product of those three efficiencies

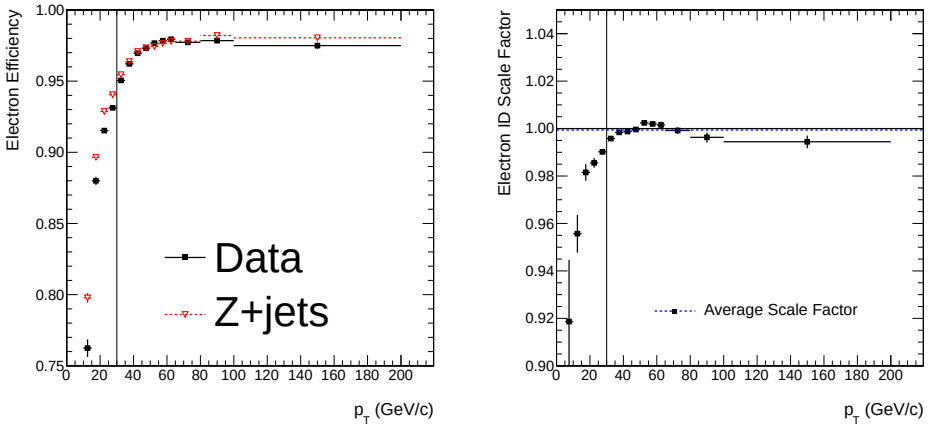
$$\epsilon^{\mathrm{sel.}} = \epsilon^{\mathrm{ident.}} \cdot \epsilon^{\gamma\,\mathrm{conv.\,rej.}} \epsilon^{\mathrm{iso.}},$$

where $\epsilon^{\mathrm{ident.}}$ is the identification efficiency, $\epsilon^{\gamma\,\mathrm{conv.\,rej.}}$ is the efficiency for prompt-electron rejection due to photon conversion rejection methods given the electron passes the identification selection and $\epsilon^{\mathrm{iso.}}$ is the isolation efficiency given the electron passes the identification selection and the photon conversion rejection. The efficiency for the reconstruction of electrons ϵ_{reco} in the final analysis is taken from the simulation.

5.2.1. Measurement of the Geometrical Electron Identification Efficiency

The geometrical electron identification efficiency based on calorimeter-cluster-to-track matching, cluster shape and calorimeter energy distribution as described in Section 4.1.1 is measured at the given working point. Figure 5.11(a) shows the measured selection efficiency on data and simulation as a function of offline reconstructed electron transverse momentum p_{T}. A turn-on is observed which leads to a plateau efficiency of 97% above a transverse momentum threshold of $p_{\mathrm{T}} > 60\,\mathrm{GeV}/c$. The behaviour on collision data is well modelled in the simulation as can be seen from the scale factor plot in Figure 5.11(b).

Figure 5.12(a) shows the electron identification efficiency as a function of the pseudorapidity η of the supercluster associated to the electron. The efficiency in the barrel region of the detector reaches values of between 96% and 97.5% and is very well modelled in the simulation as can be seen from Figure 5.12(b). This leads to a scale factor of approximately 1.0 in the barrel region. In the endcap region the efficiency observed in data is with 95% approximately 1% higher than the one



(a) Electron identification efficiency.

(b) Electron identification scale factor.

Figure 5.11.: Electron identification efficiency and scale factor as a function of p_{T} of the offline electron. The black vertical line indicates the selection threshold for events.

measured in the simulation.

Figure 5.13(a) compares the efficiency behaviour in the simulation to that in the collision data as a function of the number of primary vertices in the event. The efficiency in simulation shows a small downwards trend at the 1% level towards higher numbers of primary vertices while the efficiency on data is flat. Both efficiencies are in average 97%. The scale factor in Figure 5.13(b) is flat within an uncertainty of 1%.

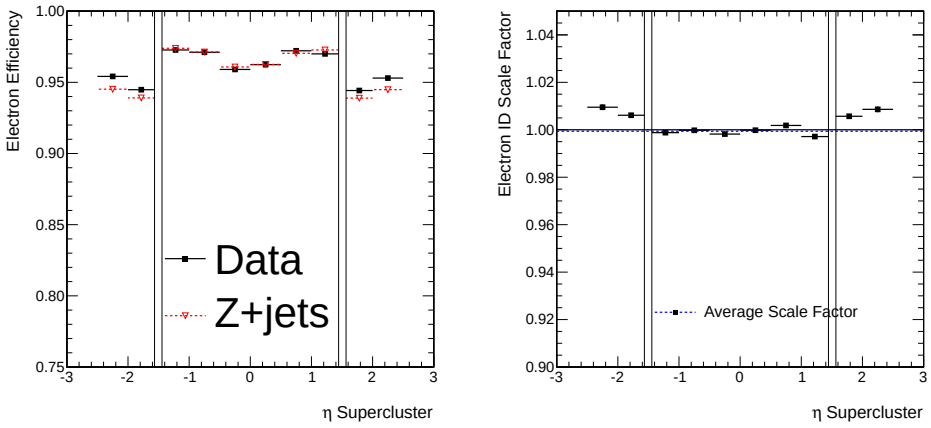
Figure 5.14 shows the electron identification efficiency and scale factor as a function of the number of jets in the event. The efficiency is flat within the statistical uncertainties.

The measured scale factor for the electron identification efficiency is 1.000 ± 0.005.

5.2.2. Measurement of the Prompt Electron Rejection Efficiency due to Photon Conversion Rejection

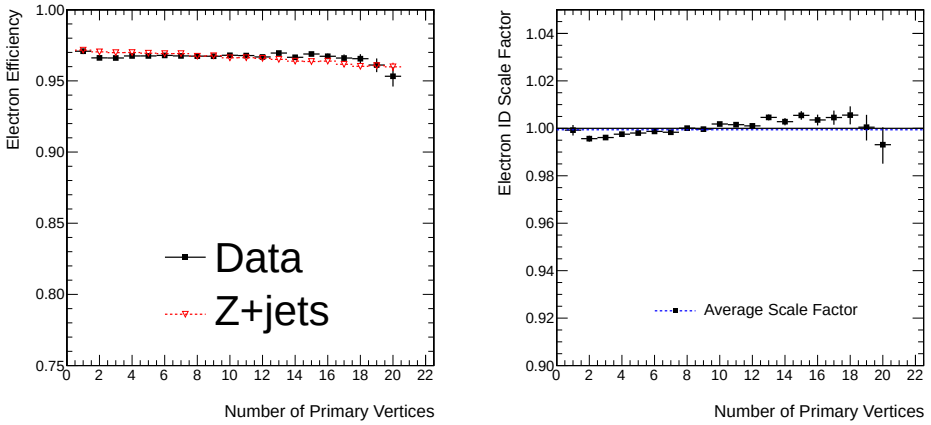
Photon conversion rejection is applied as described in Section 2.4.7. Figure 5.15(a) shows the efficiency after photon conversion rejection for prompt electrons from the $Z \rightarrow ee$ decay. While the efficiencies on simulation and on collision data agree well at values of transverse momentum of up to $70\,\mathrm{GeV}/c$, at higher transverse momentum the simulation and collision data start to diverge by 1%. The overall efficiency decreases as a function of transverse momentum p_{T} of the electron and changes from 99% at lowest transverse momenta to 98% in the simulation and to 97% in the collision data. Figure 5.15(b) shows the corresponding scale factor.

Figure 5.16(a) compares the efficiency observed in collision data and simulation



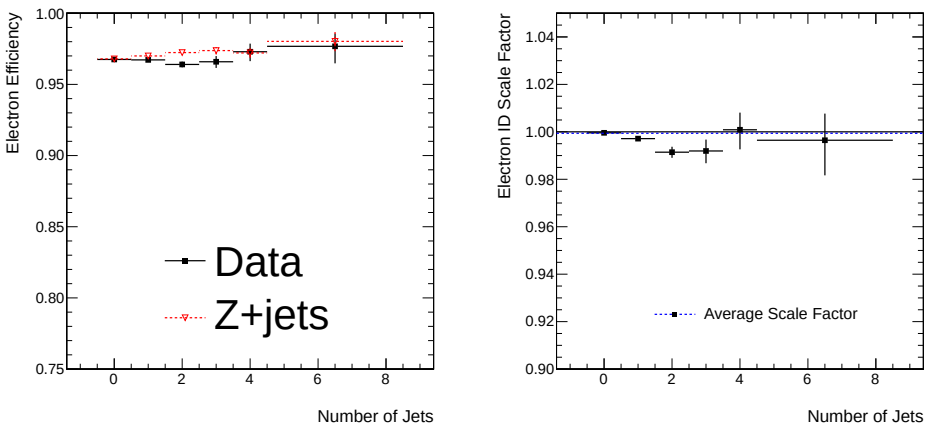
(a) Electron identification efficiency. (b) Electron identification scale factor.

Figure 5.12.: Electron identification efficiency as a function of electron supercluster pseudorapidity η. The black lines indicate the excluded transition region between detector barrel and endcap in electron supercluster pseudorapidity.



(a) Electron identification efficiency. (b) Electron identification scale factor.

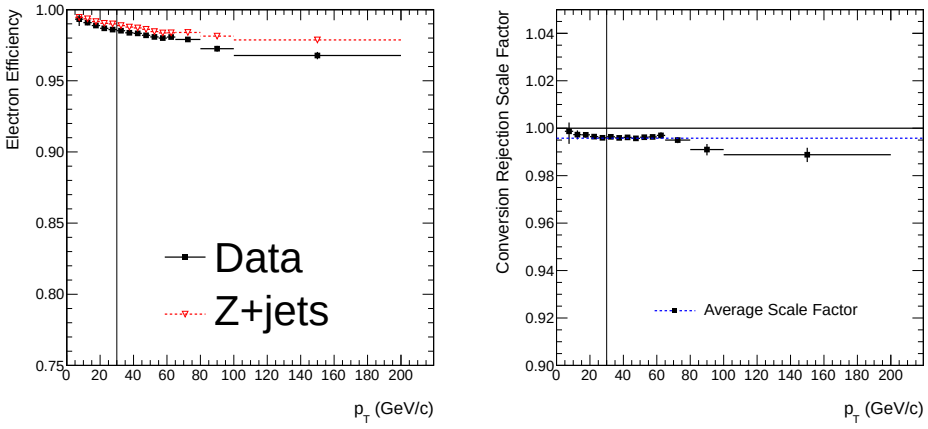
Figure 5.13.: Electron identification efficiency as a function of the number of primary vertices in the event.



(a) Electron identification efficiency.

(b) Electron identification scale factor.

Figure 5.14.: Electron identification efficiency as a function of the number of jets in the event.

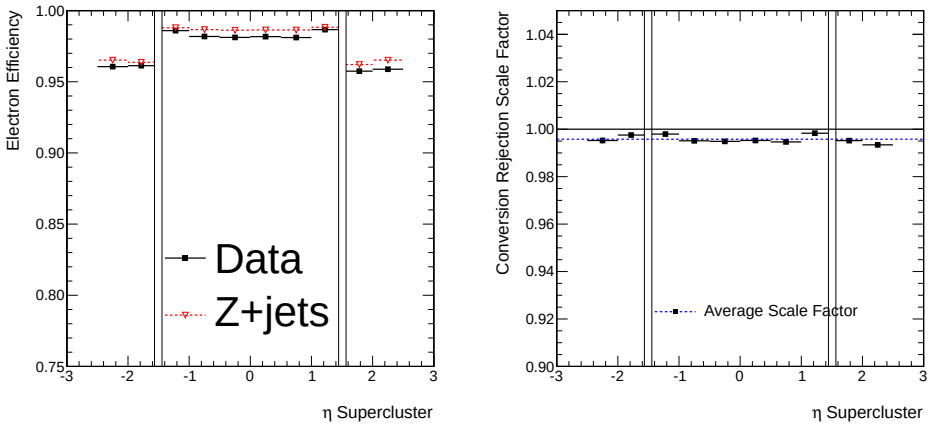


(a) Electron selection efficiency after photon conversion rejection.

(b) Scale factor after photon conversion rejection.

Figure 5.15.: Efficiency for electrons after photon conversion rejection and corresponding scale factor as a function of p_{T} of the offline electron. The black vertical line indicates the selection threshold for events.

as a function of the pseudorapidity η of the supercluster associated to the electron. In the barrel region the efficiency is 99% while it is only approximately 96% in the endcap region. The scale factor 5.16(b) shows a flat behaviour of 0.996 within an uncertainty of 0.005.



(a) Electron selection efficiency after photon conversion rejection.

(b) Scale factor after photon conversion rejection.

Figure 5.16.: Photon conversion rejection efficiency as a function of electron supercluster pseudorapidity η. The black vertical lines indicate the excluded transition region between detector barrel and endcap in electron supercluster pseudorapidity.

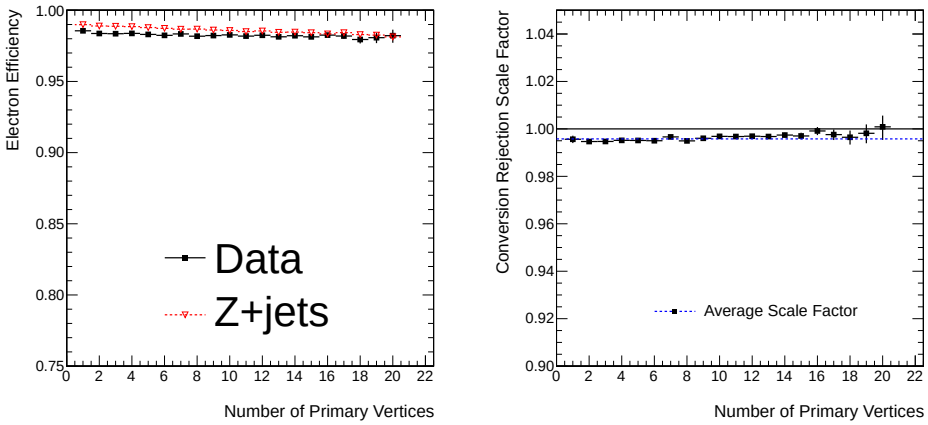
The behaviour of the conversion rejection efficiency as a function of the number of primary vertices in the event is illustrated in Figure 5.17(a). The efficiency in the simulation shows a trend on the sub-0.5% level towards higher values of primary vertices.

Within an uncertainty of 0.005 the scale factor in Figure 5.17(b) is flat at a value of 0.996.

Figure 5.18 shows the electron selection efficiency and scale factor after the application of photon conversion rejection as a function of the number of jets in the event. The efficiency is within the statistical uncertainties flat as a function of the number of jets in the event.

5.2.3. Measurement of the Electron Isolation Efficiency

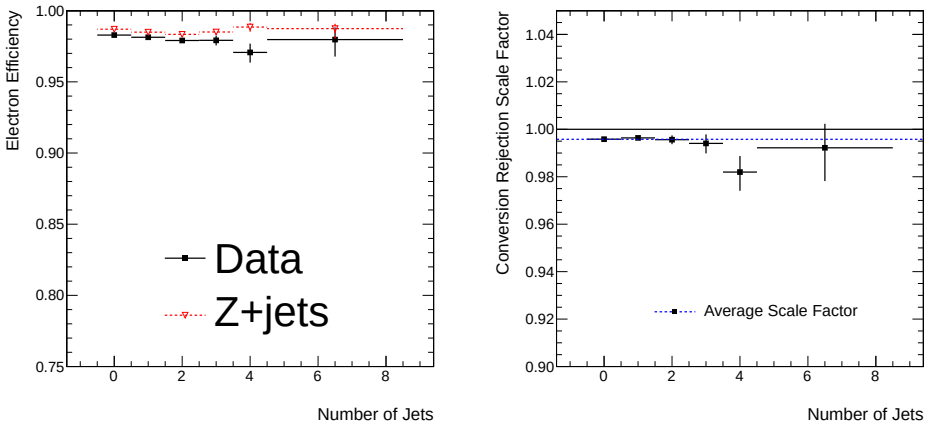
The electron isolation efficiency is measured for the nominal electron isolation used in this analysis. This includes the application of effective area corrections as discussed in Section 4.1.1. The effect of the effective area corrections on the isolation efficiency is discussed by comparing the isolation efficiencies without the correction applied to those that make use of it. The advantage of those corrections by flat-



(a) Electron selection efficiency after conversion rejection.

(b) Scale factor after photon conversion rejection.

Figure 5.17.: Photon conversion rejection efficiency as a function of the number of primary vertices in the event.



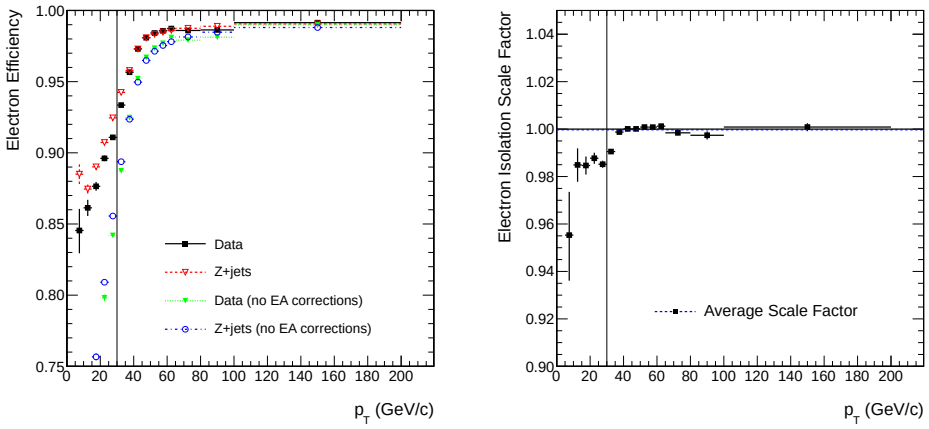
(a) Electron selection efficiency after conversion rejection.

(b) Electron selection scale factor after conversion rejection.

Figure 5.18.: Electron selection efficiency after application of conversion rejection as a function of the number of jets in the event.

tening the pile-up dependence becomes apparent later in this section. The scale factors that are shown are calculated for isolation efficiencies with applied effective area corrections as this is what is used in the remainder of the analysis.

Figure 5.19(a) shows the electron isolation efficiency as a function of electron transverse momentum in collision data and in simulation with and without the application of effective area corrections. For all curves a plateau efficiency of 99% is reached. At electron transverse momentum below $p_{\mathrm{T}} = 100\,\mathrm{GeV}/c$ the turn-on behaviour of the isolation definition without effective area corrections is considerably lower with absolute differences of 1% to more than 10% at lowest transverse momenta. The simulation to collision data agreement in case of applied effective area corrections is very good as can be seen in Figure 5.19(b).



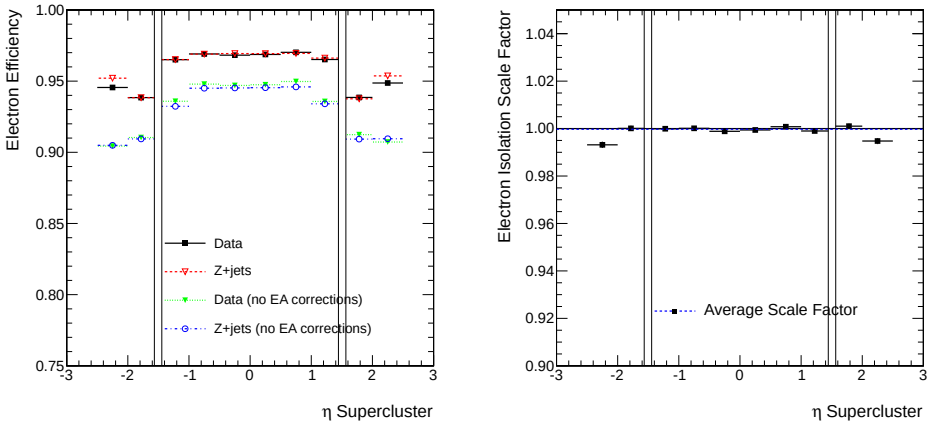
(a) Electron isolation efficiency.

(b) Electron isolation scale factor with effective area corrections.

Figure 5.19.: Electron isolation efficiency and scale factor as a function of p_{T} of the offline electron. The black vertical line indicates the selection threshold for events.

Figure 5.20(a) shows the electron isolation efficiency as a function of the pseudorapidity of the supercluster associated with the electron. One can see that the isolation efficiency is overall higher when effective area corrections are used. In the barrel region the isolation efficiency is 97% with effective area corrections and 95% without effective area corrections. In the endcap region this discrepancy increases to up to an absolute difference of 5%. In the barrel region the scale factor in Figure 5.20(b) is 1.0 and flat over the full barrel pseudorapidity region.

Figure 5.21(a) shows the effect of the effective area corrections is to reduce the pile-up interaction dependence of the isolation efficiency. For a low numbers of primary vertices in the event the two isolation definitions agree. At large numbers of primary vertices in the event and hence in events with many pile-up interactions the efficiency without effective area corrections reduces by 12%. The efficiency



(a) Electron isolation efficiency.

(b) Electron isolation scale factor.

Figure 5.20.: Electron isolation efficiency as a function of electron supercluster pseudorapidity η. The black vertical lines indicate the excluded transition region between detector barrel and endcap in electron supercluster pseudorapidity.

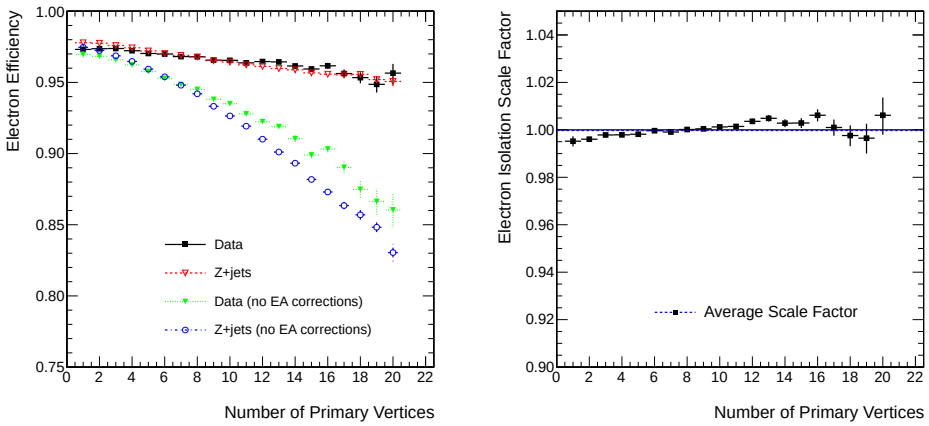
for the isolation definition with effective area corrections shows only a reduction on the order of 2% in efficiency. From this it becomes clear why effective area corrections are applied in this analysis. They make the isolation criterion applied to the electrons nearly independent of the number of pile-up interactions observed in the event.

Figure 5.22 shows the isolation efficiency and scale factor as a function of the number of jets in the event. The efficiency is within the statistical uncertainties flat as a function of the number of jets in the event.

For the isolation efficiency a flat scale factor of 1.000 ± 0.005 is used for the analysis. The uncertainty covers residual discrepancies in the various distributions.

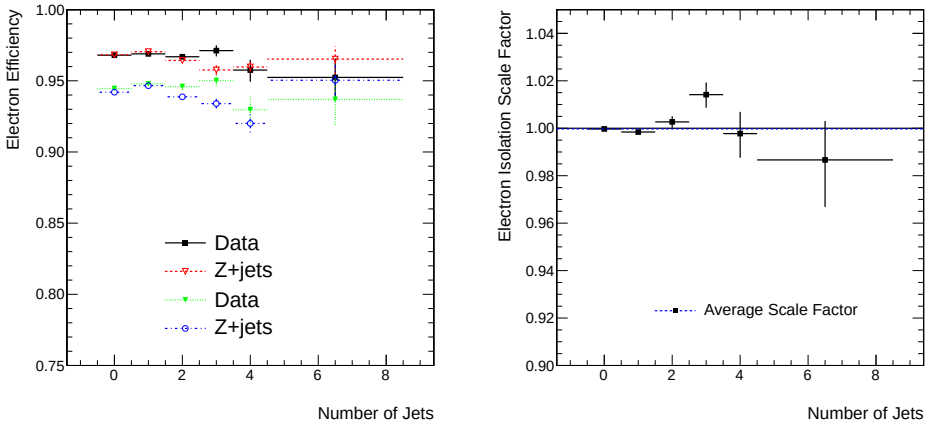
5.2.4. Summary of the Measurement of the Electron Selection Efficiencies

In summary, the electron selection criteria are well modelled in the simulated samples. All scale factors are compatible with 1.0 within the uncertainties as quoted in Table 5.4. As the production of the simulated samples used in this analysis has been started only after a significant amount of the 2011 data has been collected, this good modelling of the efficiencies could be achieved. Larger deviations of more than 5% in the efficiencies have been observed in previous versions of simulated samples. The electron reconstruction efficiency is taken purely from the simulation. A measurement from low level objects, for example tracks, with the tag and probe method is possible but not performed within the scope of this thesis.



(a) Electron isolation efficiency. (b) Electron isolation scale factor.

Figure 5.21.: Electron isolation efficiency as a function of the number of primary vertices in the event.



(a) Electron isolation efficiency. (b) Electron isolation scale factor.

Figure 5.22.: Electron isolation efficiency as a function of the number of jets in the event.

Table 5.4.: Summary of the measured electron selection scale factors given the selection in the preceding row in the table and their uncertainties.

Selection	Scale Factor	Uncertainty
Identification	1.000	±0.005
Conversion Rejection	0.996	±0.010
Isolation	1.000	±0.005

Part III.

Cross Section Measurement

6. Measurement of the Top Quark Pair Production Cross Section in the Electron and Jets Channel

The top quark pair production cross section is measured using a template fit technique, performing a binned maximum log-likelihood fit to a distribution characteristic for the top quark pair decay.

The following chapter introduces the measurement of the top quark production cross section, describes the theory of the techniques used and a method that estimates one of the dominant background contributions, the multijet production, in a data-driven way. Closure tests of the method are discussed and the measurement of the top quark pair production cross section in the electron + jets channel is performed. Control plots of kinematic distributions of event and physics object properties are provided.

6.1. Method

The total top quark pair production cross section is measured based on the equation

$$\sigma_{t\bar{t}}^{\text{total}} = \frac{N_{t\bar{t}}^{\text{sel.}}}{\epsilon_{\text{sel.}} \cdot \int \mathrm{d}tL}, \tag{6.1}$$

where $N_{t\bar{t}}^{\text{sel.}}$ is the number of selected top quark pair events in any decay channel in a given amount of data $\int \mathrm{d}tL$. The overall selection efficiency for top quark events including branching fractions is given as $\epsilon_{\text{sel.}}$.

In the collision data the number of observed events selected after applying the event selection discussed in Chapter 4 yields

$$N^{\text{obs.}} = N_{t\bar{t}}^{\text{sel.}} + N_{\text{background}}^{\text{sel.}},$$

where $N_{\text{background}}^{\text{sel.}}$ is the amount of selected events from other physics processes that pass the top quark pair event selection.

In general, based on simulation one can estimate the number of background events expected after the selection $N_{\text{background}}^{\text{sel.}}$ and calculate $N_{t\bar{t}}^{\text{sel.}}$ from $N^{\text{obs.}}$ in the collision data. This makes the measurement very simple but has several disadvantages. First, the uncertainties of each individual background contribution are propagated to the final top quark pair production cross section measurement. Second, it is necessary to have a quantitative prediction of the number of events

passing the selection from each individual background contribution. This is especially challenging for the multijet background contribution as simulation-based estimates are problematic due to the vast amount of simulated events that become necessary and the difficulty to model the subtle effects of lepton identification in multijet events. Data driven methods for the prediction of the normalisation of the multijet background contribution are available, for example with the *Matrix Method* [127], but suffer from large uncertainties.

6.1.1. Template Fit Technique for the Cross Section Measurement

The *template fit technique* offers the possibility to avoid some of the disadvantages of a counting experiment. A binned maximum likelihood template fit as described in this section is performed on a distribution characteristic for top quark pair events. For each of the relevant background contributions surviving the event selection, the shape for the variable is derived. It is desirable to choose a distribution that is robust against systematic variations. Furthermore, it is important that a separation in terms of a shape difference is given for the individual contributions. One inherent feature of the distributions of top quark pair events and single top quark events is that their shapes are very similar which leads to high correlations of the fit parameters. However, the single top quark cross section has been measured to be in good agreement with the theory prediction [128, 129]. Therefore, it is possible to constrain the amount of single top quark events within a conservative 30% to the theory prediction with a Gaussian constraint of the number of single top quark events in the template fit. A likelihood function is constructed of the form

$$L(N_k) = \prod_{i=1}^{N_{\text{bins}}} \frac{(\sum_k N_{k,i})^{n_i} \cdot e^{-(\sum_k N_{k,i})}}{n_i!} \cdot \frac{1}{\sqrt{2\pi\Delta_{\text{st}}^2}} e^{-\frac{(N_{\text{st}} - N_{\text{Sim. st}})^2}{2\Delta_{\text{st}}^2}},$$

where the fit parameters N_k are the estimated number of events for processes k, n_i is the observed number of data events in bin i of the histogram of the distribution under consideration, $N_{k,i} = N_k \cdot r_{k,i}$ with $r_{k,i}$ the normalised template shape value for process k in bin i, and Δ_{st} denotes the uncertainty of the single top constraint. N_{st} is the number of single top events measured in the template fit and $N_{\text{Sim. st}}$ is the theory prediction on the number of single top events after the event selection. The likelihood describes the Poisson distribution in each bin i of the histogram for each of the contributions k. The factor of the likelihood function is the Gaussian constraint of the single top contribution in the template fit.

The parameters in the likelihood are optimised using the MINUIT [130] package from within ROOFIT to find the optimal value for the individual fit parameters N_k. The uncertainty on the estimate from the fit is derived through the method of MINOS [130].

6.2. Shape Prediction for Signal and Background Distributions and Data Driven Multijet Shape Estimation

6.2.1. Choice of Distribution for Template Fit

Distributions which reflect the high mass of the top quark are particularly different for top quark events compared to Standard Model background processes and provide in general a good separation. One has multiple choices for the reconstruction of the top quark mass in semi-leptonically decaying top quark pair events. Either one reconstructs the mass of the leptonically decaying top quark, the mass of the hadronically decaying top quark or a combination thereof. The problem arises that in the decay of the leptonically decaying top quark a neutrino is emitted which cannot be measured directly with the CMS detector. Only the transverse component of the neutrino four-vector is accessible through the reconstructed missing transverse energy in the event. Missing transverse energy suffers from many sources of uncertainties such as energy loss in the inactive medium of the calorimeter or through gaps for cabling, noise in the calorimeter, and similar effects that make the detector not fully hermetic and active for the energy measurement. All these uncertainties propagate to the final uncertainty on the cross section measurement which is undesirable. Hence, it is advantageous to focus on the reconstruction of the mass of the hadronically decaying top quark. The signature consists of one jet from the hadronisation of the *bottom* quark from the top quark decay and two light flavoured jets from the hadronisation of the decay products of the hadronically decaying W boson.

The invariant mass of the three-jet system out of all jets in the event with the largest transverse momentum p_T,

$$\mathrm{M3} = M_{\mathrm{inv.}}(p_i + p_j + p_k), \text{ maximising } p_{\mathrm{T}}(p_i + p_j + p_k),$$

called *M3* mass, is a robust measure to reconstruct the mass of the hadronically decaying top quark. $p_{i,j,k}$ denotes the four-vector of a given jet in the event. The M3 mass definition is not strongly dependent on any fine tuning and detector understanding like for example a full kinematic fit of the top quark mass that requires the understanding and correct measurement of resolution functions for all objects in the event. Therefore, the M3 mass allows a robust approximation of the top quark mass, providing separation power with respect to the various background distributions.

Figure 6.1 shows the M3 distribution in the 2011 collision data after the electron + jets event selection. The distribution reaches its maximum in the region of the value of the top quark mass. The tail in the distribution is explained by wrong combinations in the jet assignment. The different contributions of physics processes to this distribution are described in the following section.

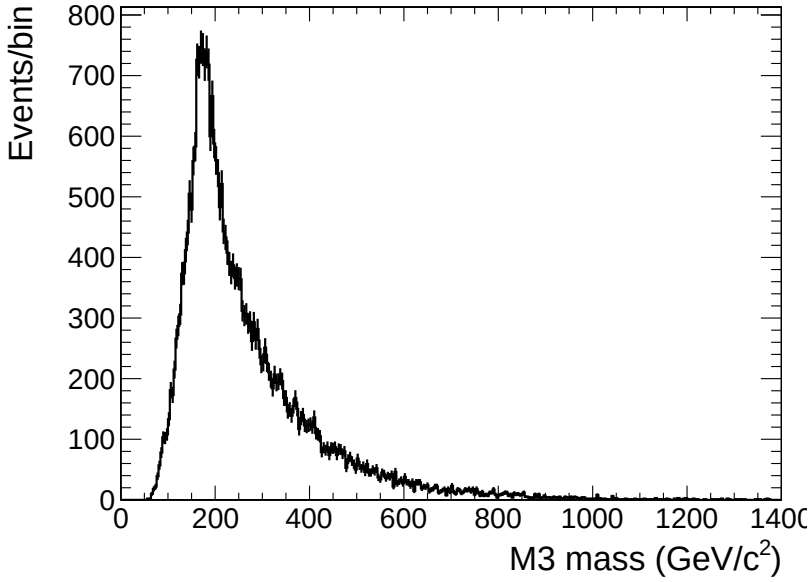


Figure 6.1.: M3 distribution in the 2011 collision data. The distribution reaches its maximum close to the value of the top quark mass.

6.2.2. Shape Estimation from Simulation

Simulated samples are used to predict the shape of the top quark pair signal and dominant background processes. A list of the samples can be found in Appendix A.2.2. The event samples are simulated using MadGraph. In the case of the top quark pair event sample events are generated with up to three additional partons in the final state. The W + jets and Z + jets event samples are generated with up to four additional partons in the final state. PowHeg is used to generate single top quark events. All samples use Pythia for the parton shower generation. Simulated multijet events are available but due to the low selection efficiency the amount of events after the selection is insufficient for an adequate shape estimate. Table A.7 in Appendix A.2.3 shows the raw amount of simulated events after the selection.

6.2.3. Multijet Shape Estimation from Data

The term multijet event refers to events with no prompt electron but rather a $qq/gg \rightarrow qq/gg$ scattering process in which multiple jets are produced. This kind of process has a cross section several orders of magnitude higher than the process of top quark pair production as explained in Section 1.7.1 and shown in Figure 1.4.

The large abundance of multijet events leads to cases where a lepton is identified in the event although there was no prompt lepton produced. Sources for these types of fake leptons are mainly jets with a large electromagnetic fraction which are reconstructed as electrons. Another source of non-prompt leptons are jets with

decay-in-flight electrons that are produced in the decay of heavy-flavoured hadrons within a jet. Furthermore, in events with large track multiplicities also photons may be misidentified as electrons.

In the nominal event selection leptons are required to be well isolated. This removes largely multijet events with fake and decay-in-flight leptons. However, a certain fraction of these events still pass the event selection given their large abundance. This makes it very difficult to simulate this type of event. In the simulated samples on the order of 100 million multijet events are available in the relevant phase space for top quark physics. These events are already electron enriched at Monte Carlo generator level as discussed in Appendix A.2.2. Yet, after the event selection the amount of simulated events is about two orders of magnitude smaller than what is expected in $4.6\,\mathrm{fb}^{-1}$ of collision data. An adequate simulation of this amount of multijet events is not feasible with the current computing resources. This is what makes a data driven measurement of the multijet background contribution desirable, besides the advantage of avoiding systematic uncertainties due to the multijet simluation.

For multijet events out of 109 million simulated events only 25 events pass the event selection.

Measurement of the Multijet Shape using the Anti-Relative-Isolation Method

The assumption is made that the kinematic properties of multijet events in the phase space relevant for top quark physics do not depend on the Particle Flow isolation of the lepton in the event. This hypothesis will be studied in later parts of this section.

Leptons from top quark events and from W + jets and Z + jets events tend to be well isolated. This is due to the fact that the lepton in those events originates from the decay of a W boson or a Z boson and there is no intrinsic reason to expect a large energy deposition in the proximity of the lepton and hence a large relative isolation value for the lepton.

Figure 6.2 illustrates the data composition as a function of the relative isolation of the electron with all other selection cuts applied. The signal range $0 \leq \mathrm{Iso}_{\mathrm{rel.}}(e) < 0.1$ is the nominal region used in the signal selection. The intermediate range $0.1 < \mathrm{Iso}_{\mathrm{rel.}}(e) < 0.25$ is used neither in the signal selection, nor in the extraction of the multijet shape. This is due to a non-negligible prompt-lepton contribution from top quark pair events and W/Z + jets events. The coloured areas depict the contribution of top quark pair events, W/Z + jets and single top quark events based on simulation and Standard Model predictions for the cross sections. The discrepancy between these contributions and the observed number of data events is referred to as multijet events.

Selecting events with $\mathrm{Iso}_{\mathrm{rel.}}(e) > 0.25$, the contamination from top quark signal events or W + jets and Z + jets events and single top events is very low and the selected events are almost exclusively multijet events. Detailed studies on the contamination are presented in a later part of this section. It should be noted that

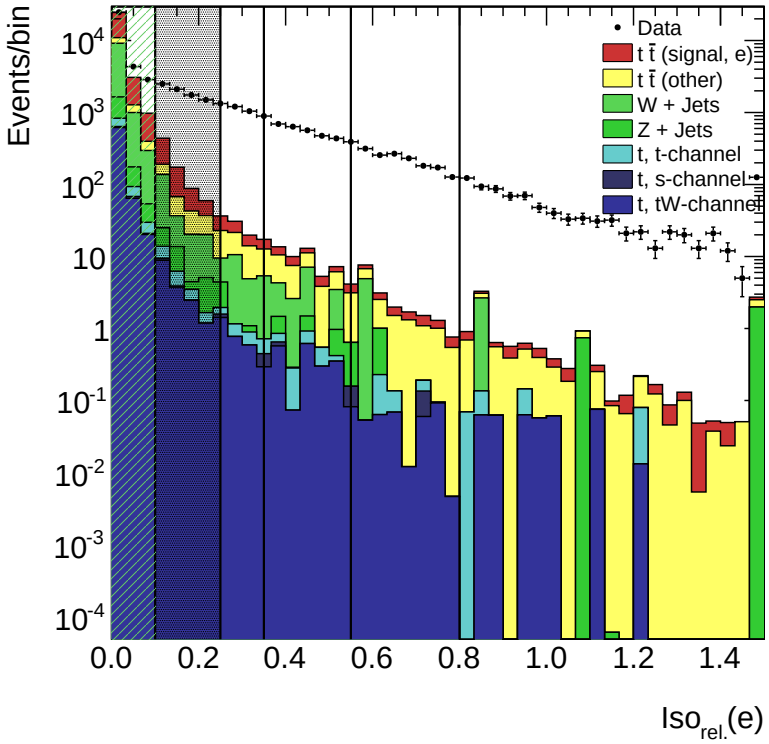


Figure 6.2.: Distribution of the relative Particle Flow isolation for electrons. The region between 0 and 0.1 is the nominal region for the signal selection. The black shaded region between 0.1 and 0.25 is used neither for the signal selection, nor for the determination of the multijet shape. The black vertical lines indicate the four different regions of relative isolation as listed in Table 6.2 and explained in the text.

Table 6.1.: Definition of slices of relative isolation for the extraction of the multijet shape. A simulation based residual signal contamination removal has been performed. Details on the composition of the contamination are presented in Table 6.2.

$\mathrm{Iso}_{\mathrm{rel.}}(e)$	Raw Data Events	After Contamination Removal	Contamination (%)
0.25-0.35	3855	3679	4.7
0.35-0.55	5291	5147	2.8
0.55-0.80	3360	3299	1.9
> 0.80	3227	3146	2.6

an isolation criterion is still implied in terms of the isolation requirement at the High-Level Trigger stage. The isolation definition in the trigger is not identical to the one used for the selection of the offline reconstructed lepton as described in Section 4.1.1 and Section 5.1. Furthermore, the trigger definition of the isolation is looser than that used in the selection of the offline reconstructed electron. Therefore, it is possible to select significant amounts of events with electrons with a large Particle Flow relative isolation value although there is an isolation requirement in the trigger.

One needs to check the dependence of the background shape, as a function of the relative isolation of the lepton. If there is such a dependence, it has to be taken into account by means of a systematic uncertainty. Slices of relative isolation are considered to study the behaviour of the M3 mass shape as a function of the relative isolation of multijet events. If the M3 mass shape is truly independent of the relative isolation for multijet events, the shapes are expected to agree for various slices of relative isolation in the top quark, W + jets and Z + jets depleted region. The choice of the slices in relative isolation is based on approximately equal event numbers per slice.

The definition of the slices in relative isolation together with the number of multijet events per slice can be found in Table 6.1. The slices in relative isolation are separated by black vertical lines in Figure 6.2. Although the overall contamination of events with prompt electrons is low in the anti-relative isolation region, a residual contamination is expected. The signal contamination removal based on simulated events is performed in the anti-relative isolation region. For this, the top quark pair shape, the W + jets shape, the Z + jets shape and the single top quark shape are derived from the simulation in the anti-relative isolation region under study. This shape is then subtracted from the number of observed data events in the region which leads to the pure multijet shape. One has to assume cross section values for the normalisation of the prompt lepton sample shapes that are subtracted. Here, the Standard Model cross section prediction for top quark pair events, W + jets events, Z + jets events and single top quark events is assumed as listed in

Table 6.2.: Prompt lepton contamination in multijet slices for top quark pair events, W + jets events, Z + jets events and single top quark events. The total number of selected events for each slice is listed in Table 6.1.

$\mathrm{Iso}_{\mathrm{rel.}}(e)$	$t\bar{t}$ (%)	W + jets (%)	Z + jets (%)	single top (%)	Overall (%)
0.25-0.35	3.4	0.9	0.2	0.1	4.7
0.35-0.55	1.9	0.7	0.1	0.1	2.8
0.55-0.80	1.4	0.2	0.3	0.1	1.9
> 0.80	1.6	0.5	0.4	0.1	2.6

Table 4.1. An 100% uncertainty on the top quark pair production cross section is assumed and propagated as systematic uncertainty on the final cross section measurement to ensure an unbiased measurement of the top quark cross section independent of the Standard Model prediction. The influence of this uncertainty is suppressed by the small contamination in the selected anti-relative isolation sample. A conservative uncertainty of 50% on the W + jets and Z + jets cross section is assumed and propagated to the final cross section measurement. Residual single top quark contamination is on the 0.1% level in all slices and therefore not explicitly subtracted.

Table 6.2 summarises the percentage of the prompt lepton contamination observed in the four relative isolation regions. The overall prompt lepton contamination is low and varies between 4.6% in the slice closest to the signal region down to values of 1.9% to 2.6% in regions of higher relative isolation.

Figure 6.3 shows the overlay of the normalised M3 mass shape estimate for the multijet background contribution from different slices of relative isolation. All shapes agree within the statistical uncertainties. Residual differences and a possible trend are taken into account as a systematic uncertainty on the final cross section measurement. More details of the effect of the multijet shape on the uncertainty of the final cross section measurement can be found in Section 7.2.11.

6.2.4. Summary of Templates for M3 Distribution

Figure 6.4 shows the normalised template shapes for the M3 distribution that are used for the template fit. Top quark pair events from all top quark pair decay channels are combined into one template due to the similarity of the shape for the M3 mass. The W + jets distribution and the Z + jets distribution have been combined to one single W/Z + jets template shape for the same reason. The ratio for the combination of W + jets and Z + jets events is based on the Standard Model prediction. A conservative uncertainty of 30% on this prediction is propagated as systematic uncertainty on the final measurement. The single top quark shape is based on simulation and combines all three single top production modes, the s-Channel, the t-Channel and the tW-Channel. The single top contribution is

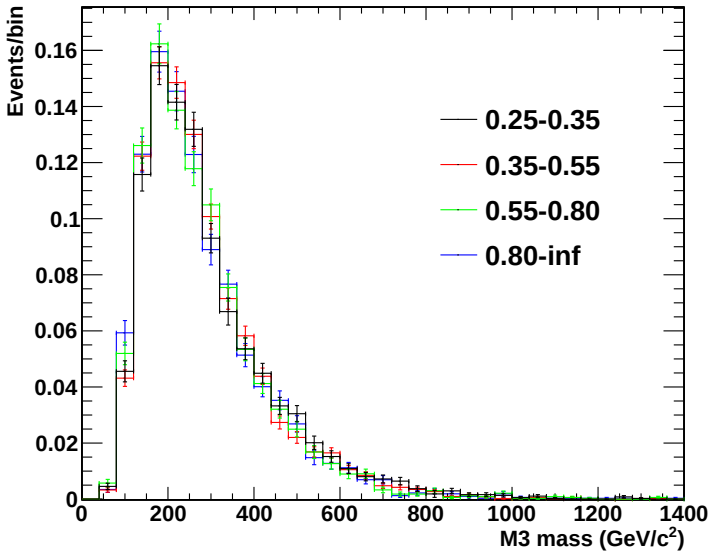


Figure 6.3.: Extracted multijet shape in different regions of Particle Flow relative isolation. The extracted shapes agree within the statistical uncertainties.

constrained within a conservative uncertainty of 30% to the Standard Model prediction as described in Section 6.1.1. Finally, the multijet background contribution is derived as described in Section 6.2.3.

6.3. Validation of the Template Fit Technique in Pseudo-Experiments

An extensive validation of the template fit technique is necessary, given the complexity of the method. In the following section, it will be demonstrated that the employed template fit technique is unbiased and the uncertainties are well estimated. Furthermore, the dependence of the maximum likelihood fit on the binning of the distribution under study is analysed. The studies are performed on ensembles of generated pseudo-experiments based on simulated events.

6.3.1. Generation of Pseudo-Experiments for the Validation of the Template Fit Technique

Pseudo-experiments are generated to test the behaviour of the binned maximum likelihood fit under predefined conditions. In general, there are several ways for the generation of pseudo-experiments. It is possible to produce vast amounts of simulated events and split the dataset into several independent sub-samples of

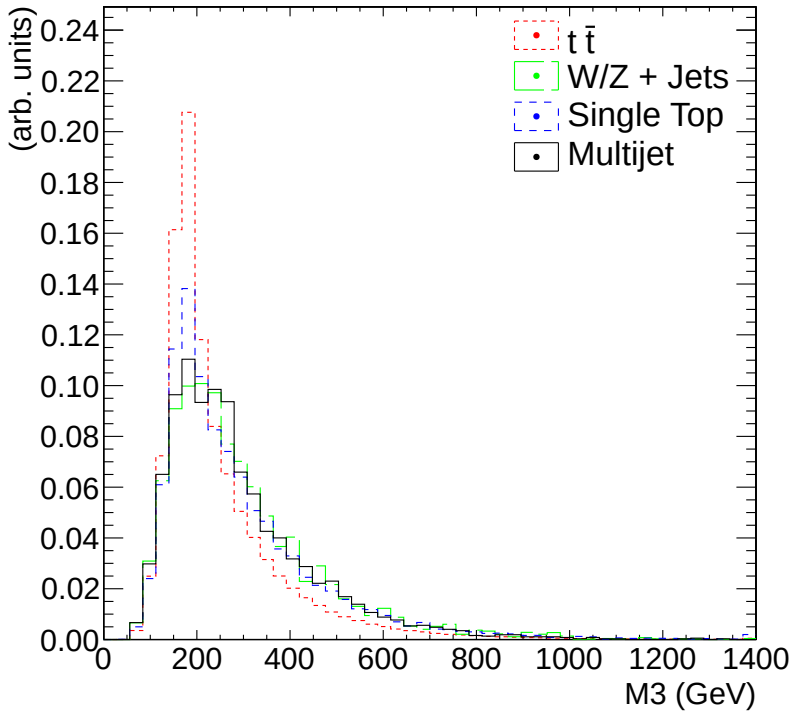


Figure 6.4.: Template shapes for top quark pair events, single top quark events, W/Z + jets events and multijet events.

events. One can derive the template shape from one of the sub-samples, repeat the template fit on the remaining sub-samples and compare the results. This technique is not feasible due to the large amount of time necessary for the simulation of events. Alternatively, events can be randomly drawn from a given simulated event sample to extract the template shapes and perform template fits. Here, the disadvantage lies in the remaining correlation due to the re-appearance of events in no longer independent pseudo-experiments and the still large requirement for numbers of simulated events.

In this thesis, an approach is chosen in which pseudo-data is generated based on the extracted template shapes described in Section 6.2. The template shapes are individually scaled to their theory prediction based on selection efficiencies from simulated samples according to

$$N_i = \int \mathrm{d}tL \cdot \epsilon_i \cdot \sigma_i,$$

where $\int \mathrm{d}tL$ is the integrated luminosity, ϵ_i the selection efficiency from the simulation for the given physics process and σ_i the theory cross section for the given process. The individual templates are summed bin-wise to construct an overall estimation of the shape, normalised to $4.6\,\mathrm{fb}^{-1}$ of collision data.

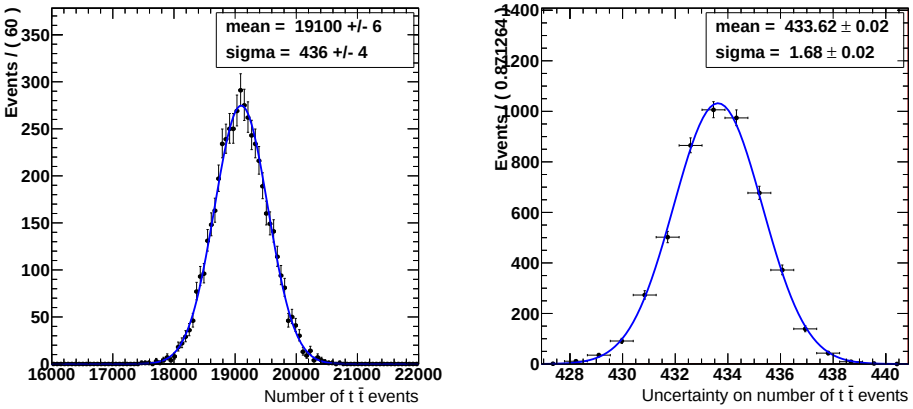
For each bin a random number is generated following a Poisson distribution to generate individual pseudo-experiments. The mean of the Poisson distribution in each bin is the bin content of the original histogram.

6.3.2. Validation of the Template Fit Technique using Pseudo-Experiments

The procedure described in Section 6.3.1 is repeated 5000 times to generate 5000 distributions with Poisson distributed fluctuations. The original templates with high statistical precision from Section 6.2 are then used to perform a template fit to the generated pseudo-data set. After each individual template fit, the top quark production cross section is measured based on the extracted number of top quark pair events, as described in Section 6.1. The estimated number of top quark pair events is shown for 5000 template fits to pseudo-experiments in Figure 6.5(a). The mean value for the measured number of events is 19103 for $4.6\,\mathrm{fb}^{-1}$ of data which corresponds, using Equation 6.1 and an selection efficiency of 0.02502 as described in Section 4.1.2, to a top quark pair production cross section of 165 pb. This demonstrates that the template fit is able to provide an unbiased estimate of the theory prediction of the cross section of 165 pb which is used to generate the pseudo-experiments. Figure 6.5(b) shows the uncertainty on the estimated number of top quark events from the template fit, estimated with the method of MINOS, that means evaluating the negative logarithm of the likelihood function at an increased value of 0.5 with respect to its minimum. 434 events are the mean of the distribution of the fit uncertainty and agree with the width of the distribution of the number of top quark events from the fit with 436 ± 4 events in Figure 6.5(a).

The width of the distribution of the number of events and the mean of the fit uncertainty distribution are not exactly equal because in pseudo-experiments the mean of the single top quark constraint is randomly sampled which leads to a small increase of width of the distribution of the estimated top quark pair events. Both values agree if the random sampling of the number of single top quark events is disabled.

The measured variation corresponds to a statistical uncertainty of 4 pb in 4.6 fb^{-1} of data. A summary of the average fit result for all parameters of the fit in terms of cross sections is given in Table 6.3. The quoted multijet cross section is the cross section for the accepted multijet events only.



(a) Number of top quark pair events as estimated by template fit.

(b) Uncertainty on number of top quark pair events estimated by the template fit.

Figure 6.5.: Number of top quark pair events and its uncertainty as a result of 5000 template fits to pseudo-experiments.

It is useful to define the *pull* [131] of for a contribution k as

$$\text{pull}_k = \frac{N_k - N_{\text{Sim.}\,k}}{\sigma_{N_k}},$$

where N_k is the measured number of events of a contribution k in the template fit, $N_{\text{Sim.}\,k}$ is the theory prediction for the given contribution and σ_{N_k} is the uncertainty on the parameter N_k from the template fit. The pull follows a Gaussian distribution. If the fit is unbiased one expects the pull distribution to be centred around zero. The width σ of the Gaussian fit to the pull distribution is expected to be $\sigma = 1$ if the uncertainties on the extracted parameters from the template fit are well estimated. The pull distribution for the top quark pair fit parameter is shown in Figure 6.6. The width of the Gaussian fit is compatible with $\sigma = 1$ and the pull distribution is centred around zero within the uncertainties which indicates the validity of the results from the template fit technique on pseudo-experiments.

Table 6.3.: Cross sections measured on 5000 pseudo experiments for different sample contributions. The cross section has been calculated from the absolute number of events of each contribution using Equation 6.1. The uncertainty on the measured cross section reflects the width of the distribution of measured cross sections in pseudo-experiments.

Contribution	$\sigma_{\text{Standard Model}}$ (pb)	σ_{Observed} (pb)
$t\bar{t}$	165	165.0 ± 3.8
W/Z + jets	34362	34299 ± 3341
single top	84.9	85.6 ± 24.7
multijet	1886.5	1903.5 ± 1068.4

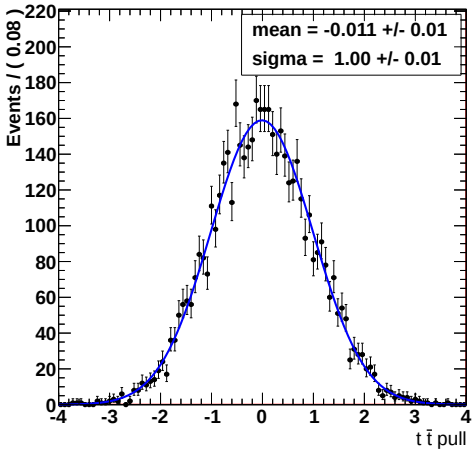
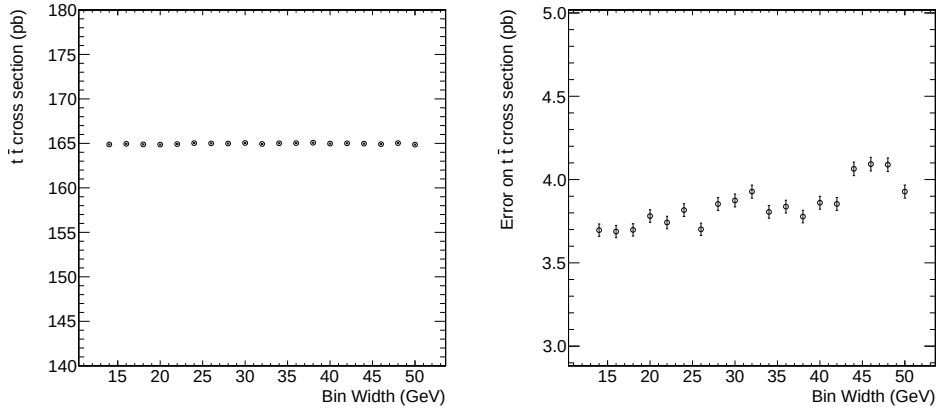


Figure 6.6.: Pull distribution for top quark pair fit parameter. The pull is centred around zero with a width of $\sigma = 1$ within the uncertainties. This means the fit results are unbiased and the uncertainties are well estimated.

6.3.3. Study of the Effect of the Template Bin Width on the Template Fit Estimate

On pseudo-experiments the influence of the template binning on the final fit result is studied by repeating the template fit with different binning. It is desirable that the number of events resulting from the fit is independent of the bin width. An increase of the uncertainty of the fit is expected with increasing bin width due to loss of shape information with fewer bins. A region of bin widths with small variations of the uncertainty is favourable for a robust fit result.



(a) Dependence of measured top quark production cross section on bin width for pseudo-data.

(b) Dependence of uncertainty on measured top quark production cross section on bin width for pseudo-data.

Figure 6.7.: Measured top quark pair production cross section and uncertainty on pseudo-data as a function of the bin width used for the fit. The increase in the fit uncertainty with increasing bin width is due to a loss of shape information with larger bin width. A bin width of $28\,\mathrm{GeV}/c^2$ is used for the final measurement.

Figure 6.7(a) shows the measured cross section as a function of the bin width used for the template shapes and pseudo-data. The cross section measurement is independent of the bin width. One does not expect any variations on the cross section measurement because the pseudo-data is generated based on a distribution with the same binning as the one used in the final template fit. That means that possible fluctuations in the bins appear in the pseudo-data and at the same time in the templates that are used for the fit. Figure 6.7(b) shows the statistical uncertainty of the measured cross section as a function of the bin width used for the template fit. The fit uncertainty is stable over bin width variations. For all further studies a bin width of

$$\mathrm{width(bin)} = 28\,\mathrm{GeV}/c^2$$

is chosen. At this bin width the template fit result is still in the stable plateau of minimal uncertainty starting at finer binnings. Yet, the bin width is wide enough to avoid large statistical fluctuations in the individual bins.

6.4. Measurement of the Top Quark Production Cross Section

The top quark production cross section is measured in $4.6 \pm 0.1\,\mathrm{fb}^{-1}$ of data from the 2011 data taking period. The template fit method as described in Section 6.1 and validated in Section 6.3 is used to extract the number of top quark pair and background events.

Table 6.4 lists the expected and fitted number of events for each contribution after the template fit and the correlation coefficients of the four fit parameters.

One observes that the measured number of W/Z + jets events is 25% lower than expected. The observed number of multijet events after the selection is by a factor two higher than the prediction from the simulation. This behaviour can be understood when taking the correlation between the different contributions into account. The correlation between the parameters for the number of W/Z + jets events and the number of multijet events is negative and large. This is due to the similarity of the template shapes of the W/Z + jets and the multijet processes. The parameter used for the estimation of the number of top quark pair events is reasonably uncorrelated to the remaining parameters of the fit. The largest correlation is the one with the number of multijet events. The sum of the observed number of W/Z + jets, single top quark and multijet events yields 12608 ± 1464. This is in good agreement with the Standard Model expectation of 12884 events.

The estimated number of W/Z + jets and multijet events cannot be considered independently to draw conclusions on either one of those cross sections, but the estimation of the number of top quark pair events is not affected by those correlations. When the fit is repeated fixing the number of W/Z + jets events to the Standard Model expectation a top quark pair production cross section of 156.0 pb is measured. This demonstrates that the measured top quark pair cross section is unaffected by the correlation between the W/Z + jets and multijet template shape. The small change of 0.5 pb in the measured cross section is well below the statistical uncertainty of the template fit.

The results of the template fit to the collision data are shown in Figure 6.8. A good agreement between the result of the template fit and the collision data can be observed. The numerical values of the cross section measurement derived from the event numbers from the fit to data are given in Table 6.5. The cross sections have been derived by using the simulation based selection efficiencies listed in Table 6.6. For multijet events no selection efficiency is estimated and the quoted cross section corresponds to the cross section for accepted events.

The measurement of the top quark pair production cross section in the semi-

Table 6.4.: Expected and observed number of events for different contributions and correlations between individual fit parameters for the fit to 2011 collision data. Only one half of the table is filled due to the symmetry of the correlation factors.

Contribution	$N_{\text{exp.}}$	$N_{\text{obs.}}$	$t\bar{t}$	W/Z + jets	single top	multijet
$t\bar{t}$	19104	18002 ± 431	1	0.07	−0.21	−0.37
W/Z + jets	9952	7497 ± 970		1	−0.11	−0.90
single top	1011	1191 ± 290			1	−0.09
multijet	1887	3920 ± 1060				1

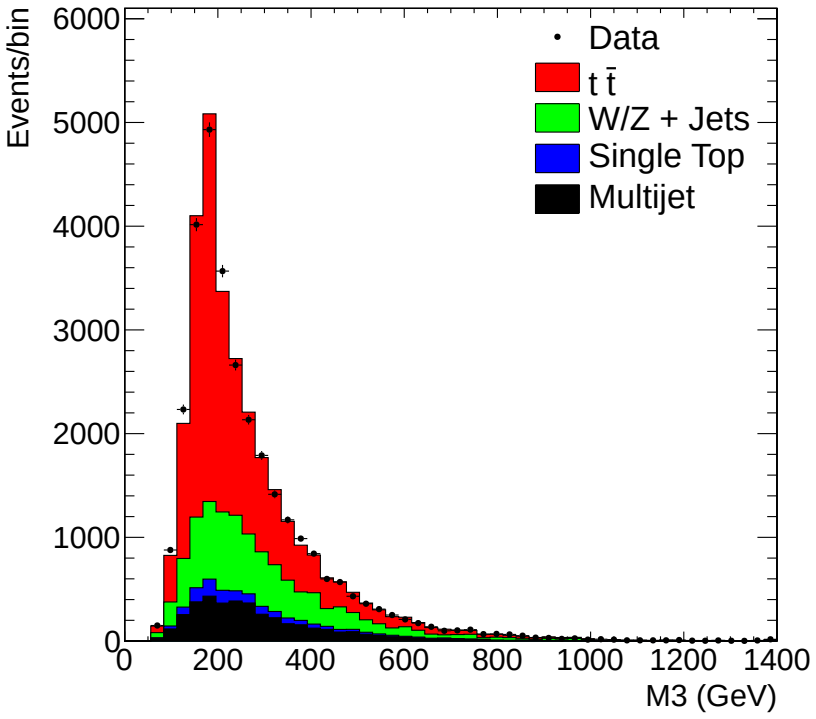


Figure 6.8.: Result of the template fit to the M3 distribution in the 2011 collision data. A good agreement between the result of the template fit and the collision data shape is reached.

Table 6.5.: Cross sections measured on 2011 collision data. The cross section has been calculated from the absolute number of events of each contribution using Equation 6.1.

Contribution	$\sigma_{\text{Standard Model}}$ (pb)	σ_{Observed} (pb)
$t\bar{t}$	165	155.5 ± 3.7
W/Z + jets	34362	23975 ± 3400
single top	84.9	98.8 ± 24
multijet	1886.5	3920 ± 1060

Table 6.6.: Selection efficiencies based on simulation for individual fit contributions.

Contribution	Selection Efficiency
$t\bar{t}$	$(2.502 \pm 0.002) \cdot 10^{-2}$
W/Z + jets	$(6.04 \pm 0.08) \cdot 10^{-5}$
single top	$(2.89 \pm 0.02) \cdot 10^{-3}$

leptonic electron + jets decay channel yields

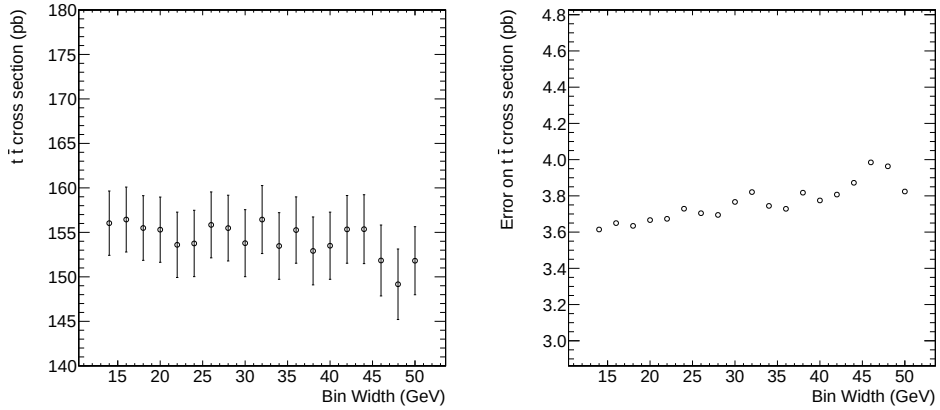
$$\sigma_{t\bar{t}} = 155.5 \pm 3.7\,\text{pb}.$$

The quoted uncertainty reflects only the statistical uncertainty from the template fit. A detailed study of the systematic uncertainties on the measurement is presented in Chapter 7.

6.5. Confirmation of Fit Stability on Collision Data

As an additional verification of the stability of the fit as a function of the bin width, the cross section measurement is repeated on data with varying bin width.

Figure 6.9(a) shows the dependence of the measured top quark pair production cross section as a function of the bin width used in the template fit. The measured cross section is very stable as a function of the bin width. Figure 6.9(b) shows the variation of the uncertainty on the estimated parameter from the template fit. One can see that the variation of the uncertainty is merely 0.2 pb for bin width between $15\,\text{GeV}/c^2$ and $35\,\text{GeV}/c^2$. Towards higher bin widths the uncertainty increases up to values of 4.0 pb due to the loss of shape information of the distribution. Overall, the behaviour expected from simulation is reproduced on collision data and the parameter estimation is stable.



(a) Dependence of measured top quark production cross section on bin width for collision data.

(b) Dependence of uncertainty on measured top quark production cross section on bin width for collision data.

Figure 6.9.: Measured top quark pair production cross section and uncertainty on collision data as a function of the bin width used for the fit. A bin width of $28\,\mathrm{GeV}/c^2$ is used for the nominal measurement.

6.6. Kinematic Distributions Scaled to Cross Sections from Template Fit

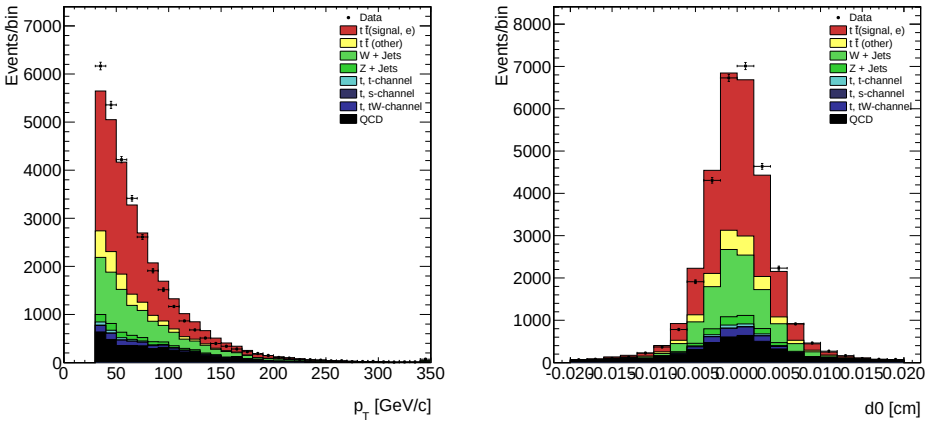
This section gives a simulation-to-collision-data comparison of kinematic distributions in top quark pair events. The individual contributions are based on the template shapes explained in Section 6.2 and scaled to an integrated luminosity of $4.6\,\mathrm{fb}^{-1}$ based on the measured cross sections listed in Table 6.5.

Figure 6.10(a) shows the transverse momentum spectrum of selected electrons. The slope of the distribution is described reasonably well. In the transverse momentum range $30\,\mathrm{GeV}/c < p_\mathrm{T} < 50\,\mathrm{GeV}/c$ a disagreement on the order of 5% is observed. This is most likely an artefact of the multijet model. In the derivation of the multijet shape the relative isolation criterion is inverted. The relative isolation is a function of the transverse momentum of the lepton $\mathrm{Iso}_\mathrm{rel.}(e) \propto 1/p_\mathrm{T}$. Therefore, it is expected that the transverse momentum in the anti-isolation region is not necessarily identical to that in the nominal selection region. Figure 6.10(b) shows the impact parameter distribution for selected electrons. A good agreement is observed between prediction and collision data.

The pseudorapidity distribution of selected electrons is shown in Figure 6.11(a). Disagreements are observed in the pseudorapidity region $0.5 < |\eta| < 1.5$. An explanation to this is the large correlation between the W/Z + jets and multijet shape in the template fit discussed in Section 6.4. In the M3 mass distribution used for the template fit the multijet and the W/Z + jets shape are very similar

which results in a correlation coefficient of the fit parameters of 0.9. This leads to an over-prediction for multijet events and an under-prediction of W/Z + jets events. In the distribution of the pseudorapidity of the electron the shapes of the two contributions are very different. Therefore, the shift in normalisation due to the template fit to the M3 mass leads here to a noticeable disagreement. If the top quark cross section extracted from the template fit is used but the Standard Model predictions for all background contributions are assumed, a very good agreement is achieved as shown in Figure 6.12.

Figure 6.11(b) shows the distribution of the azimuthal angle ϕ of the electron. The distribution is flat with a good prediction-to-data agreement which is expected due to the cylindrical symmetry of the CMS detector.



(a) Electron transverse momentum.

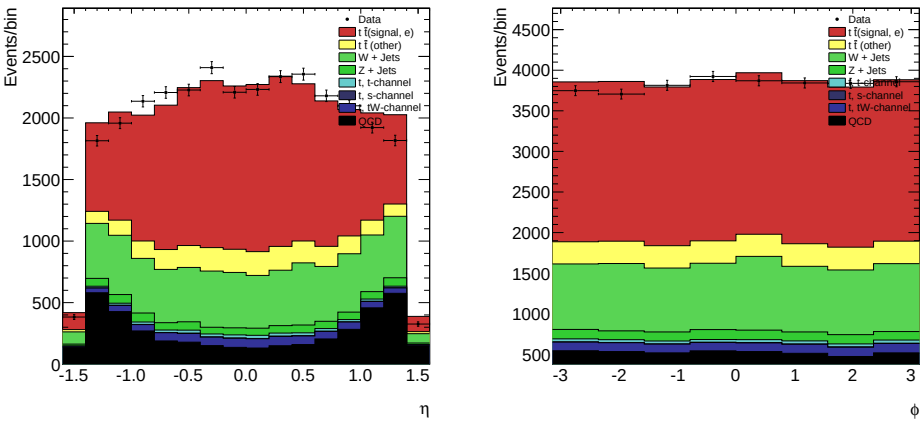
(b) Electron impact parameter with respect to the primary vertex.

Figure 6.10.: Distribution of transverse momentum p_{T} and impact parameter $d0$ of the selected electron.

Figure 6.13(a) shows the distribution of the missing transverse energy in an event. A good agreement between prediction and data is observed. Multijet events which have intrinsically no source of missing transverse energy are observed mainly in the region $E_T^{\mathrm{miss}} < 60\,\mathrm{GeV}$. Top quark pair events and W + jets events contain larger values of missing transverse energy due to the neutrino which is emitted in the leptonic decay of the W boson. Figure 6.13(b) shows the distribution of the total transverse energy H_{T} in the event. The total transverse energy is defined as

$$H_{\mathrm{T}} = \sum_{\mathrm{jets,\,electron}} p_{\mathrm{T}},$$

this means the sum of the transverse momenta of all jets in the event and the transverse momentum of the selected electron. The agreement above $350\,\mathrm{GeV}/c$



(a) Electron pseudorapidity.

(b) Electron azimuthal angle.

Figure 6.11.: Distribution of pseudorapidity η and azimuthal angle ϕ of the selected electron.

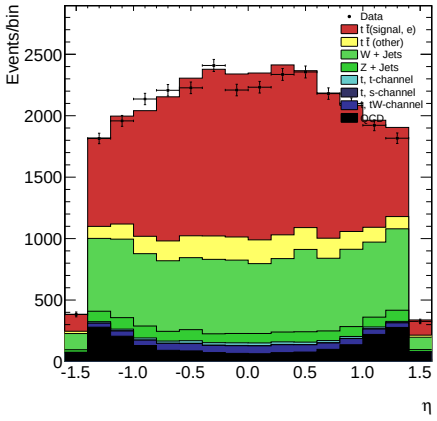
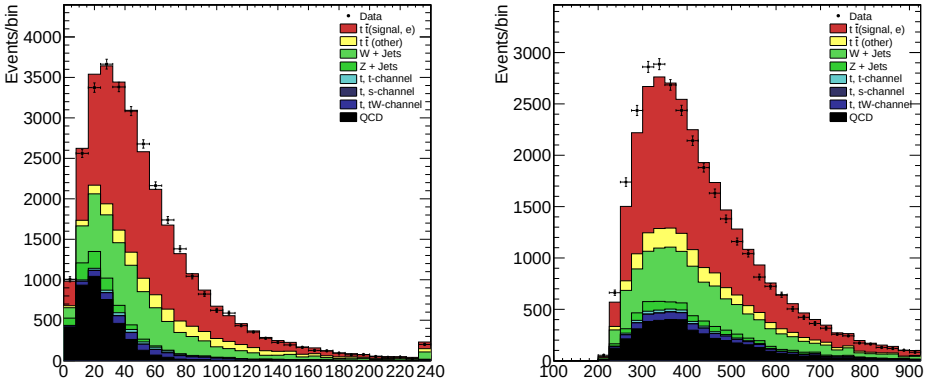


Figure 6.12.: Distribution of pseudorapidity η of the selected electron for measured top quark pair cross section from the template fit and Standard Model expectations for background estimate.

is good while up to 10% more events per bin are observed than expected in the region $H_{\mathrm{T}} < 350\,\mathrm{GeV}/c$. This discrepancy is most likely due to the fact that the electron transverse momentum is used in the calculation of the total transverse energy. Hence, the discrepancy in the electron p_{T} distribution is propagated to the H_{T} distribution.



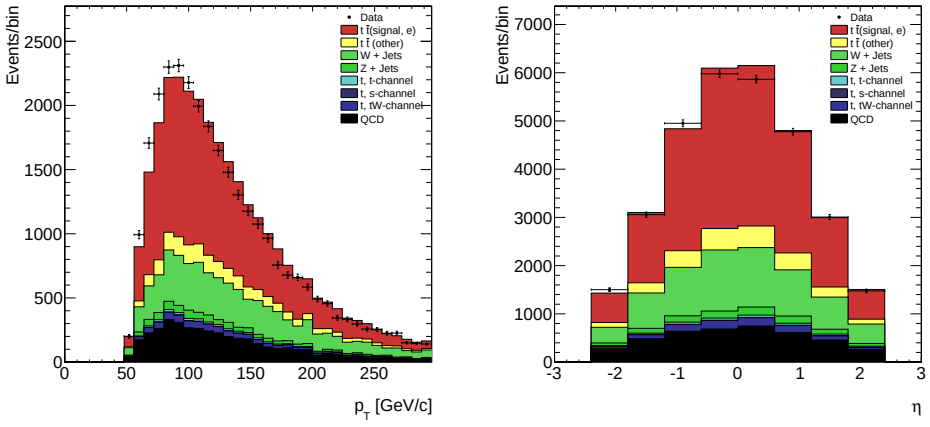
(a) Missing transverse energy.

(b) Total transverse momentum in the event H_T.

Figure 6.13.: Distribution of missing transverse energy and total transverse momentum in the event.

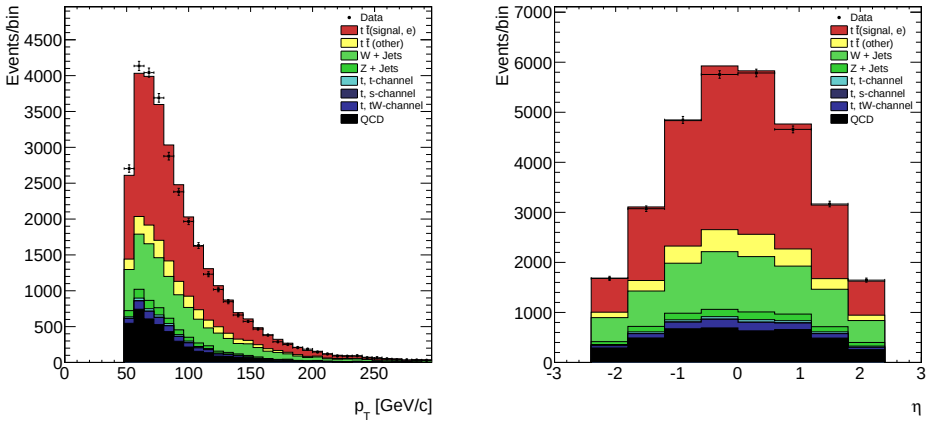
Figures 6.14 to 6.17 show the transverse momentum and the pseudorapidity distribution for the first four leading jets in transverse momentum in each event. For all four jets the agreement between prediction and collision data is good. The transverse momenta of the second, third and fourth jet agree very well, with maximum per-bin disagreements on the order of 2%, compatible within the statistical uncertainties. The leading jet in transverse momentum shows disagreements up to 10% per bin in the transverse momentum range $50\,\mathrm{GeV}/c < p_{\mathrm{T}} < 100\,\mathrm{GeV}/c$. The spectrum above $p_{\mathrm{T}} > 100\,\mathrm{GeV}/c$ is well described. The spectrum of the transverse momentum of all jets in the event and the associated pseudorapidity is very well modelled as can be seen from Figure 6.18.

Figure 6.19(a) shows the azimuthal angle ϕ for all jets in the event. The distribution is flat and shows a good agreement with the prediction as expected due to the cylindrical symmetry of the CMS detector. The number of jets observed in each event is shown in Figure 6.19. At large jet numbers with more than seven jets in the event the number of events is over-estimated. In the simulation only up to four additional jets are produced in the matrix element generator in W/Z + jets events and only up to three additional jets in top quark pair events. Additional jets are approximated through hadron showers in Pythia as described in Section 3.1. The disagreement at large jet quantities is most likely due to inaccuracies in



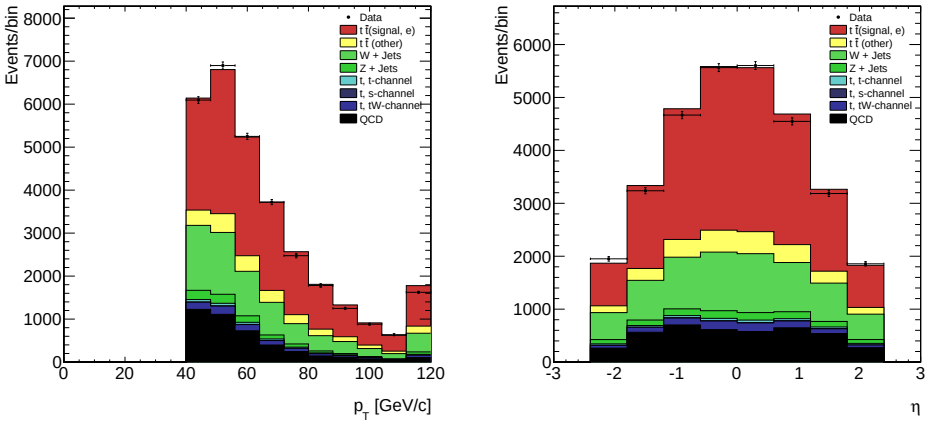
(a) First jet transverse momentum. (b) First jet pseudorapidity.

Figure 6.14.: Distribution of transverse momentum p_{T} and pseudorapidity η for leading jet in transverse momentum.



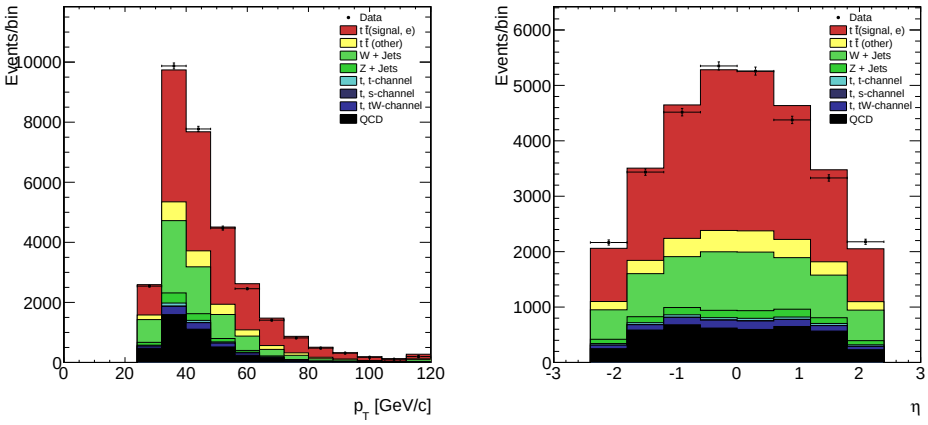
(a) Second jet transverse momentum. (b) Second jet pseudorapidity.

Figure 6.15.: Distribution of transverse momentum p_{T} and pseudorapidity η for second-leading jet in transverse momentum.



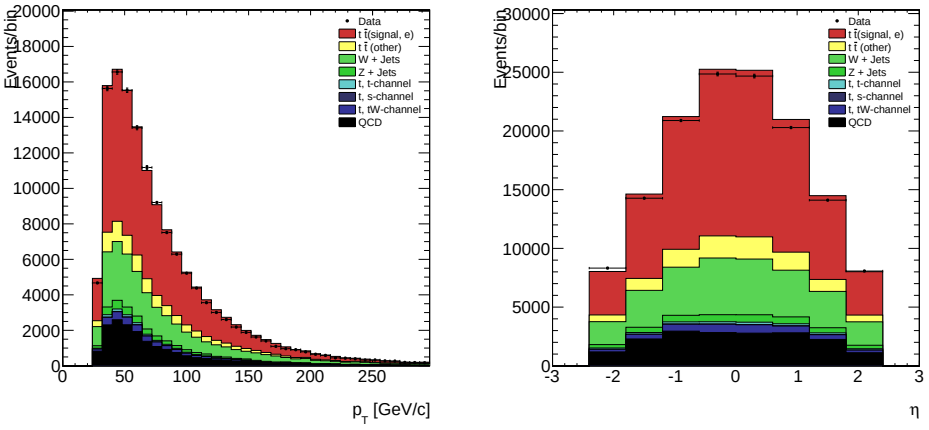
(a) Third jet transverse momentum. (b) Third jet pseudorapidity.

Figure 6.16.: Distribution of transverse momentum p_{T} and pseudorapidity η for third-leading jet in transverse momentum.



(a) Fourth jet transverse momentum. (b) Fourth jet pseudorapidity.

Figure 6.17.: Distribution of transverse momentum p_{T} and pseudorapidity η for fourth-leading jet in transverse momentum.



(a) All jets transverse momentum.

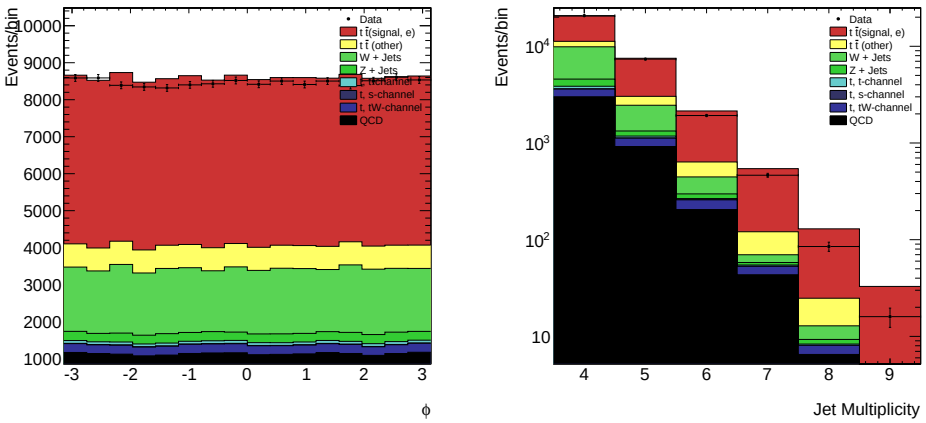
(b) All jets pseudorapidity.

Figure 6.18.: Distribution of transverse momentum p_{T} and pseudorapidity η for all jets in the event.

the modelling of the hadron shower, but does not affect the measurement requiring four or more jets.

The number of primary vertices in the event as an indicator for the number of pile-up interactions per event is shown in Figure 6.20(a). A good agreement is observed which confirms the performance of the pile-up re-weighting procedure described in Section 4.2.2.

Figure 6.20(b) shows the distance $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ between the selected electron and the closest jet. The distribution of multijet events is very different from what is observed in the collision data after the final event selection. This behaviour is expected due to the multijet shape extraction using the anti-relative-isolation method described in Section 6.2.3. The assumption is made that the multijet M3 mass shape is independent of the isolation of the lepton in the event. However, by definition multijet events do not feature any prompt leptons. This implies a physics correlation between the jet and the lepton in multijet events. Figure 6.21 shows the distance ΔR between the electron and closest jet in the event as a function of the relative isolation of the electron. No events with a low relative isolation are observed within the distance $\Delta R(e, \text{jet}) < 0.5$. A low relative isolation value implies an average distance of $\Delta R \approx 1.5$ between electron and closest jet. A close distance of electron and jet result in a large relative isolation value. Applying an upper limit to the electron isolation implicitly prevents electrons to be close to jets. When inverting the isolation requirement to extract the multijet shape a certain energy must be deposited near the lepton which can come only from the jet. In conclusion, relative Particle Flow isolation and the distance ΔR(jet, electron) are both isolation definitions which are largely correlated and lead to the observed behaviour in Figure 6.20(b). A hypthetical requirement of a minimum distance in

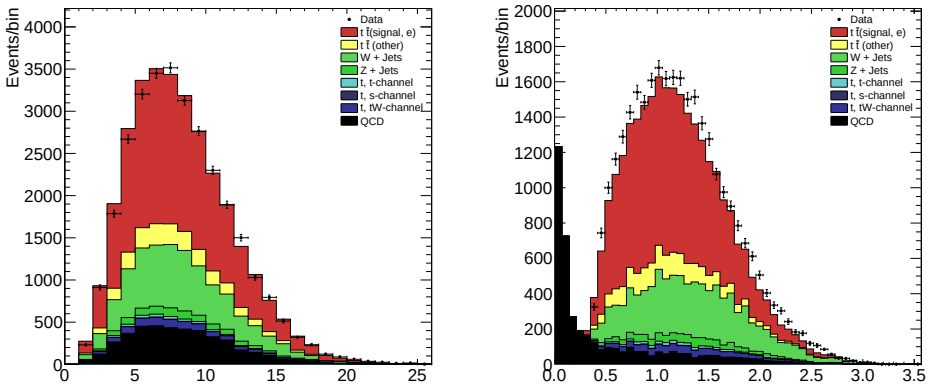


(a) All jets azimuthal angle.

(b) Number of Jets in the Event.

Figure 6.19.: Distribution of azimuthal angle of all jets in the event and the number of jets per event.

ΔR between the electron and the closest jet in the multijet shape extraction would result in a requirement of isolation and non-isolation at the same time.



(a) Number of Primary Vertices per Event. (b) Minimal distance between jet and electron ΔR.

Figure 6.20.: Distribution of the number of primary vertices per event and minimal distance $\Delta R(\text{jet}, e)$ between electron and closest jet.

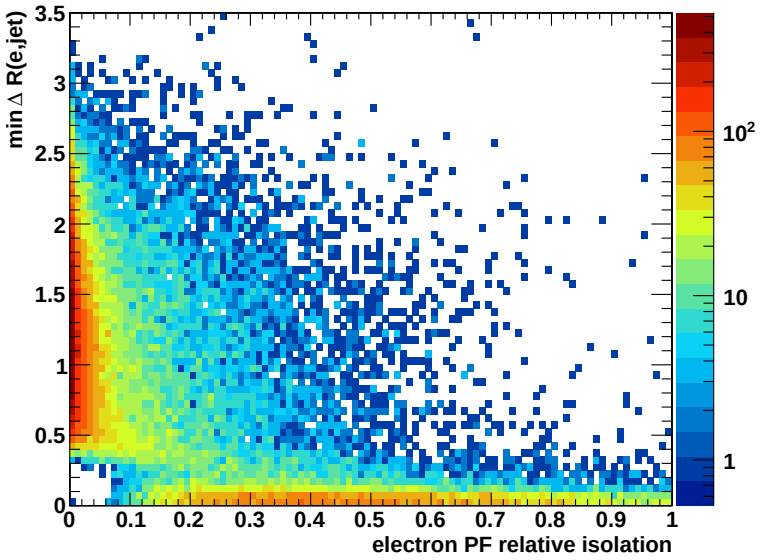


Figure 6.21.: Distribution of ΔR(jet, electron) for the electron and the closest jet in the event as a function of the Particle Flow relative isolation of the electron in the event.

7. Systematic Uncertainties on the Measurement of the Top Quark Pair Production Cross Section

The measurement of the top quark production cross section is described in Chapter 6. The cross section is measured in the semi-leptonic electron + jets channel as $155.5 \pm 3.7\,\text{pb}$. The uncertainty quoted on the measurement is the statistical uncertainty due to the template fit technique. Here, the relevant systematic uncertainties that influence the cross section measurement are evaluated.

7.1. Methodology for the Evaluation of Systematic Uncertainties

In general, a systematic uncertainty changes the result of the cross section measurement in two ways. First, a systematic change in a specific variable can change the top quark pair selection efficiency. The effect of the change in selection efficiency propagates directly to the final cross section measurement from Equation 6.1. Second, the measured cross section can change due to a change in the shape of the M3 mass templates used in the template fit. These two effects can enhance or counteract each other.

Template shapes for given distributions are re-derived after a systematic variation of the parameter under study. If top quark pair events are influenced by the systematic uncertainty, the signal selection efficiency is also re-calculated. The template fit is repeated on collision data with the varied template shapes. This yields an updated measured top quark pair production cross section value of

$$\sigma_{t\bar{t}}^{\text{systematic}} = \frac{N_{t\bar{t}}^{\text{systematic}}}{\int \mathrm{d}tL \cdot \varepsilon_{t\bar{t}}^{\text{systematic}}}, \tag{7.1}$$

where $N_{t\bar{t}}^{\text{systematic}}$ is the number of top quark pair events estimated from the template fit using systematically varied M3 mass template shapes, $\int \mathrm{d}L = 4.6\,\text{fb}^{-1}$ is the integrated luminosity of the data under consideration and $\varepsilon_{t\bar{t}}^{\text{systematic}}$ is the top quark pair selection efficiency after re-evaluation on the systematically varied simulated sample.

The systematic uncertainty on the final measurement δ for a given source of

systematic uncertainty, is then calculated as

$$\delta = \sigma_{t\bar{t}}^{\text{systematic}} - \sigma_{t\bar{t}},$$

where $\sigma_{t\bar{t}} = 155.5\,\text{pb}$ is the result of the nominal cross section measurement and $\sigma_{t\bar{t}}^{\text{systematic}}$ is the result of the cross section measurement after systematically varying a given parameter.

The systematic uncertainties are also evaluated on 5000 pseudo-experiments. The pseudo-experiments are generated according to the description in Section 6.3.1 using the nominal distributions for the pseudo-experiment generation. For each pseudo-experiment a cross section measurement is performed, fitting the systematically changed template shapes to the pseudo-data. The overall systematic uncertainty on the final cross section measurement in pseudo-experiments is evaluated as the difference of the nominal cross section measured in pseudo-experiments and the mean of a Gaussian function fit to the distribution of the systematically changed cross section measurements. The estimate on pseudo-experiments gives a prediction of the sensitivity of the result on the source of systematic uncertainty.

7.2. Sources of Systematic Uncertainties

An overview of the sources of systematic uncertainties and their impact on the cross section estimation is presented in this section.

7.2.1. Statistical Uncertainty on Template Shapes

The template shapes from data and simulation are not exactly known. They are subject to statistical fluctuations in the individual bins of the given distribution. The top quark template shape is available with a statistical precision sufficiently high as discussed in Appendix A.2.3. However, the statistical uncertainty on the shape of the dominant background processes is not negligible.

All template shapes derived according to the explanation in Section 6.2 are individually bin-by-bin fluctuated according to a Poisson distribution within their statistical uncertainty of $\sigma_{\text{Poisson}} = \sqrt{N_{\text{bin}}}$, where N_{bin} is the number of events in a given bin. The template shapes constructed by this method are used in the template fit to re-evaluate the top quark pair production cross section. This procedure is repeated 5000 times. This systematic uncertainty is estimated only on collision data and not on pseudo-experiments as the pseudo-data are generated from the same templates that are used to generate the statistically varied template shapes in the estimation of this systematic uncertainty. Possible fluctuations would be propagated to both shapes and the estimate would not be meaningful.

Figure 7.1 shows the estimated number of top quark pair events after each fit to collision data with statistically varied templates. The resulting systematic uncer-

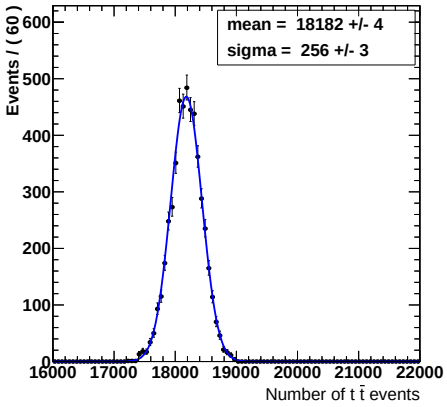


Figure 7.1.: Spread of the number of top quark pair events on collision data after varying the bin content of every template within its Poisson distributed uncertainty.

tainty on the cross section measurement is

$$\delta_{\text{template unc.}} = \pm 2.2\,\text{pb}.$$

7.2.2. Pile-Up Interactions

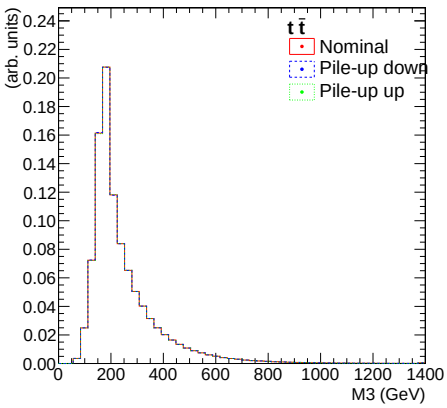
The estimate on the number of pile-up interactions in data has an overall systematic uncertainty of 5%. The dominant sources for systematic uncertainties are the luminosity uncertainty, the measurement of the total inelastic cross section in 2011, smearing effects of the pile-up interaction distribution and Monte Carlo generator modelling of pile-up interactions [132].

The measured uncertainty on the final cross section measurement is $+0.3\,\text{pb}/-0.1\,\text{pb}$ on collision data and $\pm 0.3\,\text{pb}$ on pseudo-experiments.

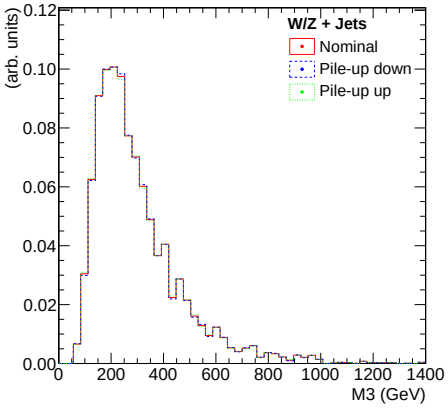
The top quark pair event selection efficiency changes from its nominal value of 0.02502 to 0.02499 for a downward variation of the effect and to 0.02510 for an upward variation of the effect.

7.2.3. Luminosity Uncertainty

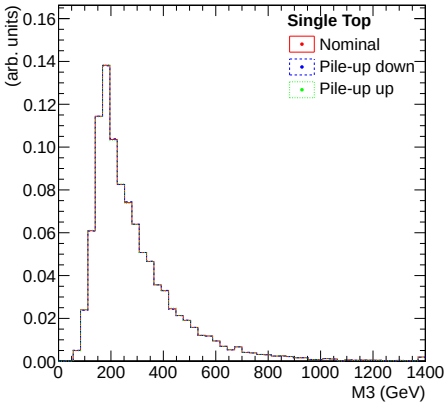
The uncertainty on the measurement of the integrated luminosity in the 2011 data taking period is estimated as 2.2% [68]. The sources of the uncertainty are explained together with the method for the luminosity measurement in Section 2.5. This uncertainty propagates directly to the measurement of the top quark pair production cross section with Equation 6.1. The systematic uncertainty on the final cross section measurement is $-3.3\,\text{pb}/+3.5\,\text{pb}$ on collision data and $-3.6\,\text{pb}/+3.7\,\text{pb}$ on pseudo-experiments.



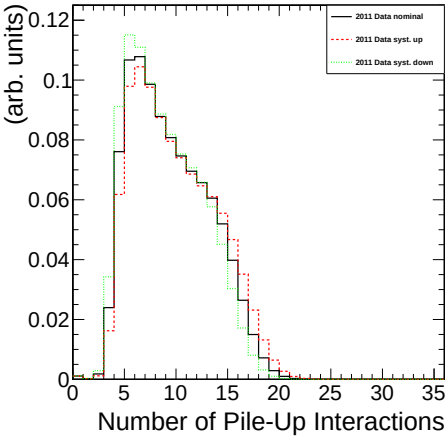
(a) Top quark pair template shape variation.



(b) W/Z + jets template shape variation.



(c) Single top quark template shape variation.



(d) Variation of estimated number of pile-up interactions in collision data.

Figure 7.2.: Figures 7.2(a) to 7.2(c) show the systematically varied template shapes due to the uncertainty on the number of pile-up interactions. The multijet shape is unchanged as it is measured in collision data. Figure 7.2(d) shows the variation of the estimation of the number of pile-up interactions in collision data due to a change of ±5% to the minimum bias event cross section.

7.2.4. Factorisation Scale

The factorisation scale defines the amount of the squared transverse momentum used in the evolution of the parton shower simulation. It is defined as

$$Q^2 = M^2_{t/W/Z} + \sum_{\text{jets}} (p_{\text{T}}^{\text{jet}})^2,$$

where $M_{t/W/Z}$ is the mass of the top quark, the W boson or the Z boson. This value is varied by a factor of two up and down. Simulated samples with varied factorisation scale are used to estimate the systematic uncertainty. They are listed in Table A.5 Appendix A.2.2. These samples contain one order of magnitude less events than the nominal samples. To estimate the statistical accuracy of the observed systematic uncertainty, the extracted template shapes are fluctuated per bin within their Poisson distributed uncertainties and the fit to collision data is repeated to estimate the effect of the low statistical precision of the shapes. The procedure is repeated 5000 times. The spread of the result of the fit is quoted as statistical uncertainty on the systematic uncertainty on collision data. This information is not used in the final estimation of the uncertainty on the cross section measurement and merely gives an estimate of the statistical significance of the estimate of the systematic uncertainties. The uncertainty due to factorisation scale is treated correlated for W + jets and Z + jets but is treated uncorrelated to top quark pair events.

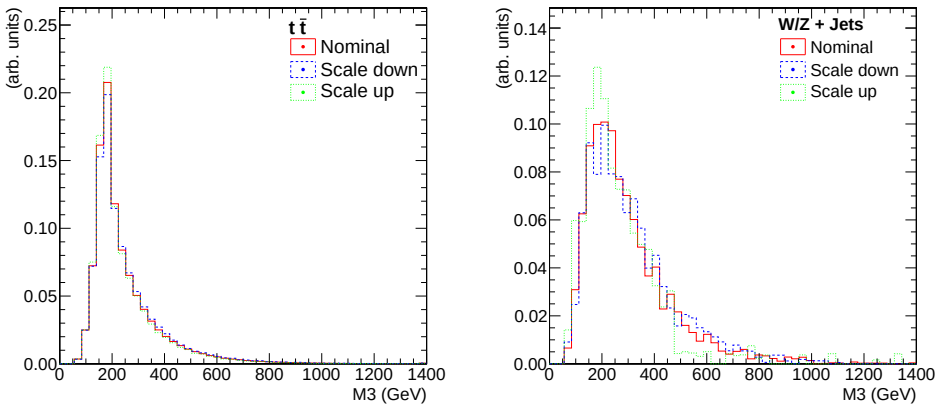
The estimated uncertainty on the final cross section measurement due to the factorisation scale uncertainty in top quark pair events is $-4.0 \pm 2.0\,\text{pb} / + 5.5 \pm 1.9\,\text{pb}$ in collision data and $-4.0\,\text{pb} / + 5.2\,\text{pb}$ in pseudo-experiments for upward and downward variations respectively.

The estimated uncertainty on the cross section measurement due to the factorisation scale uncertainty in W/Z + jets events is $-5.3 \pm 1.5\,\text{pb} / + 4.9 \pm 1.6\,\text{pb}$ in collision data and $-5.5\,\text{pb} / + 1.0\,\text{pb}$ in pseudo-experiments for upward and downward variations respectively.

Figure 7.3 shows the variation of the template shapes due to the uncertainty on the factorisation scale in top quark pair events and W/Z + jets events. The top quark pair event selection efficiency changes from its nominal value of 0.02502 to 0.02684 for a downward variation of the effect and to 0.02334 for an upward variation of the effect.

7.2.5. Jet - Parton Matching Threshold

The matching between the underlying partons of the hard interaction in Matrix Element calculations with the jets created in the parton showering done with Pythia is described in Sections 1.3.2 and 3.1. It is done according to the so-called *MLM scheme* [133]. The systematic uncertainties on the cross section measurement are estimated by adjusting the parameters for minimum transverse momentum for jets to be considered in the matching $p_{\text{T}}^{\text{min}}(\text{jet})$. Simulated samples are used with var-



(a) Top quark pair template shape variation. (b) W/Z + jets template shape variation.

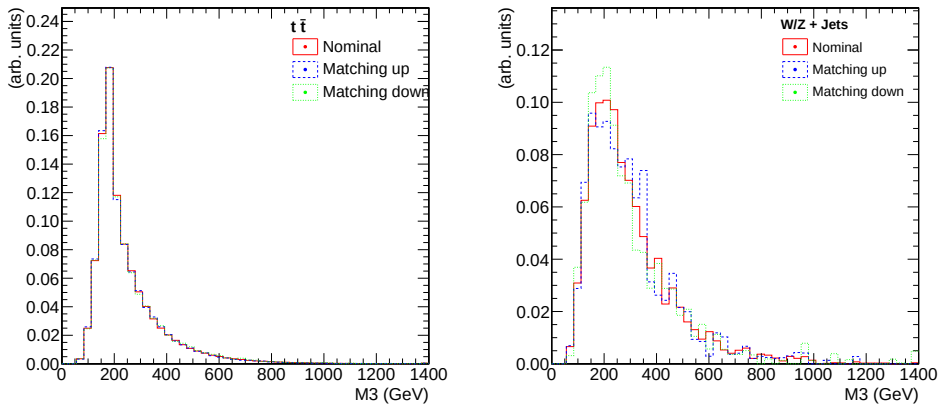
Figure 7.3.: Systematically varied template shapes due to the uncertainty on factorisation scale. The multijet shape is unchanged as it is measured in collision data. The single top quark templates remain unvaried as the single top contribution is small and the additional uncertainty not significant.

ied jet-parton matching threshold to estimate this systematic effect. A list of the samples used can be found in Table A.4. As in the case of the estimation of the systematic uncertainty due to the factorisation scale, described in Section 7.2.4, the statistical precision is lower in the systematically varied samples than in the nominal samples. The statistical uncertainty on the evaluation of the systematic uncertainty due to the jet-parton matching threshold is estimated following the description in Section 7.2.4. The uncertainty due to jet-parton matching is treated correlated for W + jets and Z + jets but is treated uncorrelated to top quark pair events.

For W/Z + jets events the nominal value for the jet-parton matching threshold is $10\,\mathrm{GeV}/c$. This value is varied to be $20\,\mathrm{GeV}/c$ and $5\,\mathrm{GeV}/c$. For top quark pair events the nominal value for the matching threshold is $20\,\mathrm{GeV}/c$. It is varied to be $40\,\mathrm{GeV}/c$ and $10\,\mathrm{GeV}/c$ [134].

The variation of the jet-parton matching threshold in W/Z + jets events results in a systematic uncertainty of $-1.2 \pm 1.1\,\mathrm{pb} / -4.4 \pm 0.9\,\mathrm{pb}$ on collision data and $-1.9\,\mathrm{pb} / -5.4\,\mathrm{pb}$ on pseudo-experiments. The measured cross section value changes systematically in the same direction after increasing and decreasing the threshold value. This is explained by the shape changes in the M3 mass template shapes. The resulting shape peaks towards lower values of M3 mass after both up and down changes as can be seen in Figure 7.4(b).

The variation of the jet-parton matching threshold in top quark pair events results in a systematic uncertainty of $-0.6 \pm 1.9\,\mathrm{pb} / -3.6 \pm 2.3\,\mathrm{pb}$ on collision



(a) Top quark pair template shape variation. (b) W/Z + jets template shape variation.

Figure 7.4.: Systematically varied template shapes due to the uncertainty on jet-parton matching threshold. The multijet shape is unchanged as it is measured in collision data. The single top quark templates remain unvaried as the single top contribution is small and the additional uncertainty not significant.

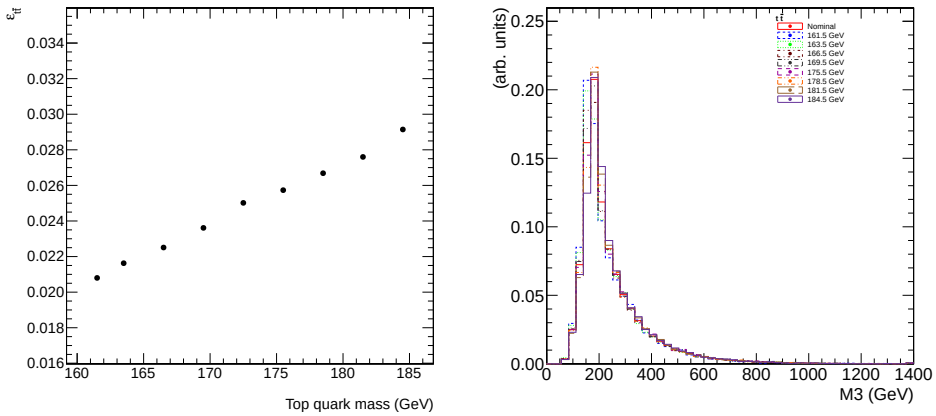
data and $-0.8\,\text{pb}/-3.4\,\text{pb}$ on pseudo-experiments. As for the W/Z + jets, the measured cross section value changes systematically in the same direction after increasing and decreasing the threshold value. The given explanation also holds here. The change of the top quark pair shape is shown in Figure 7.4(a). The simulation predicts the observed asymmetric uncertainty for W/Z + jets and $t\bar{t}$, which serves as validation of the evaluation of the matching threshold uncertainty.

The top quark pair event selection efficiency changes from its nominal value of 0.02502 to 0.02603 for a downward variation of the effect and to 0.02482 for an upward variation of the effect.

7.2.6. Top Quark Mass

The mass of the top quark is varied within the uncertainties of the current world average of $173.5 \pm 0.6 \pm 0.8\,\text{GeV}/c^2$ [2]. In the simulation a value of $172.5\,\text{GeV}/c^2$ is assumed. To compensate for this discrepancy an asymmetric error band corresponding to $172.5^{+2.0}_{-0.0}\,\text{GeV}/c^2$ is propagated.

Eight simulated samples are used with varied top masses between 161.5 GeV and 184.5 GeV. Table A.6 lists the samples used. Figure 7.5(a) shows the variation of the top quark pair selection efficiency as a function of the top quark mass. Figure 7.5(b) shows the change of the top quark pair M3 mass template shape with different top quark mass assumptions. The M3 mass peak shifts towards lower values with decreasing top mass assumption and towards higher values with increasing top mass



(a) Top quark pair selection efficiency variation.

(b) Top quark pair template shape variation.

Figure 7.5.: Change of top quark pair selection efficiency and template shape as a function of the top quark mass.

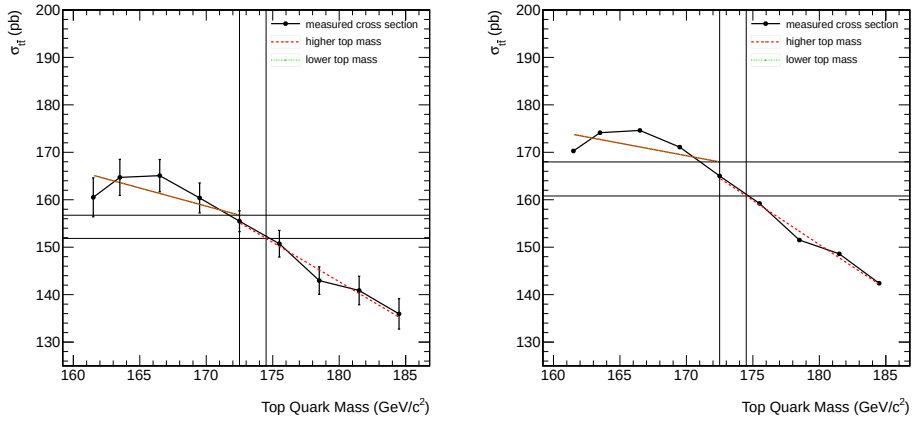
value. Figure 7.6 shows the simultaneous variation of template shape and selection efficiency.

The variation of the fitted cross section is interpolated between the eight mass points in positive and negative direction of systematically shifted top quark masses to the reference mass by a straight line fit. Figure 7.6 presents the results of the procedure. Figure 7.7 repeats the procedure twice, once keeping the nominal top quark pair template shape constant and varying the top quark pair selection efficiency, and once keeping the nominal top quark pair selection efficiency constant and changing the top quark pair template shape. This helps to understand the composition of the systematic uncertainty due to the two effects. The systematic variation due to the change of the selection efficiency is a linear function of the top quark mass while the systematic variation due to the change of the shape of the top quark template is rather flat and decreases towards lower top quark mass assumptions.

The uncertainty on the final measurement due to the top quark mass uncertainty is $-3.6\,\text{pb}/+1.3\,\text{pb}$ on collision data and $-4.2\,\text{pb}/+2.9\,\text{pb}$ on pseudo-experiments.

7.2.7. W/Z boson Cross Section Ratio and Single Top Quark Cross Section

When constructing the combined W/Z + jets template the W + jets template and the Z + jets template are added according to the ratio of the selection efficiencies and the production cross section of the two processes listed in Table 4.1. A conservative uncertainty of 30% is assumed on this ratio. A new set of combined



(a) Change of selection efficiency and shape for collision data.

(b) Change of selection efficiency and shape for pseudo-experiments.

Figure 7.6.: Systematic uncertainty on cross section measurement in electron + jets channel due to the top quark mass on collision data and on pseudo-experiments.

W/Z + jets templates is constructed, one with a 30% lower cross section ratio and one assuming a 30% higher ratio. The top quark cross section is re-evaluated using these templates for W/Z + jets. Figure 7.8 shows the re-derived template shapes after the variation of the W/Z + jets cross section ratio. The effect is very small due to the dominance of the W + jets contribution.

The resulting systematic uncertainty on the final cross section measurement is $+0.2\,\text{pb}/-0.3\,\text{pb}$ on collision data and $+0.1\,\text{pb}/-0.4\,\text{pb}$ on pseudo-experiments.

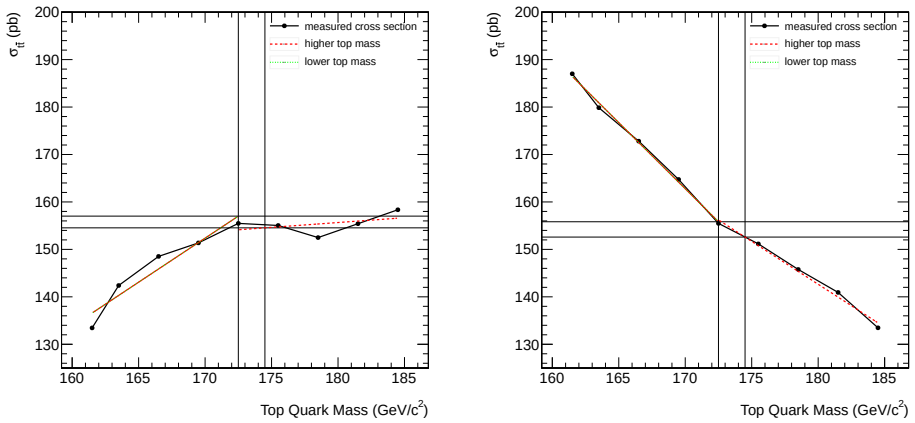
7.2.8. Jet Energy Scale and Jet Energy Resolution

The influence of the jet energy scale and resolution on the cross section measurement is studied because the M3 mass shape involves three jets with corrections applied.

Jet Energy Resolution

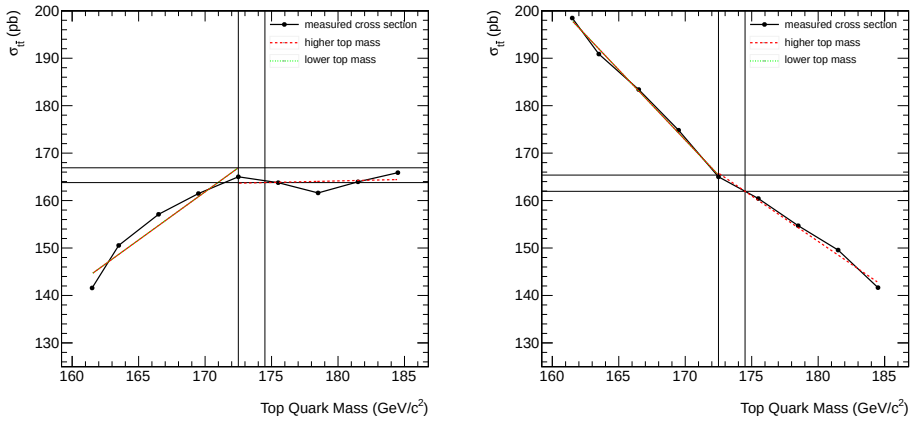
Section 4.2.5 describes how the jet energy resolution in the simulated samples is tuned to the average value on collision data. The uncertainties on the jet energy resolution are estimated by varying the core resolution scale factor according to its uncertainties listed in Table 4.5. Then the jet energy resolution for all jets is adapted following the prescription from Section 4.2.5 and the selection is re-run, including the M3 mass calculation.

Figure 7.9 shows the variation of the template shapes due to the change of the jet energy resolution in the simulation. The multijet shape is unchanged as it is mea-



(a) Change of shape with constant selection efficiency for collision data.

(b) Change of selection efficiency with constant mass shape for collision data.



(c) Change of shape with constant selection efficiency for pseudo-experiments.

(d) Change of selection efficiency with constant mass shape for pseudo-experiments.

Figure 7.7.: Systematic uncertainty on cross section measurement in electron + jets channel due to the top quark mass on collision data and on pseudo-experiments. The effect of the shape influence and of the selection efficiency influence are decoupled. Figure 7.7(a) and Figure 7.7(c) show the cross section measurement after systematically changing the M3 mass shape and keeping the top quark pair selection efficiency constant. Figure 7.7(b) and Figure 7.7(d) show the cross section measurement after systematically changing the selection efficiency and using the nominal shape of the M3 mass.

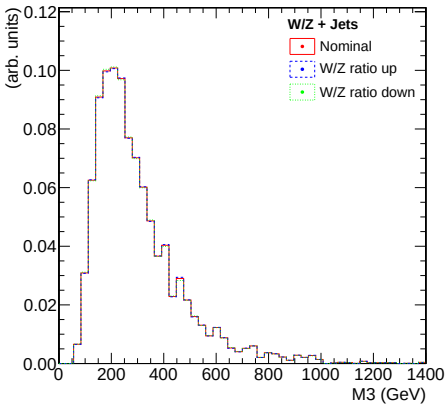


Figure 7.8.: Systematically varied template shapes due to the uncertainty on W/Z + jets cross section ratio.

sured in collision data. The effect of the variation of the simulated samples for the prompt-electron contamination removal is well below the statistical uncertainties of the multijet shape and is therefore not propagated. Table 7.1 lists the uncertainties on the final cross section measurement on collision data and pseudo-experiments.

The top quark pair event selection efficiency changes from its nominal value of 0.02502 to 0.02496 for a downward variation of the effect and to 0.02511 for an upward variation of the effect.

Jet Energy Scale

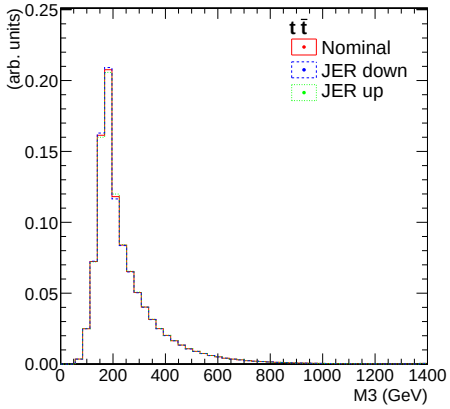
The uncertainty on the jet energy calibration constants has been discussed in Section 2.4.6 and is propagated to the final cross section measurement. The four vector of all jets is scaled up and down within one sigma of the uncertainty on the jet energy. The jet energy scale uncertainty, like the jet energy scale, is parametrised as a function of jet transverse momentum and pseudorapidity of the jet. The selection is repeated and the M3 mass re-calculated.

Figure 7.10 shows the change in the M3 mass shape due to the uncertainties on the jet energy scale. Table 7.1 lists the uncertainties on the final cross section measurement on collision data and pseudo-experiments.

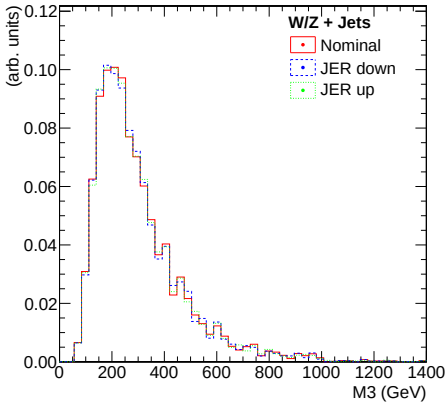
The top quark pair event selection efficiency changes from its nominal value of 0.02502 to 0.02381 for a downward variation of the effect and to 0.02612 for an upward variation of the effect.

7.2.9. Electron Energy Scale

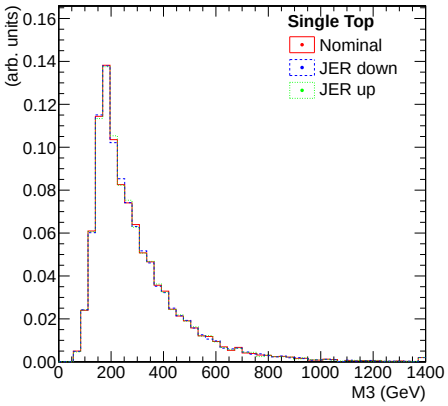
An uncertainty of 1% on the electron energy scale is propagated to the final cross section measurement to account for residual loss of crystal transparency during



(a) Top quark pair template shape variation.

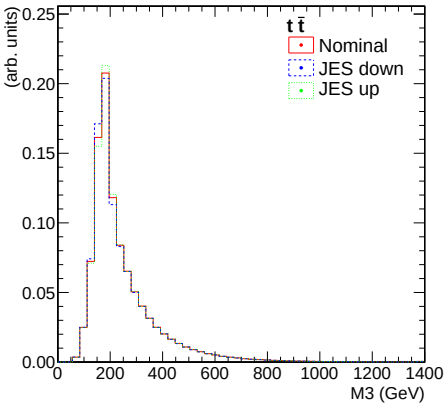


(b) W/Z + jets template shape variation.

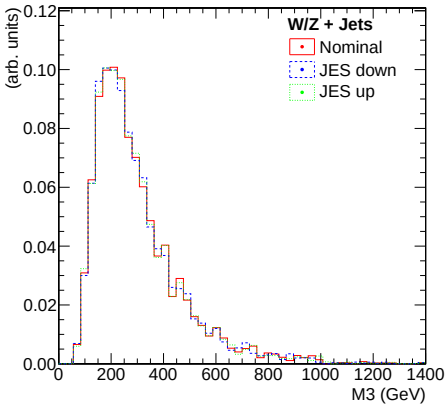


(c) Single top quark template shape variation.

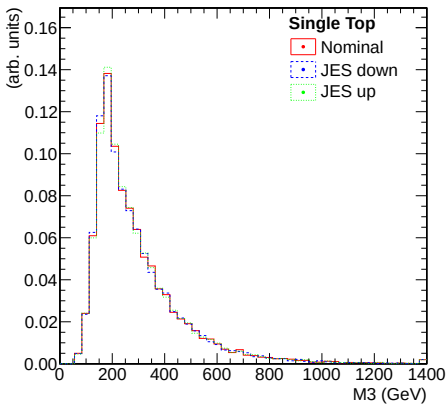
Figure 7.9.: Systematically varied template shapes due to the uncertainty on jet energy resolution. The multijet shape is unchanged as it is measured in collision data.



(a) Top quark pair template shape variation.



(b) W/Z + jets template shape variation.



(c) Single top quark template shape variation.

Figure 7.10.: Systematically varied template shapes due to the uncertainty on jet energy scale. The multijet shape is unchanged as it is measured in collision data.

Table 7.1.: Summary of the systematic uncertainty due to the jet energy scale and resolution uncertainty. *Up* and *Down* refer to an upward and downward variation of the parameter under study within its uncertainties with respect to the nominal value.

Source	Uncertainty (pb)			
	Data		Pseudo-Experiment	
	Up	Down	Up	Down
Jet Energy Resolution	+0.8	−1.7	+0.6	−1.7
Jet Energy Scale	−7.8	+4.2	−8.8	+4.1

data taking compared to the simulation. The uncertainty on the cross section measurement is $-1.1\,\text{pb}/+1.0\,\text{pb}$ in collision data and $-1.2\,\text{pb}/+1.0\,\text{pb}$ in pseudo-experiments.

The top quark pair event selection efficiency changes from its nominal value of 0.02502 to 0.02484 for a downward variation of the effect and to 0.02519 for an upward variation of the effect.

7.2.10. Lepton Identification and Trigger Efficiency Measurements

The uncertainties on the lepton identification and trigger scale factor measurement have been discussed in detail in Chapter 5. The final scale factors and their uncertainties are listed in Table 5.3 and in Table 5.4. In summary, the uncertainty on the lepton identification and isolation scale factor is 0.5% each. The uncertainty on the conversion rejection scale factor is 1%. The uncertainty on the trigger scale factor is 1% on the scale factor measurement of the hadronic part of the trigger and 1% on the scale factor estimate of the leptonic part. The systematic uncertainties do not affect the template shapes and can be directly propagated to the cross section measurement by means of a change of the top quark pair selection efficiency. Table 7.2 summarises the uncertainties on the top quark pair production cross section measurement.

7.2.11. Multijet Shape Estimation

There are two sources of systematic uncertainty considered on the estimate of the multijet shape. First, the stability of the fit using different slices of relative isolation for the multijet shape extraction is studied. Second, the influence of the assumed cross section values on the contamination removal is evaluated.

Table 7.2.: Summary of the systematic uncertainty due to the lepton and trigger scale factor uncertainty.

Source	Uncertainty (pb)			
	Data		Pseudo-Experiment	
	Up	Down	Up	Down
Trigger (e part)	-1.6	$+1.6$	-1.7	$+1.6$
Trigger (jet part)	-1.6	$+1.6$	-1.7	$+1.6$
Conv. Rej. Eff.	-1.6	$+1.6$	-1.7	$+1.6$
Electron Iso. Eff.	-0.8	$+0.8$	-0.9	$+0.9$
Electron ID Eff.	-0.8	$+0.8$	-0.9	$+0.9$

Stability of Multijet Shape in Slices of Relative Particle Flow Isolation

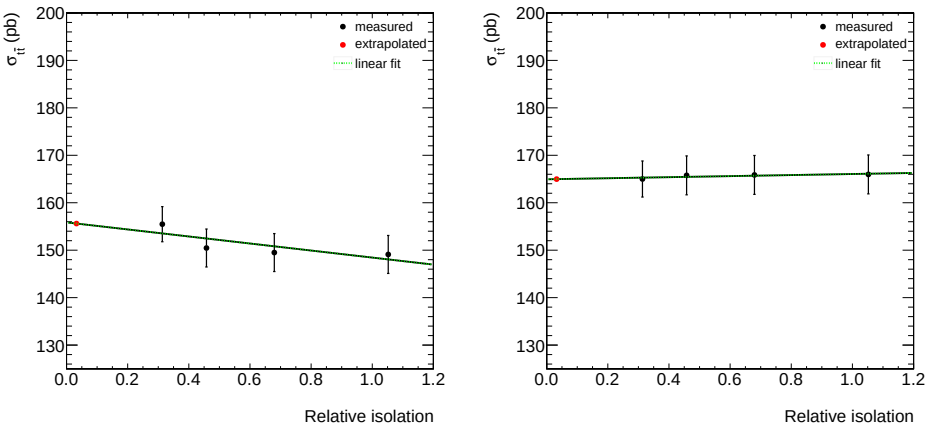
In Section 6.2.3 it is explained that the multijet shape is extracted in different slices of relative isolation. The individual slices are listed in Table 6.1 and depicted in Figure 6.3. The slice closest to the signal region is chosen as nominal multijet shape because the signal contamination is already very low and possible systematic effects are minimised. The cross section measurement is repeated with the multijet shape from each individual slice. A cross check for the cross section in the signal region is derived through a linear fit to the individual measurements as a function of the mean value of the relative isolation slice. This makes it possible to extrapolate the cross section measurement to the signal region. The difference between the cross section measurement using the nominal slice of $0.25 < \mathrm{Iso}_{\mathrm{rel.}}(e) < 0.35$ and the extrapolated value is quoted as systematic uncertainty.

Figure 7.11 shows the result of the cross section measurement for the individual relative Particle Flow isolation slices. The systematic uncertainty is $-0.1\,\mathrm{pb}$ on collision data and $-0.1\,\mathrm{pb}$ on pseudo-experiments. Although the slope of the extrapolation is different on data and on pseudo-experiments, the comparison of the nominal cross section measurement point and the extrapolated point yields a value of $-0.1\,\mathrm{pb}$ in both cases.

Influence of Contamination Removal on Multijet Shape

The procedure for the contamination removal of the residual contamination of prompt-electron events in the multijet shape is explained in Section 6.2.3. An uncertainty of 100% on the top quark pair production cross section and an uncertainty of 50% on the W/Z + jets cross section is assumed. The multijet template shape is re-derived using the modified normalisation for the prompt-electron contamination.

Figure 7.12 shows the varied template shape of the multijet background after subtracting no top quark pair events and only half the amount of W/Z + jets



(a) Measurement on collision data. (b) Measurement on pseudo-experiments.

Figure 7.11.: Cross section measurement with extracted multijet shapes from different relative Particle Flow isolation slices. An extrapolation to the signal region is performed using a linear function fit.

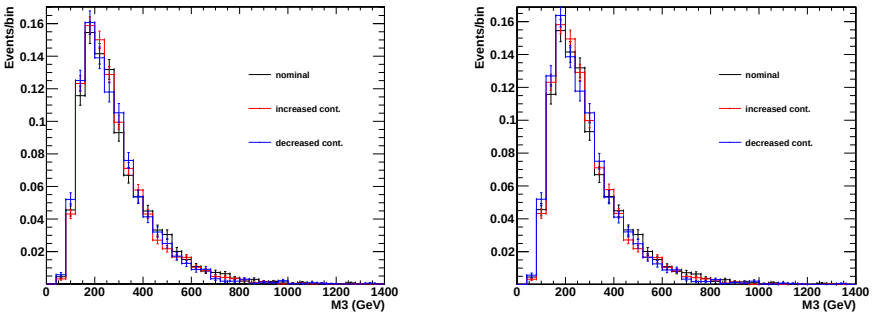


Figure 7.12.: Re-derived multijet shape after assuming a 100% uncertainty on the number of top quark pair events (left) and a 50% uncertainty on the number of W/Z + jets events (right).

events and after subtracting twice the amount of expected top quark pair events and 50% more W/Z + jets events than expected. The variations of the top quark pair cross section and of the W/Z + jets cross section are treated as uncorrelated. The variation of the W + jets and Z + jets cross section is treated correlated.

The resulting systematic uncertainty on the final cross section measurement due to the uncertainty on the top quark pair cross section is $+0.9 / -1.4\,\text{pb}$ on collision data and $+0.4 / -0.5\,\text{pb}$ on pseudo-experiments. The resulting systematic uncertainty on the final cross section measurement due to the uncertainty on the W/Z + jets cross section is $-0.9 / -1.3\,\text{pb}$ on collision data and $-0.5 / -0.5\,\text{pb}$ on pseudo-experiments. The shape change leads for upward and downward variations to a small decrease of the measured cross section. This is possibly due to the statistical uncertainty on the template shapes.

7.2.12. Parton Distribution Functions

The theoretical uncertainties on parton distribution functions of the colliding protons need to be considered for the evaluation of the uncertainty on the cross section measurement. The so-called Hessian Method [135] is employed to propagate the theoretical uncertainties from the parton distribution functions to the analysis. The Hessian Method both constructs a N eigenvector basis of parton distribution functions and provides a method from which uncertainties on observables are calculated [136]. To determine parton distribution functions and their uncertainties, a χ^2 fit to collision data using N free parameters describing the parton distribution functions is performed. The parameter values in the minimum of the χ^2 of the fit are used as the central value. Diagonalising the Hessian error matrix yields N eigenvectors. Each eigenvector is then varied within its uncertainties to obtain $2N$ new parameter sets, $S_i^{\pm}$. The parton distribution function set and uncertainties used within this analysis are taken from the CTEQ library in version 6.6 [72], where the Hessian error matrix from the CTEQ group contains $N = 22$ eigenvectors.

One employs the so-called "Master equations" [136] to evaluate the uncertainty on the final cross section number.

$$\Delta X^{+}_{\text{max}} = \sqrt{\sum_{i=1}^{N} [\max(X_i^{+} - X_0, X_i^{-} - X_0, 0)]^2} \tag{7.2}$$

$$\Delta X^{-}_{\text{max}} = \sqrt{\sum_{i=1}^{N} [\max(X_0 - X_i^{+}, X_0 - X_i^{-}, 0)]^2} \tag{7.3}$$

X_0 is the central value of the observable, $X_i^{\pm}$ are the values obtained for the observable after varying a certain eigenvector within its uncertainties. Here, X is the extracted top quark production cross section. It accounts for shape and signal efficiency changes determined via re-weighting the simulated events produced with

CTEQ in version 6.l to correspond to CTEQ in version 6.6 for the central X_0 and the two times 22 CTEQ 6.6 uncertainty sets for the $X_i^{\pm}$.

Evaluating the Master equations yields an uncertainty on the cross section measurement of $+8.4\,\text{pb}/-6.8\,\text{pb}$ on collision data and $+9.7\,\text{pb}/-8.0\,\text{pb}$ on pseudo-experiments. The variation of the template shapes due to the uncertainty on the parton distribution functions can be seen in Figure 7.13. The change in the selection efficiency for top quark pair events as a function of the parton distribution function uncertainty parameter set is presented in Figure 7.13(b).

The template shape change is the dominant source of the uncertainty as can be seen in Table 7.3.

7.3. Summary of Systematic Uncertainties on Cross Section Measurement

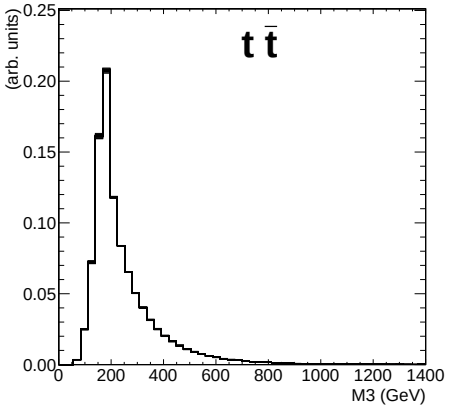
Table 7.3 compares the effect of the systematically changed template shapes to the effect of the top quark selection efficiency changes. The change of the selection efficiency is the dominant source of the systematic uncertainty for the jet energy scale, the top quark jet–parton matching uncertainty and the top quark mass uncertainty. The template shape change is the dominant source of the systematic uncertainty for the jet energy resolution, the parton distribution functions and the pile-up interaction uncertainty. A counteracting effect of shape change and selection efficiency change is observed for the top quark factorisation scale uncertainty which results in an reduced overall uncertainty compared to the individual components of the uncertainty. The systematically varied top quark selection efficiencies are listed in Table 7.4.

A summary of all systematic uncertainties on the top quark pair production cross section measurement in the electron + jets channel is given in Table 7.5. The table lists the uncertainties for the measurement on collision data and in the pseudo-experiment.

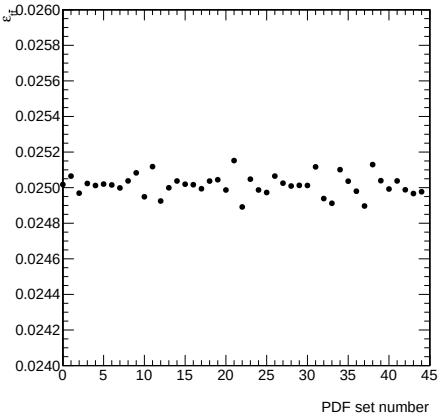
The dominant systematic uncertainties on the cross section measurement are the uncertainty due to the parton distribution functions, the jet energy scale, jet-parton matching and factorisation scale. One notices that all uncertainties are theoretical uncertainties except for the uncertainty due to the jet energy scale which can be reduced by improved jet calibrations and the uncertainty on the parton distribution functions. The latter is a combination of theoretical and experimental uncertainties which could be improved by more precise measurements in deep inelastic scattering or from measurements of W and Z boson production at the LHC. For all uncertainties the prediction from pseudo-experiments is comparable to the result obtained in collision data.

The expected sensitivity for the measurement based on pseudo-experiments is

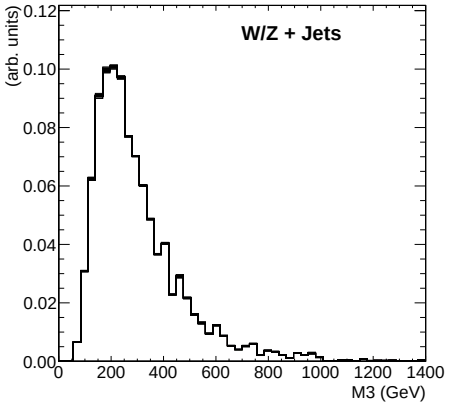
$$\sigma_{t\bar{t}} = 165.0 \pm 3.8\,(\text{stat.})\,^{+12.2}_{-16.2}\,(\text{syst.}) \pm 3.5\,(\text{lumi.})\,\text{pb}.$$



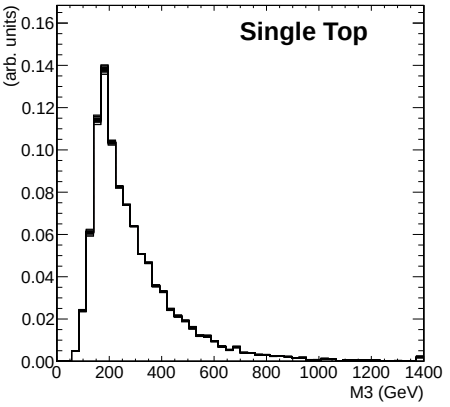
(a) Top quark pair template shape variation.



(b) Top quark pair selection efficiency.



(c) W/Z + jets template shape variation.



(d) Single top quark template shape variation.

Figure 7.13.: Systematically varied template shapes due to the uncertainty on parton distribution functions. The multijet shape is unchanged as it is measured in collision data. Figure 7.13(b) shows the top quark pair selection efficiency for different uncertainty sets of the parton distribution functions.

The total systematic uncertainty combines the individual systematic uncertainties in quadrature. In relative uncertainties the result of the measurement on pseudo-experiments can be expressed as

$$\sigma_{t\bar{t}} = 165.0\,\text{pb} \pm 2.3\,\%\,(\text{stat.})\,^{+7.4}_{-9.8}\,\%\,(\text{syst.}) \pm 2.2\,\%\,(\text{lumi.}).$$

Together with the information from Chapter 6 the result of the measurement in $L = 4.6\,\text{fb}^{-1}$ of collision data collected in 2011 yields

$$\sigma_{t\bar{t}} = 155.5 \pm 3.7\,(\text{stat.})\,^{+12.6}_{-14.8}\,(\text{syst.}) \pm 3.4\,(\text{lumi.})\,\text{pb}.$$

In relative uncertainties the result of the measurement on collision data is

$$\sigma_{t\bar{t}} = 155.5\,\text{pb} \pm 2.4\,\%\,(\text{stat.})\,^{+8.1}_{-9.5}\,\%\,(\text{syst.}) \pm 2.2\,\%\,(\text{lumi.}).$$

Table 7.3.: Comparison of systematic change by template shape change or top quark selection efficiency change in the electron + jets decay channel on collision data.

Source	Uncertainty (pb)						Dominant Effect
	Overall		Shape		Efficiency		
	Up	Down	Up	Down	Up	Down	
Jet Energy Resolution	+0.8	−1.7	+1.4	−2.0	−0.5	+0.4	Shape
Jet Energy Scale	−7.8	+4.2	−1.4	−3.5	−6.6	+7.9	Efficiency
$t\bar{t}$ Matching	−0.6	−3.6	−1.8	+2.5	+1.2	−6.1	Efficiency
$t\bar{t}$ Fact. Scale	−4.0	+5.5	−14.2	+17.3	+11.2	−10.6	Shape
Top Quark Mass	−3.6	0.0	−0.9	0.0	−2.9	0.0	Efficiency
Parton Distribution Functions	+8.4	−6.8	+9.2	−7.6	+1.7	−1.6	Shape
Pile-Up	+0.3	−0.1	−0.3	0.4	+0.1	−0.1	Shape

Table 7.4.: Comparison of systematic change of top quark selection efficiency for different sources of systematic uncertainties in the electron + jets decay channel on collision data. The nominal top quark pair selection efficiency is 0.02502.

Source	Efficiency	
	Up	Down
Jet Energy Resolution	0.02511	0.02496
Jet Energy Scale	0.02612	0.02381
$t\bar{t}$ Matching	0.02482	0.02603
$t\bar{t}$ Fact. Scale	0.02334	0.02684
Top Quark Mass	0.02541	n/a
Pile-Up	0.02499	0.02510

Table 7.5.: Summary of the systematic uncertainties on the measurement of the top quark pair production cross section in the electron + jets decay channel.

Source	Uncertainty (pb)			
	Data		Pseudo-Experiment	
	Up	Down	Up	Down
Jet Energy Resolution	$+0.8$	-1.7	$+0.6$	-1.7
Jet Energy Scale	-7.8	$+4.2$	-8.8	$+4.1$
$t\bar{t}$ Matching	-0.6 ± 1.9	-3.6 ± 2.3	-0.8	-3.4
W/Z + jets Matching	-1.2 ± 1.1	-4.4 ± 0.9	-1.9	-5.4
$t\bar{t}$ Fact. Scale	-4.0 ± 2.0	$+5.5 \pm 1.9$	-4.0	$+5.2$
W/Z + jets Fact. Scale	-5.3 ± 1.5	$+4.9 \pm 1.6$	-5.5	$+1.0$
Multijet Cont. Rem. ($t\bar{t}$)	$+0.9$	-1.4	$+0.4$	-0.5
Multijet Cont. Rem. (W/Z + jets)	-0.9	-1.3	-0.5	-0.5
Multijet Stability	0.0	-0.1	0.0	-0.1
Stat. Shape Unc.	$+2.2$	-2.2	n/a	n/a
Top Quark Mass	-3.6	0.0	-4.2	0.0
Parton Distribution Functions	$+8.4$	-6.8	$+9.7$	-8.0
Pile-Up	$+0.3$	-0.1	$+0.3$	-0.3
Trigger (e part)	-1.6	$+1.6$	-1.7	$+1.6$
Trigger (jet part)	-1.6	$+1.6$	-1.7	$+1.6$
Conv. Rej. Eff.	-1.6	$+1.6$	-1.7	$+1.6$
Electron Iso. Eff.	-0.8	$+0.8$	-0.9	$+0.9$
Electron ID Eff.	-0.8	$+0.8$	-0.9	$+0.9$
Electron Energy Scale	-1.1	$+1.0$	-1.2	$+1.0$
W/Z + jets Ratio	$+0.2$	-0.3	$+0.1$	-0.4
σ(Single Top)	$+0.8$	-0.8	incl. in stat.	incl. in stat.
Luminosity	-3.3	$+3.5$	-3.6	$+3.7$
Statistical Uncertainty	$+3.7$	-3.7	$+3.8$	-3.8

8. Simultaneous Measurement of the Top Quark Pair Production Cross Section in the Combined Electron/Muon and Jets Channels

This thesis focuses on the measurement of the top quark pair production cross section in the electron + jets channel. A similar measurement of the production cross section was performed in the muon + jets channel [3].

It is desirable to perform a combined measurement of the top quark pair production cross section using both channels, the electron + jets channel and the muon + jets channel. This can help to reduce statistical uncertainties due to the combined amount of data selected. Furthermore, systematic uncertainties can be reduced due to the additional information that becomes available in the combined consideration of the two channels.

8.1. Summary of the Measurement of the Top Quark Pair Production Cross Section in the Semi-Leptonic Muon + Jets Channel

Reference [3] presents the measurement of the top quark pair production cross section in the semi-leptonic muon + jets decay channel. The cross section is measured using a binned likelihood fit following the same definition in terms of the likelihood function as the one presented in this thesis in Equation 6.1. The templates are extracted in the same way as in this thesis, using simulation based shapes for the top quark pair contribution, the single top quark contribution and the W/Z + jets contribution. The multijet shape is extracted from collision data using the anti-relative isolation method. Details on the event selection can be found in Reference [3]. The measurement in Reference [3] yields a cross section of

$$\sigma_{t\bar{t}} = 157.4 \pm 3.8\,(\text{stat.})\,^{+10.6}_{-12.4}\,(\text{syst.}) \pm 3.5\,(\text{lumi.})\,\text{pb}.$$

This result is compatible with the measurement performed in this thesis which yields

$$\sigma_{t\bar{t}} = 155.5 \pm 3.7\,(\text{stat.})\,^{+12.6}_{-14.8}\,(\text{syst.}) \pm 3.4\,(\text{lumi.})\,\text{pb}.$$

Hence, a combination of the two results can be performed.

8.2. Measurement of the Top Quark Pair Production Cross Section in the Semi-Leptonic Electron/Muon and Jets Channels

The individual steps to measure the cross section in the combined electron/muon + jets channels are discussed in this section. First, the template shapes are explained, then the stability of the simultaneous template fit is confirmed. The cross section measurement is performed on collision data and the systematic uncertainties are evaluated.

8.2.1. Template Shapes for the Simultaneous Fit

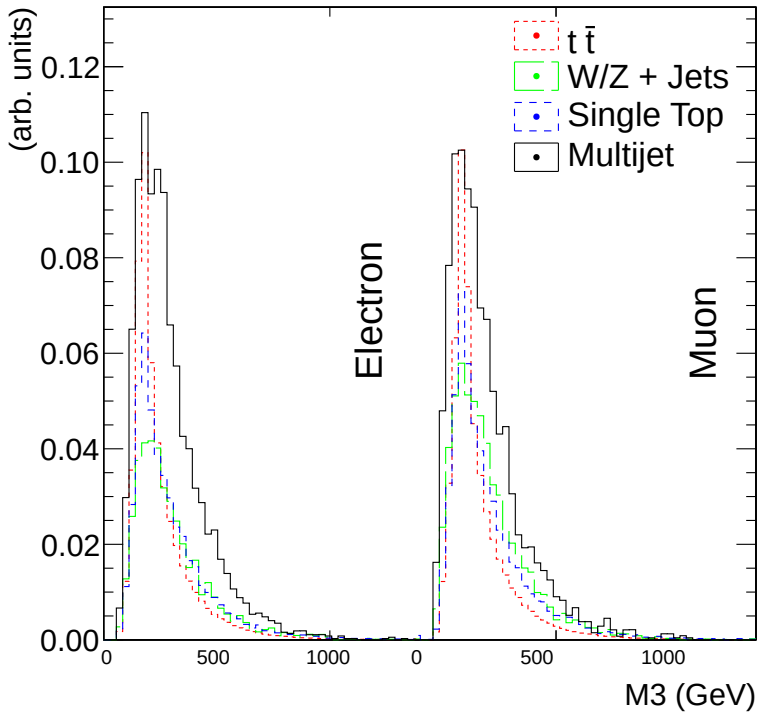


Figure 8.1.: Template shapes for top quark pair events, single top quark events, W/Z + jets events and multijet events for the combined electron/muon + jets measurement.

The measurements are combined by performing a simultaneous template fit to the combined M3 mass distribution in the semi-leptonic electron + jets and muon + jets decay channel. The likelihood function is given by Equation 6.1. The two

M3 mass distributions from the two decay channels are combined to one single data distribution. Also the individual template shapes are combined. This is done by letting the two distributions share the same x-axis. The distribution of the M3 mass shape for the muon + jets channel has been shifted along the x-axis by $1400\,\mathrm{GeV}/c^2$. This is to prevent any overlap of the two distributions. The resulting template shapes are shown in Figure 8.1. The individual shape contributions are identical to the ones in the pure electron + jets decay channel as described in Section 6.2 and in the pure muon + jets decay channel as described in Reference [3]. The number of parameters in the fit is increased by one compared to the fit in an individual channel. The parameters for the number of top quark pair events, the number of W/Z + jets events and the number of single top quark events are the same as in the fit in the individual channel. The parameters for these processes are directly correlated through the ratio of the selection efficiencies between the electron + jets channel and the muon + jets channel, as they are used to extract the same cross section values. The templates are constructed by normalising the template shapes from the electron and muon channels individually and weighting them by the selection efficiency for the given channel. Then the two shapes are combined and normalised to unit area. The number of multijet events with an electron and the number of multijet events with a muon are considered independently as the sources of those background contributions are different and do not necessarily yield the same cross section. Multijet events with electrons are dominated by jets with large electromagnetic fractions and decay-in-flight leptons. Multijet events with muons are typically due to decay-in-flight of hadrons and so-called *punch-through* muons, where particles are identified in the muon chamber due to late hadronic cascades in the material between the calorimeter and the muon chambers. Resulting hadron leakage into the muon chambers can be misidentified as muons. In summary, the template fit in the combined channel features five fit parameters, $N_{t\bar{t}}$, $N_{W/Z\,+\,\mathrm{jets}}$, N_{st}, which are the sum of the events in the electron and muon channel, $N_{\mathrm{multijet\,e}}$ and $N_{\mathrm{multijet}\,\mu}$, in contrast to the four parameters $N_{t\bar{t}}$, $N_{W/Z\,+\,\mathrm{jets}}$, N_{st}, N_{multijet} in the individual channels, taking the selection efficiency ratios for $\epsilon_{t\bar{t}}$, $\epsilon_{W/Z\,+\,\mathrm{jets}}$ and ϵ_{st} from simulation.

8.2.2. Stability of the Template Fit

The stability of the fit is studied on 5000 pseudo-experiments as in the case of the template fit in an individual channel as described in Section 6.3.1. The *pull* distribution is derived for the top quark pair fit parameter and shown in Figure 8.2. It is centred around zero with a width of $\sigma = 1$ which means that the fit is unbiased and the uncertainties are well estimated.

Table 8.1 summarises the result of the template fit on pseudo-experiments. The measured cross sections agree within the uncertainties with the cross sections that were used to generate the pseudo-experiments.

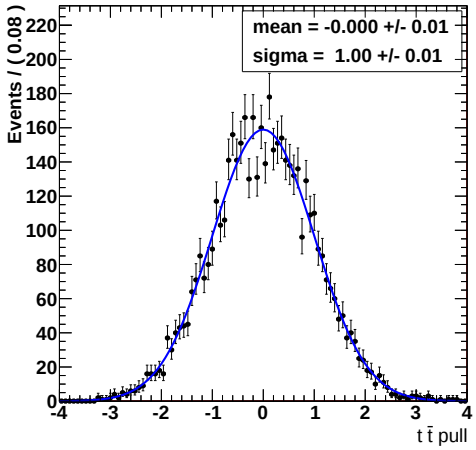


Figure 8.2.: Pull distributions for top quark pair template fit parameter in combined electron/muon + jets channels. The pull is centred around zero with a width of $\sigma = 1$ within the uncertainties. This means the fit results are unbiased and the uncertainties are well estimated.

Table 8.1.: Cross sections measured on 5000 pseudo-experiments in the combined electron/muon + jets channels. The cross section has been calculated from the absolute number of events of each contribution using Equation 6.1.

Contribution	σ_{Expected} (pb)	σ_{Observed} (pb)
$t\bar{t}$	165	165 ± 2.9
W/Z + jets	34362	34270 ± 1900
single top	85	86 ± 24
multijet (μ)	345	389 ± 790
multijet (e)	1886	1925 ± 610

8.2.3. Result of the Measurement in the Combined Electron/Muon and Jets Channels

The top quark pair production cross section is measured in the combined electron/muon + jets channels. The result of the fit to collision data is presented in Figure 8.3.

Table 8.2 shows the Standard Model expectation and the observed number of events from the fit for each contribution and the correlation between the different parameters of the template fit. The correlation between the top quark pair contribution and all remaining parameters is reasonably low. The highest correlation can be observed with respect to the multijet background process featuring electrons with a value of -0.38. Larger correlations on the order of 0.8 are obtained between the W/Z + jets parameter and the two multijet contributions. This does not affect the measurement of the top quark pair production cross section because all three contributions have reasonably low correlations with respect to the top quark pair fit parameter. The sum of the expected background contributions is 28485 which is in good agreement with the observed number of 27411 ± 1750.

Table 8.3 lists the results of the cross section measurement in the combined channel. The measurement of the top quark pair production cross section yields

$$\sigma_{t\bar{t}} = 155.9 \pm 2.7\,\text{pb},$$

where the uncertainty is statistical only. Like in the measurements in the electron + jets channel and in [3] the cross section for the multijet contribution is overestimated and the cross section for the W/Z + jets process is underestimated compared to the Standard Model prediction.

8.3. Systematic Uncertainties on the Measurement of the Top Quark Pair Production Cross Section in the Combined Electron/Muon and Jets Channels

The systematic uncertainties on the combined measurement are extracted following the description in Chapter 7. The systematically varied top quark pair selection efficiencies and template shapes are not explicitly displayed for the combined fit because they are merely the combination of the figures shown in Chapter 7 and the ones for the muon + jets channel described in detail in Reference [3]. The summary of all systematic uncertainties will be given at the end of this section in Table 8.5. Some comments on the derivation of the systematic uncertainties follow for selected uncertainties where the uncertainty is not a trivial result of the application of the procedure of Chapter 7.

Table 8.2.: Expected and observed number of events for different contributions and correlations between individual fit parameters for the fit to 2011 collision data in the combined electron/muon + jets channels. Only one half of the table is filled due to the symmetry of the correlation factors.

Contribution	$N_{\text{exp.}}$	$N_{\text{obs.}}$	$t\bar{t}$	W/Z + jets	single top	multijet (μ)	multijet (e)
$t\bar{t}$	38887	36749 ± 650	1	0.05	−0.33	−0.30	−0.38
W/Z + jets	24077	15790 ± 1300		1	−0.30	−0.87	−0.80
single top	2176	2537 ± 600			1	0.02	−0.03
multijet (μ)	345	4232 ± 800				1	0.93
multijet (e)	1887	4852 ± 600					1

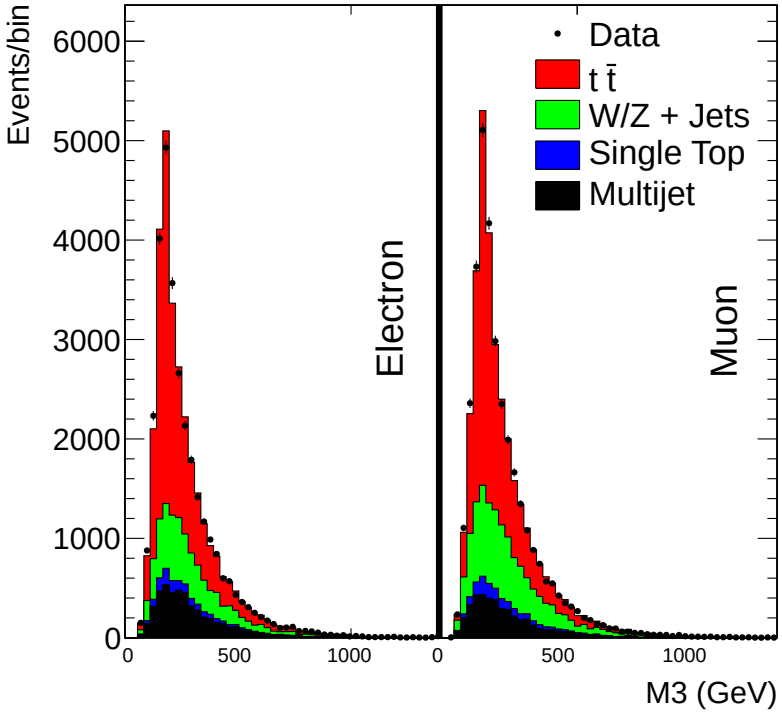


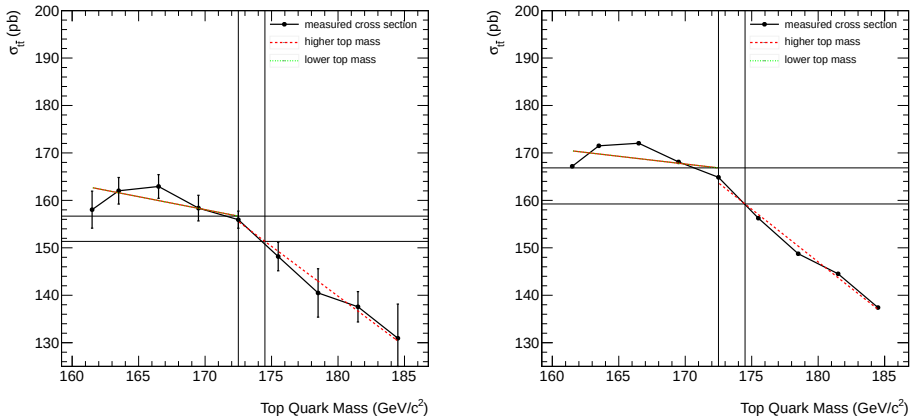
Figure 8.3.: Result of the template fit to the M3 distribution in the 2011 collision data in the combined electron/muon + jets channels. A good agreement between the result of the template fit and the collision data shape is reached.

Table 8.3.: Cross sections measured on 2011 collision data in the combined electron/muon + jets channels. The cross section has been calculated from the absolute number of events of each contribution using Equation 6.1.

Contribution	σ_{Expected} (pb)	σ_{Observed} (pb)
$t\bar{t}$	165	155.9 ± 2.7
W/Z + jets	34362	22541 ± 1900
single top	84.9	99.0 ± 23
multijet (μ)	344.9	4232.3 ± 800
multijet (e)	1886.5	4852.1 ± 600

8.3.1. Top Quark Mass

The systematic uncertainty on the measurement of the top quark production cross section due to the uncertainty on the knowledge of the top quark mass follows the description in Section 7.2.6. For reference, Figure 8.4 shows the extrapolation of the measurement of the cross section as a function of the top quark mass. Figure 8.5 shows the decoupled effect of the influence of the top quark pair template shape change and of the top quark pair selection efficiency change. The behaviour follows the one described in Section 7.2.6 with the same interpretation of the results.



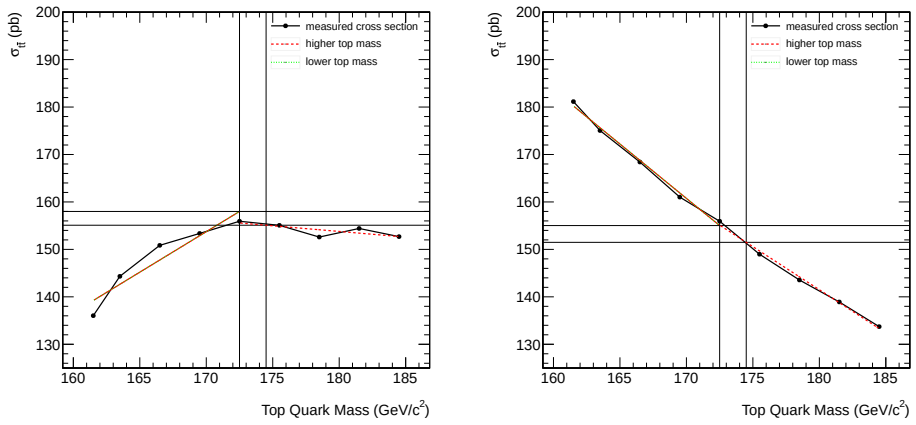
(a) Change of selection efficiency and shape for collision data.

(b) Change of selection efficiency and shape for pseudo-experiments.

Figure 8.4.: Systematic uncertainty on cross section measurement in combined electron/muon + jets channels due to the top quark mass uncertainty on collision data and on pseudo-experiments. Figure 8.4(a) and Figure 8.4(b) show the cross section measurement after systematically changing the selection efficiency and the M3 mass shape.

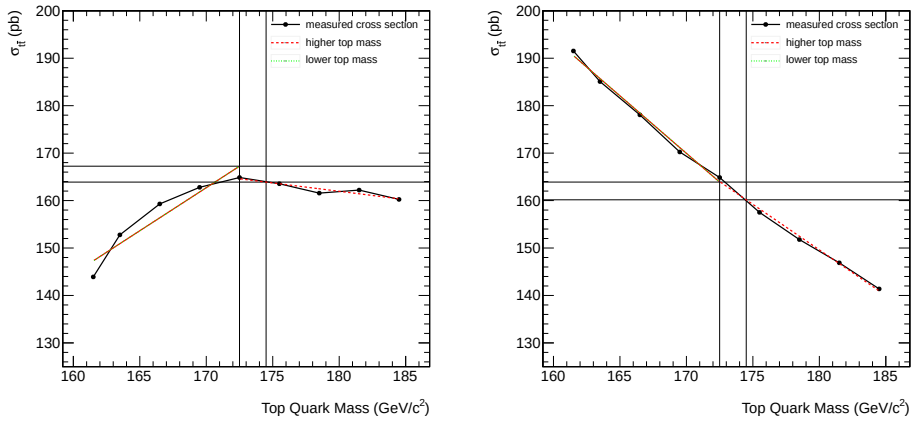
8.3.2. Multijet Background Contribution Estimation

The multijet shapes in the electron + jets channel and in the muon + jets channel are treated separately as motivated in Section 8.2.1. For this reason, the systematic uncertainties on the electron multijet shape and on the muon multijet shape for the isolation dependence are evaluated separately in the simultaneous fit. The uncertainty on the contamination removal is evaluated in a correlated way as described later in this section. The resulting uncertainties on the isolation dependence are added in quadrature. First, the fit is repeated keeping the muon multijet shape constant and varying the electron multijet shape for isolation stability uncertainty. Then, the electron shape is kept constant in its nominal form and the muon multijet shape is varied according to its isolation stability uncertainty. The results are listed



(a) Change of shape with constant selection efficiency for collision data.

(b) Change of selection efficiency with constant mass shape for collision data.



(c) Change of shape with constant selection efficiency for pseudo-experiments.

(d) Change of selection efficiency with constant mass shape for pseudo-experiments.

Figure 8.5.: Systematic uncertainty on cross section measurement in the combined electron/muon + jets channels due to the top quark mass on collision data and on pseudo-experiments. The effect of the shape influence and of the selection efficiency influence are decoupled. Figure 8.5(a) and Figure 8.5(c) show the cross section measurement after systematically changing the the M3 mass shape and keeping the top quark pair selection efficiency constant. Figure 8.5(b) and Figure 8.5(d) show the cross section measurement after systematically changing the selection efficiency and using the nominal shape of the M3 mass.

in Table 8.5. For reference, Figure 8.6 shows the result of the cross section measurement in different slices of relative isolation for the electron multijet background contribution and the muon multijet background contribution. In both cases, the variation is shown on collision data and pseudo-experiments and an extrapolation towards the signal region is performed with a straight line fit.

For the contamination removal uncertainty the contamination removal is increased first by increasing the top quark pair cross section by 100% in the electron and muon channel in parallel and keeping the W/Z + jets cross section at its nominal value. Then the top quark pair cross section is kept at its nominal value and the W/Z + jets cross section is varied up and down by 50% in both channels in parallel. Note that in the measurement in the muon + jets channel presented in [3], an uncertainty of 50% on the top quark pair cross section is assumed.

8.3.3. Summary of Systematic Uncertainties

A summary of all systematic uncertainties on the top quark pair production cross section measurement in the combined electron/muon + jets channels is given in Table 8.5. The table lists the uncertainties for the measurement on collision data and in the pseudo-experiments.

As in the cross section measurement in the electron + jets decay channel, the dominant systematic uncertainties are the uncertainty due to the jet energy scale, jet-parton matching, factorisation scale and parton distribution functions.

8.4. Conclusions on the Measurement of the Top Quark Pair Production Cross Section in the Combined Electron/Muon and Jets Channels

The expected sensitivity for the combined measurement based on pseudo-experiments is

$$\sigma_{t\bar{t}} = 165.0 \pm 2.9\,(\text{stat.})\,^{+11.5}_{-15.0}\,(\text{syst.}) \pm 3.6\,(\text{lumi.})\,\text{pb}.$$

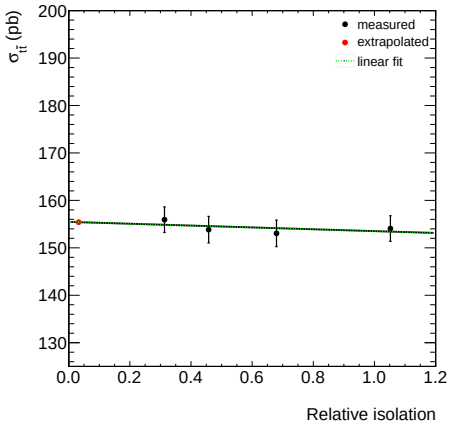
In relative uncertainties the result of the measurement on pseudo-experiments can be expressed as

$$\sigma_{t\bar{t}} = 165.0\,\text{pb} \pm 1.8\,\%\,(\text{stat.})\,^{+7.0}_{-9.1}\,\%\,(\text{syst.}) \pm 2.2\,\%\,(\text{lumi.}).$$

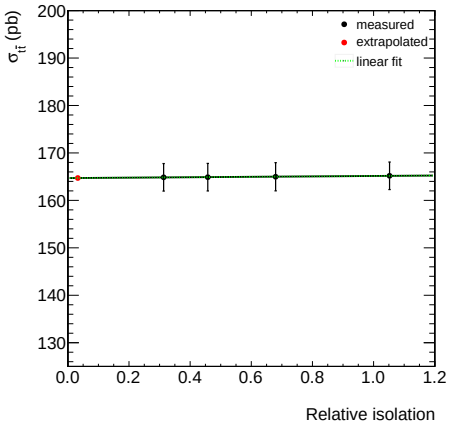
The measurement in the combined electron/muon + jets channels on collision data yields

$$\sigma_{t\bar{t}} = 155.9 \pm 2.7\,(\text{stat.})\,^{+11.2}_{-13.8}\,(\text{syst.}) \pm 3.4\,(\text{lumi.})\,\text{pb}.$$

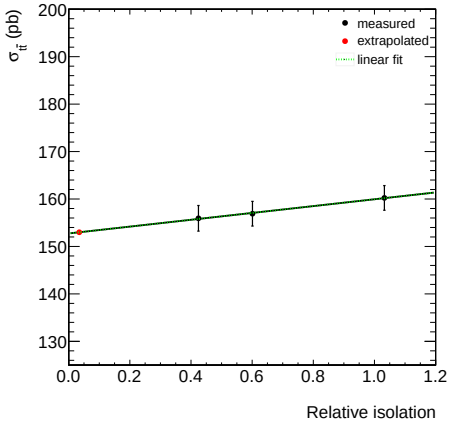
In relative uncertainties the result of the measurement on collision data can be



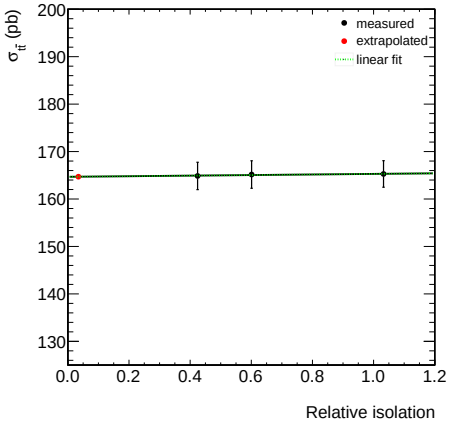
(a) Variation of electron template for measurement on collision data.



(b) Variation of electron template for measurement on pseudo-experiments.



(c) Variation of muon template for measurement on collision data.



(d) Variation of muon template for measurement on pseudo-experiments.

Figure 8.6.: Cross section measurement with extracted multijet shapes from different relative Particle Flow isolation slices. An extrapolation to the signal region is performed using a linear function fit.

Table 8.4.: Comparison of systematic change by template shape change or top quark selection efficiency change in the combined electron/muon + jets decay channels on collision data.

Source	Uncertainty (pb)						Dominant Effect
	Overall		Shape		Efficiency		
	Up	Down	Up	Down	Up	Down	
Jet Energy Resolution	+0.8	−0.6	+1.4	−2.0	−0.5	+0.4	Shape
Jet Energy Scale	−6.9	+4.5	−0.8	−2.9	−6.1	+7.5	Efficiency
$t\bar{t}$ Matching	−0.8	−4.2	−1.6	+2.2	+0.8	−6.3	Efficiency
$t\bar{t}$ Fact. Scale	−1.9	+4.6	−14.2	+17.3	+11.2	−10.6	Balanced
Top Quark Mass	−4.6	0.0	−0.8	0.0	−4.5	0.0	Efficiency
Parton Distribution Functions	+7.4	−6.0	+8.4	−7.0	+1.7	−1.7	Shape
Pile-Up	0.0	0.0	0.0	0.0	0.0	0.0	Efficiency

Table 8.5.: Summary of the systematic uncertainties on the measurement of the top quark pair production cross section in the combined electron/muon + jets decay channels.

Source	Uncertainty (pb)			
	Data		Pseudo-Experiment	
	Up	Down	Up	Down
Jet Energy Resolution	+0.8	−0.6	+0.1	0.0
Jet Energy Scale	−6.9	+4.5	−7.8	+4.7
$t\bar{t}$ Matching	-0.8 ± 1.5	-4.2 ± 1.8	−0.9	−3.2
W/Z + jets Matching	-2.9 ± 1.5	-3.5 ± 1.4	−4.4	−4.9
$t\bar{t}$ Fact. Scale	-1.9 ± 1.6	$+4.6 \pm 1.5$	−1.9	+4.8
W/Z + jets Fact. Scale	-4.9 ± 1.3	$+2.8 \pm 1.1$	−6.7	+1.8
Multijet Stability (e)	0.0	−0.5	0.0	−0.3
Multijet Stability (μ)	0.0	−2.9	0.0	−0.3
Multijet Cont. Rem. ($t\bar{t}$)	+3.4	−3.2	+0.6	−0.4
Multijet Cont. Rem. (W/Z + jets)	−0.2	−1.0	−0.1	−0.2
Stat. Shape Unc.	+1.8	−1.8	n/a	n/a
Top Quark Mass	−4.6	0.0	−5.6	0.0
Parton Distribution Functions	+7.4	−6.0	+8.9	−6.8
Pile-Up	0.0	0.0	+0.1	0.0
Trigger (e part)	−0.8	+0.8	−0.8	+0.9
Trigger (jet part)	−0.8	+0.8	−0.8	+0.9
Conv. Rej. Eff.	−0.8	+0.8	−0.8	+0.9
Electron Iso. Eff.	−0.4	+0.4	−0.4	+0.4
Electron ID Eff.	−0.4	+0.4	−0.4	+0.4
Electron Energy Scale	−0.5	+0.6	−0.5	+0.6
Muon Trigger Eff.	−0.4	+0.4	−0.4	+0.5
Muon ID Eff.	−0.8	+0.8	−0.8	+0.9
W/Z + jets Ratio	+0.2	−0.3	+0.3	−0.4
σ(Single Top)	+1.2	−1.2	incl. in stat.	incl. in stat.
Luminosity	−3.4	+3.5	−3.6	+3.7
Statistical Uncertainty	+2.7	−2.7	+2.9	−2.9

expressed as

$$\sigma_{t\bar{t}} = 155.9\,\text{pb} \pm 1.7\,\%\,(\text{stat.})\,^{+7.2}_{-8.9}\,\%\,(\text{syst.}) \pm 2.2\,\%\,(\text{lumi.}).$$

The combination of the electron/muon + jets channels leads to an improvement in the precision of the measurement. The statistical uncertainty reduces as well as the systematic uncertainty compared to the pure electron + jets channel. Yet, the precision of the measurement in the combined electron/muon + jets channels does not yield a better precision than the one of the pure muon + jets channel with a result of

$$157.4 \pm 3.8(\text{stat.})^{+10.6}_{-12.4}(\text{syst.}) \pm 3.5(\text{lumi.})\,\text{pb}\,[3].$$

Although the statistical uncertainty is reduced by the combined measurement due to the combined amount of data from the electron + jets and muon + jets channels, the systematic uncertainty of the combined measurement does not reach the one from the measurement in the muon + jets channel alone. The applied method is an approximation for the combination of the two results. The systematic uncertainties are re-evaluated individually using the combined template shape of both channels. The higher precision in terms of the overall systematic uncertainty in the muon + jets channel does not result in a larger weight of the muon + jets result in the combination. This would be the case in an alternative approach for the combination with the BLUE method [137]. The BLUE method requires symmetric uncertainties of the measurements. A symmetrisation of the uncertainties of the measurement and of the measurement presented in Reference [3] would lead to a significant increase of the systematic uncertainties and would make it difficult to compare the combined result with the initial measurements. Therefore, the approach of the approximate combination is chosen as described in this thesis.

The result in the muon + jets channel reaches a higher precision than the measurement presented in this thesis mainly because a misidentification of muons is significantly less likely than that of electrons. This leads to a largely reduced background from multijet events. In particular, the theory uncertainties due to jet-parton matching and factorisation scale have a larger influence on the electron + jets channel than on the muon + jets channel due to the effect on the shape change in the M3 distribution in the two channels.

The systematic uncertainties of the separate measurements in the electron + jets channel and in the muon + jets channel are correlated to some extent. This includes the uncertainty on jet energy scale and resolution, parton-shower matching and factorisation scale, pile-up interactions, W/Z + jets cross section ratio, single top quark cross section and luminosity. Uncorrelated systematics are the multijet shape and stability due to the different multijet background compositions, the lepton and trigger efficiency measurements, and the statistical shape uncertainty on the templates. The correlations are correctly treated in the combination procedure.

The measurements in the electron + jets channel, the muon + jets channel and the combined electron/muon + jets channels agree within the statistical uncertain-

ties alone. The measurement in the combined electron/muon + jets channels is centred between the measurement in the electron + jets channel and the muon + jets channel. All measurements agree within the uncertainties with the standard model prediction at next-to-next-to-leading log of $165 \pm 10\,\mathrm{pb}$ [117].

9. Measurement of the Top Quark Mass from the Top Quark Production Cross Section

The top quark mass is one of the free parameters of the Standard Model of particle physics as described in Section 1.4. Direct measurements of the top quark mass are available [138, 139] but these depend strongly on the kinematic properties of top quark events. As an alternative approach, the top quark mass can be extracted from comparing the top quark production cross section to higher order QCD calculations. This approach is less sensitive to the kinematic properties of the event and provides information complementary to a direct measurement.

The measured cross sections from Chapter 6 and Chapter 8 are compared to next-to-next-to-leading-logarithm (NNLO) calculations parametrised as a function of the top quark mass provided by Kidonakis [117]. The top quark mass is defined as a pole mass. Uncertainties on the theory prediction are provided, including the uncertainty due to renormalisation and factorisation scales, uncertainty of the parton luminosity and uncertainty due to the variation of the strong coupling constant α_s in the parton distribution functions.

The theory prediction $f_{\text{th}}(\sigma_{t\bar{t}}|m_t)$ of the top quark production cross section is provided at discrete points as a function of the top quark mass. The top quark production cross section $f_{\text{exp}}(\sigma_{t\bar{t}}|m_t)$ is measured for nine discrete mass points as described in Section 7.2.6. Both sets of data points are interpolated by fitting a third order polynomial of the form

$$\sigma_{t\bar{t}} = \frac{1}{m_t^4}\left(a + b \cdot m_t + c \cdot m_t^2 + d \cdot m_t^3\right),$$

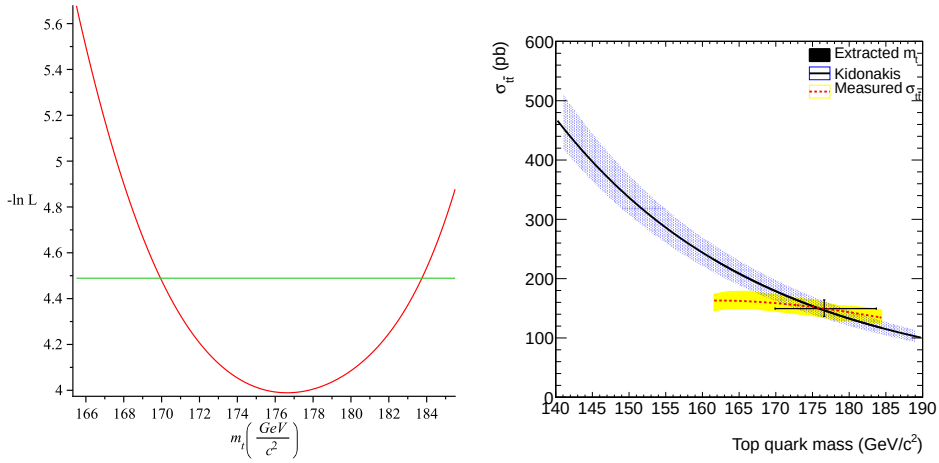
where m_t is the top quark mass. For both functions Gaussian uncertainties are assumed. In the case of asymmetric uncertainties they are symmetrised to match the larger of the two uncertainties. The uncertainty is the combined statistical and systematic uncertainty on the nominal cross section measurement, excluding the systematic uncertainty due to the top quark mass uncertainty. The uncertainty on the cross section measurement is assumed to be independent of the top quark mass assumption for the measurement.

A joint-likelihood approach is chosen to extract the top quark mass [140, 141,

142]. The joint-likelihood is defined as

$$L(m_t) = \int f_{\text{exp}}(\sigma_{t\bar{t}}|m_t) f_{\text{th}}(\sigma_{t\bar{t}}|m_t)\, \mathrm{d}\sigma_{t\bar{t}}.$$

It is evaluated by minimising the negative logarithm of the likelihood function. Figure 9.1(a) shows the negative logarithm of the likelihood function for the measurement in the electron + jets channel, Figure 9.2(a) for the combined electron/muon + jets channels. The uncertainty on the measurement is extracted by evaluating the change in top quark mass at the points where the likelihood increases by a value of 0.5 relative to its minimum.



(a) Negative logarithm of likelihood function. (b) Cross section as a function of top quark mass.

Figure 9.1.: Top quark production cross section as a function of the top quark mass and likelihood function for determination of top quark mass from cross section in the electron + jets channel.

The measurement of the top quark mass from the cross section in the electron + jets channel yields

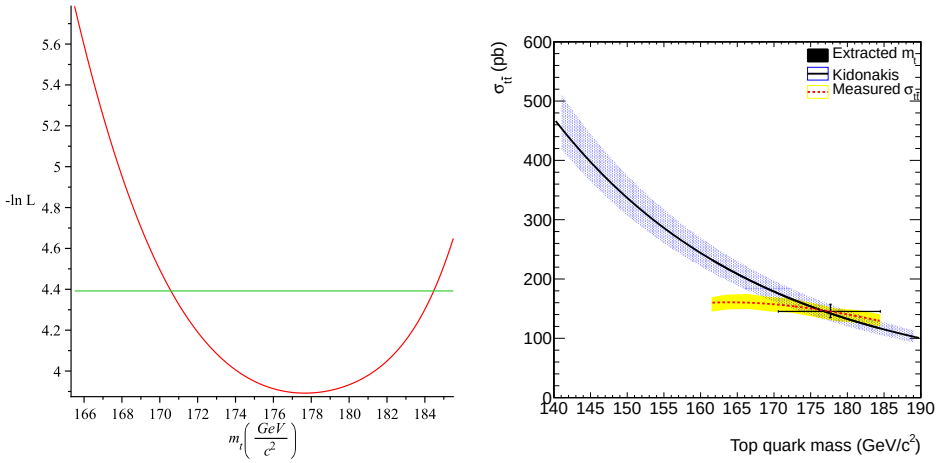
$$m_t = 176.6^{+7.1}_{-6.7}\,\text{GeV}/c^2.$$

The dependence of the top quark pair production cross section on the top quark mass is presented in Figure 9.1(b).

In the combined electron/muon + jets channel the measurement of the top quark mass from the cross section results in

$$m_t = 177.7^{+6.8}_{-7.1}\,\text{GeV}/c^2.$$

The dependence of the top quark pair production cross section on the top quark mass in the combined electron/muon + jets channel is shown in Figure 9.2(b). The uncertainty on the measurement in the combined channel yields a larger uncertainty



(a) Negative logarithm of likelihood function. (b) Cross section as a function of top quark mass.

Figure 9.2.: Top quark production cross section as a function of the top quark mass and likelihood function for determination of top quark mass from cross section in the combined electron/muon + jets channels.

than the measurement in the electron + jets channel because of the nonlinear behaviour of the measured cross section as a function of the top quark mass and the larger dependence in the combined electron/muon + jets channel.

Both measurements agree with each other and are compatible with the world average top quark mass

$$m_t = 173.5 \pm 0.6 \pm 0.8\,\mathrm{GeV}/c^2\ [2].$$

Part IV.

Conclusions

10. Comparison of the Obtained Results with Published Measurements and Theory Predictions

The measured results from Chapter 6 and Chapter 8 are compared to published results from the CMS and ATLAS collaborations and latest theory calculations.

Figure 10.1 gives an overview of the measurements in the semi-leptonic electron + jets channel and in the combined electron/muon + jets channels, compared to NNLO theory predictions [117, 143, 144] and the latest combined measurement on 2011 data from the CMS collaboration [145, 146] and the ATLAS collaboration [147, 148]. A more extensive list of cross section measurements is presented in Table 10.1.

The cross section combination presented by the CMS collaboration in [146] includes measurements in the single electron + jets and muon + jets channels, in the dilepton channels with two electrons, two muons, one electron and one muon, and one muon and one tau lepton, and in the fully hadronic channel. The combination of the ATLAS measurements [148] contains measurements from the single electron + jets and muon + jets channels, the dilepton channels with two electrons, two muons, and one electron and one muon and the fully hadronic channel. The LHC combination [149] combines all of the CMS and ATLAS measurements mentioned above.

Both measurements performed as part of this thesis are compatible with all theory predictions and the measurements performed by the CMS collaboration and the ATLAS collaboration.

Figure 10.2 shows the theory prediction of the top quark pair production cross section as a function of centre-of-mass energy comparing the results to measurements from experiments at the Tevatron [150, 151]. Again agreement of all results and the theory prediction is observed.

Table 10.1.: Comparison of cross section measurements from the CMS and ATLAS collaborations with results from this thesis and [3].

Source	Channel	Cross Section
Thesis	e + jets	$155.5 \pm 3.7\,(\text{stat.})\,^{+12.6}_{-14.8}\,(\text{syst.}) \pm 3.4\,(\text{lumi.})$ pb
Thesis	e/μ + jets	$155.9 \pm 2.7\,(\text{stat.})\,^{+11.2}_{-13.8}\,(\text{syst.}) \pm 3.4\,(\text{lumi.})$ pb
S. Walsh [3]	μ + jets	$157.4 \pm 3.8\,(\text{stat.})\,^{+10.6}_{-12.4}\,(\text{syst.}) \pm 3.5\,(\text{lumi.})$ pb
CMS [145]	e + jets	$163.0 \pm 4.4\,(\text{stat.}) \pm 12.7\,(\text{syst.}) \pm 7.3\,(\text{lumi.})$ pb
CMS [145]	e/μ + jets	$164.4 \pm 2.8\,(\text{stat.}) \pm 11.9\,(\text{syst.}) \pm 7.4\,(\text{lumi.})$ pb
CMS [146]	Combination	$165.8 \pm 2.2\,(\text{stat.}) \pm 10.6\,(\text{syst.}) \pm 7.8\,(\text{lumi.})$ pb
ATLAS [147]	e/μ + jets	$179 \pm 9.8\,(\text{stat.} + \text{syst.}) \pm 6.6\,(\text{lumi.})$ pb
ATLAS [148]	Combination	$177 \pm 3\,(\text{stat.})\,^{+8}_{-7}\,(\text{syst.}) \pm 7\,(\text{lumi.})$ pb
ATLAS+CMS [149]	Combination	$173.3 \pm 2.3\,(\text{stat.}) \pm 7.6\,(\text{syst.}) \pm 6.3\,(\text{lumi.})$ pb

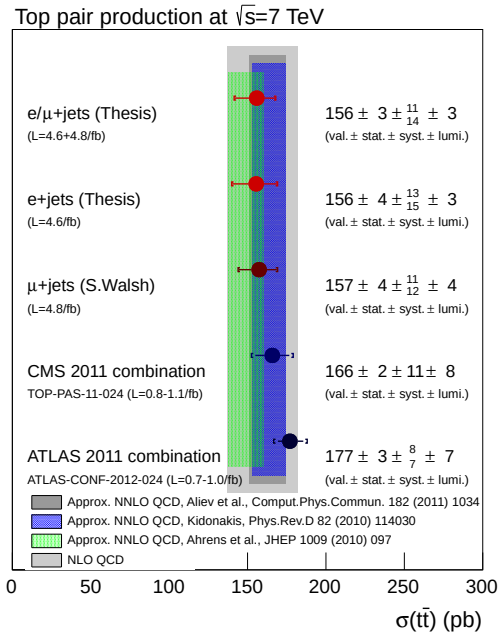


Figure 10.1.: Measured cross sections as part of this thesis compared to measured cross section from [3] and latest published results from the CMS and ATLAS collaborations.

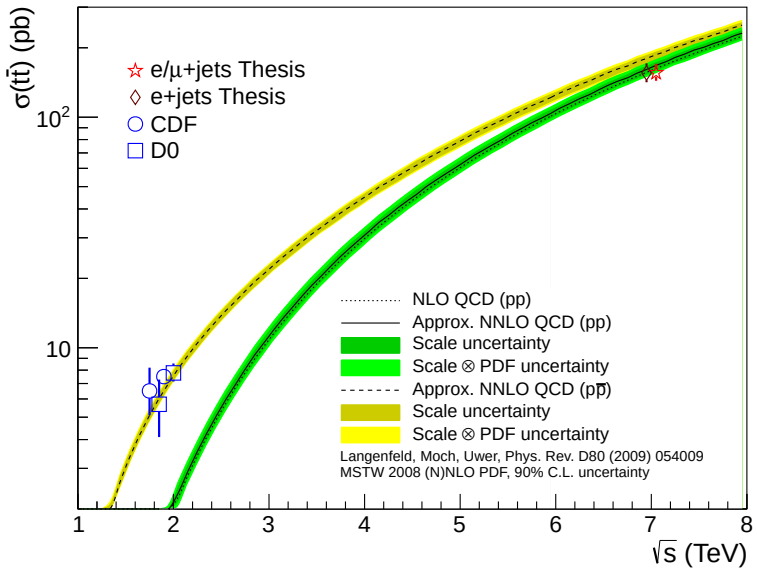


Figure 10.2.: Cross sections measured as part of this thesis compared to cross sections measured at Tevatron energies and theory prediction of the cross section as a function of the centre-of-mass energy for proton-proton and proton-anti-proton collisions. Multiple measurements at the same centre-of-mass energy have been slightly displaced for the sake of visibility.

11. Conclusion

The top quark is the heaviest fundamental particle observed. At LHC energies the top quark production cross section is increased by more than a factor of twenty compared to former particle physics experiments. The heavy mass and short lifetime of the top quark give it a special role in high energy physics. The mass of the top quark is one of the free parameters of the Standard Model. Moreover, the top quark is the only quark that can be studied directly by means of its decay products because its lifetime is shorter than the hadronisation time. This allows for precision measurements such as the measurement of the top quark pair production cross section in the electron + jets and combined electron/muon + jets channels performed in this thesis. The measurement of the top quark pair production cross section is a precision test to challenge the predictions of the Standard Model of particle physics. Moreover, a precise knowledge of the top quark pair production is crucial for other analyses where top quark pair production is one of the dominant background processes. This is valid for the measurement of the CKM matrix element $|V_{\mathrm{tb}}|$ in single top quark production and many searches for physics beyond the Standard Model.

The measurement of the top quark pair production cross section has been performed in the electron + jets channel with the CMS detector at the LHC at a centre-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$ in the full 2011 dataset.

One of the crucial parts of a cross section measurement in a decay channel featuring a lepton is the precise knowledge of the lepton selection and trigger efficiencies. Extensive data-driven studies on the behaviour of those efficiencies have been performed as a function of various kinematic variables. Scale factors were derived to correct the behaviour observed in the simulation to that measured on collision data. A two-dimensional parametrisation of the trigger efficiency has been derived and multiple scale factors for the different parts of the electron identification procedure have been calculated and applied.

The dominant background processes to top quark pair events in the semi-leptonic electron channel are W/Z + jets events, single top quark events and multijet production. A data-driven model based on measurements in side band regions has been developed and extensively validated to extract the contribution of multijet production after the nominal event selection. The stability of the model has been demonstrated and corrections for prompt-lepton contamination have been applied.

The top quark pair production cross section in the electron + jets channel has been measured using a binned maximum likelihood template fit approach as

$$\sigma_{t\bar{t}} = 155.5 \pm 3.7\,(\mathrm{stat.})\,^{+12.6}_{-14.8}\,(\mathrm{syst.}) \pm 3.4\,(\mathrm{lumi.})\,\mathrm{pb}.$$

The stability and performance of the method has been assessed in detail on pseudo-experiments. Using information from an external muon + jets analysis, the template fit approach has been extended to perform a simultaneous fit in the electron + jets and muon + jets decay channels. This leads to a reduction of the uncertainties compared to the measurement in the electron + jets channel only with the resulting combined cross section

$$\sigma_{t\bar{t}} = 155.9 \pm 2.7\,(\text{stat.})\,^{+11.2}_{-13.8}\,(\text{syst.}) \pm 3.4\,(\text{lumi.})\,\text{pb}.$$

The dominant systematic uncertainties in both measurements are the uncertainty due to the parton distribution functions, the uncertainty due to the jet energy scale, the uncertainty due to jet-parton matching and factorisation scale and the uncertainty on the luminosity measurement.

The results have been compared to various theory predictions and other measurements of the top quark pair production cross section. Agreement between the results has been observed in all cases.

The top quark mass has been estimated based on the Standard Model prediction of the top quark pair production cross section and its functional dependence on the top quark mass. This estimation has been performed in the electron + jets decay channel and in the combined electron/muon + jets decay channels. Both measurements are compatible with each other and with direct measurements of the top quark mass.

In the future, even more precise measurements of the top quark pair production cross section will become possible due to the larger amount of recorded collision data, better understanding of the detector and high precision tuning of simulated events. Measurements should also compare the predictions of alternative Monte Carlo generators like MC@NLO [76] or alternative parton shower models like the one employed in Herwig [152, 153]. Dominant systematic uncertainties as the uncertainty due to the jet energy scale will reduce over time with increasing understanding of the jet reconstruction in the CMS detector. The uncertainty on the measurement of the top quark pair production cross section is dominated by the parton distribution function uncertainty and should be reduced with more precise predictions on the parton distribution functions. These can be derived from measurements of W and Z boson production at the LHC. For future measurements, also additional parton distribution function libraries should be compared for the estimation of systematic uncertainties, such as MSTW [154], NNPDF [155] or HERAPDF [156].

Challenging the Standard Model with measurements of the top quark pair production at different centre-of-mass energies will become possible in the near future within the near-term upgrade plans of the Large Hadron Collider. First measurements of the top quark pair production cross section are already available at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ [157]. Within the next years a measurement at a centre-of-mass energy of up to $\sqrt{s} = 14\,\text{TeV}$ will become possible. One of the key points for future measurements in the electron + jets decay channel will

be the development of new models for the estimation of the multijet production. Once precise predictions on the multijet production become available, multivariate techniques can be employed to maximise the separation of the individual physics processes after the final selection and to select input variables to reduce systematic uncertainties. This will help to minimise the overall uncertainties on the final measurement.

Part V.

Appendix

A. Technical Setup for the Analysis

A.1. Calibration Constants and Software Versions

For this thesis the CMS software framework CMSSW is used in version 4.2.8 patch 7. The calibration constants, including jet energy calibrations, are set according to the *global tag* `START42_V17` in simulation and according to `GR_R_42_V25` on collision data.

A.2. Datasets with Simulated Events

A.2.1. Primary Datasets with Collision Data

Throughout the 2011 data taking period the primary dataset name changed several times. A part of the data has been re-reconstructed to compensate for various problems that occured during the data taking and to profit from latest offline reconstruction algorithms. For this reason, events from different run ranges are taken from different datasets. Table A.1 lists the run ranges and corresponding *Data Aggregation Service* names. Events that are used for the nominal analysis are taken from the *ElectronHad* datasets. Events used for the measurement of the trigger and lepton selection efficiency are taken from the *SingleElectron* dataset, as this dataset is most suitable for the *Tag & Probe* method. Muon events are selected from the *SingleMuon* dataset for the combined measurement in the electron + jets and muon + jets channels. Events that are selected for storage by a muon trigger which requires only one muon in the event are written to the *SingleMuon* dataset. Events with only one online electron requirement are written to the *SingleElectron* dataset. Top quark physics trigger algorithms with the requirement of one online electron in addition to several online jets are written to the *ElectronHad* dataset.

Data filters are employed to ensure that only high quality collision data is used in the analysis. The CMS collaboration provides quality definition files in the *JSON* file format [158]. A list of the corresponding files that have been used is given in Table A.2.

A.2.2. Description of Simulated Datasets

Table A.3 summarises the CMS internal *Data Aggregation Service* dataset names for the various samples of simulated events used throughout this thesis. Top quark pair events, W + jets and Z + jets events are simulated at leading-order of perturbation theory using the MadGraph event generator. Single top quark events

Table A.1.: List of primary datasets of collision data, re-reconstructed datasets and corresponding run ranges. Events that are used for the nominal analysis are taken from the *ElectronHad* datasets. Events used for the measurement of the trigger and lepton selection efficiency are taken from the *SingleElectron* dataset.

Run Range	DAS name
160329-163869	/ElectronHad/Run2011A-May10ReReco-v1/AOD
165071-168437	/ElectronHad/Run2011A-PromptReco-v4/AOD
170053-172619	/ElectronHad/Run2011A-05Aug2011-v1/AOD
172620-173688	/ElectronHad/Run2011A-PromptReco-v6/AOD
175973-180252	/ElectronHad/Run2011B-PromptReco-v1/AOD
160329-163869	/SingleElectron/Run2011A-May10ReReco-v1/AOD
165071-168437	/SingleElectron/Run2011A-PromptReco-v4/AOD
170053-172619	/SingleElectron/Run2011A-05Aug2011-v1/AOD
172620-173688	/SingleElectron/Run2011A-PromptReco-v6/AOD
175973-180252	/SingleElectron/Run2011B-PromptReco-v1/AOD
160329-163869	/SingleMuon/Run2011A-May10ReReco-v1/AOD
165071-168437	/SingleMuon/Run2011A-PromptReco-v4/AOD
170053-172619	/SingleMuon/Run2011A-05Aug2011-v1/AOD
172620-173688	/SingleMuon/Run2011A-PromptReco-v6/AOD
175973-180252	/SingleMuon/Run2011B-PromptReco-v1/AOD

Table A.2.: CMS collaboration internal quality definition files for collision data in corresponding datasets. The identifier *ReReco* with a given date refers to a re-reconstruction that has been performed in 2011 on a part of the data. *PromtReco* refers to data that has not been reprocessed. The corresponding dataset definitions are listed in Table A.1.

Dataset	Data Quality Definition File
May10ReReco	Cert_160404-163869_7TeV_May10ReReco_Collisions11_JSON_v3
Aug05ReReco	Cert_170249-172619_7TeV_ReReco5Aug_Collisions11_JSON_v3
PromtReco	Cert_160404-180252_7TeV_PromptReco_Collisions11_JSON

are simulated at next-to-leading-order precision with the POWHEG event generator [159, 160]. For the simulation of multijet events the PYTHIA event generator has been employed using the *Z2* tune [161]. In all cases PYTHIA is used for the generation of the parton shower. The parton distribution functions used at production stage are based on the CTEQ 6.*l* library.

The top quark pair event samples, W + jets events and Z + jets events have been re-produced with varied values according to the theoretical uncertainties for the Q^2 parton-jet matching threshold as listed in Table A.4 and with varied factorisation scale threshold as listed in Table A.5.

Top quark pair samples have been reproduced with varying values for the top quark mass to study the influence of the top quark mass uncertainty on the cross section measurement. The samples with different top quark masses are listed in Table A.6. In the nominal top quark pair sample a top quark mass of 172.5 GeV is assumed.

Table A.3.: List of nominal simulated datasets used for the cross section measurement. t-Channel, s-Channel and tW-Channel denote the given single top quark production mode. Mulijet, e/m enriched denotes multijet samples with the requirement of a large electromagnetic fraction in the event. Multijet, $b/c \rightarrow e$ contains events with b or c flavoured hadrons that decay into an electron. Both types of multijet samples are produced in different regions of momentum transfer in the hard interaction $\hat{p}_T$.

Event Content	DAS name
$t\bar{t}$	/TTJets_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v2
W + jets	/WJetsToLNu_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1
Z + jets	/DYJetsToLL_TuneZ2_M-50_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1
t-Channel, t	/T_TuneZ2_t-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1
t-Channel, $\bar{t}$	/Tbar_TuneZ2_t-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1
s-Channel, $\bar{t}$	/T_TuneZ2_s-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1
s-Channel, $\bar{t}$	/Tbar_TuneZ2_s-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1
tW-Channel, t	/T_TuneZ2_tW-channel-DR_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1
tW-Channel, $\bar{t}$	/Tbar_TuneZ2_tW-channel-DR_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v2
mulijet, e/m enriched	/QCD_Pt-20to30_EMEnriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1
mulijet, e/m enriched	/QCD_Pt-30to80_EMEnriched_TuneZ2_7TeV-pythia/Fall11-PU_S6_START42_V14B-v1
mulijet, e/m enriched	/QCD_Pt-80to170_EMEnriched_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v2
mulijet, $b/c \rightarrow e$	/QCD_Pt-20to30_BCtoE_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1
mulijet, $b/c \rightarrow e$	/QCD_Pt-30to80_BCtoE_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1
mulijet, $b/c \rightarrow e$	/QCD_Pt-80to170_BCtoE_TuneZ2_7TeV-pythia/Fall11-PU_S6_START42_V14B-v1
mulijet, μ enriched	/QCD_Pt-20_MuEnrichedPt-15_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1

Table A.4.: List of simulated datasets used for the study of the Q^2 jet-parton matching uncertainty on the cross section measurement.

Event Content	DAS name
$t\bar{t}$	/TTjets_TuneZ2_matchingup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v2
$t\bar{t}$	/TTjets_TuneZ2_matchingdown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v2
W + jets	/WJetsToLNu_TuneZ2_matchingup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1
W + jets	/WJetsToLNu_TuneZ2_matchingdown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1
Z + jets	/ZJetsToLL_TuneZ2_matchingup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1
Z + jets	/ZJetsToLL_TuneZ2_matchingdown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1

Table A.5.: List of simulated datasets used for the study of the factorisation scale uncertainty on the cross section measurement.

Event Content	DAS name
$t\bar{t}$	/TTjets_TuneZ2_scaleup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1
$t\bar{t}$	/TTjets_TuneZ2_scaledown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v2
W + jets	/WJetsToLNu_TuneZ2_scaleup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1
W + jets	/WJetsToLNu_TuneZ2_scaledown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1
Z + jets	/ZJetsToLL_TuneZ2_scaleup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1
Z + jets	/ZJetsToLL_TuneZ2_scaledown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1

Table A.6.: List of simulated datasets with varying top quark mass assumptions.

Top quark mass (GeV)	DAS name
161.5	/TTJets_TuneZ2_mass161_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3
163.5	/TTJets_TuneZ2_mass163_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3
166.5	/TTJets_TuneZ2_mass166_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3
169.5	/TTJets_TuneZ2_mass169_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3
175.5	/TTJets_TuneZ2_mass175_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3
178.5	/TTJets_TuneZ2_mass178_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3
181.5	/TTJets_TuneZ2_mass181_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3
184.5	/TTJets_TuneZ2_mass184_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3

A.2.3. Selection of Simulated Events

Table A.7 shows the number of simulated events after each selection step for the signal and background samples under consideration. When comparing the number of expected events for $4.6\,\mathrm{fb}^{-1}$ of recorded luminosity in Table 4.2 to the number of raw simulated events in Table A.7 it becomes apparent that the number of simulated multijet events is too low by almost two orders of magnitude. 25 simulated events represent the expectation of 1919 events in $4.6\,\mathrm{fb}^{-1}$ of collision data. Consequently, this background contribution is estimated via sideband regions in collision data as described in Section 6.2. For top quark pair events and single top quark events the number of simulated events is higher than the expectation for $4.6\,\mathrm{fb}^{-1}$ by one order of magnitude. For W + jets and Z + jets events the number of simulated events and the number of predicted events are of the same order of magnitude.

Table A.7.: Raw number of selected simulated events for all considered datasets. The last row shows the expectation after the final selection step in $4.6\,\mathrm{fb}^{-1}$ of collision data.

Selection	$t\bar{t}$ (e + jets)	$t\bar{t}$ (other)	W + jets	Z + jets	multijet[1]	t, s-chan	t, t-chan	t, tW-chan
Produced	12731620	73957992	90437607	39534966	109243736	452820	6581070	1972504
Trigger	6658307	2769829	269774	571728	5263	8192	99283	154020
1 iso. e	4592892	1594749	137821	260569	448	4870	61286	101511
μ veto	4591058	1250911	137812	260192	448	4870	61262	97798
loose e veto	4582133	1050393	137664	125517	448	4865	61170	94924
≥ 1 jets	4469560	991538	75427	69562	328	4431	53922	88491
≥ 2 jets	3698451	728988	36383	25888	150	2985	33779	61310
≥ 3 jets	2864153	468404	16594	9818	70	1661	16964	38721
≥ 4 jets	1902595	266165	5458	3367	25	588	5321	20138
expt. $4.6\,\mathrm{fb}^{-1}$	16812	2343	8746	1201	1919	28	242	742

[1] For multijet events a preselection with the requirement of one offline electron with a transverse momentum of $p_\mathrm{T} > 15\,\mathrm{GeV}/c$ and three offline jets with a transverse momentum of $p_\mathrm{T} > 15\,\mathrm{GeV}/c$ has been applied between the *Produced* and the *Trigger* step.

B. Electron+Jets Trigger in 2011

Section 2.2.7 describes that the trigger menu changed repeatedly in 2011. An electron + jets trigger with an isolation requirement for top quark physics was only available in the "1e33", "2e33", "3e33" and "5e33" menu. The changes of the trigger efficiency throughout the different trigger menus is discussed in the following section.

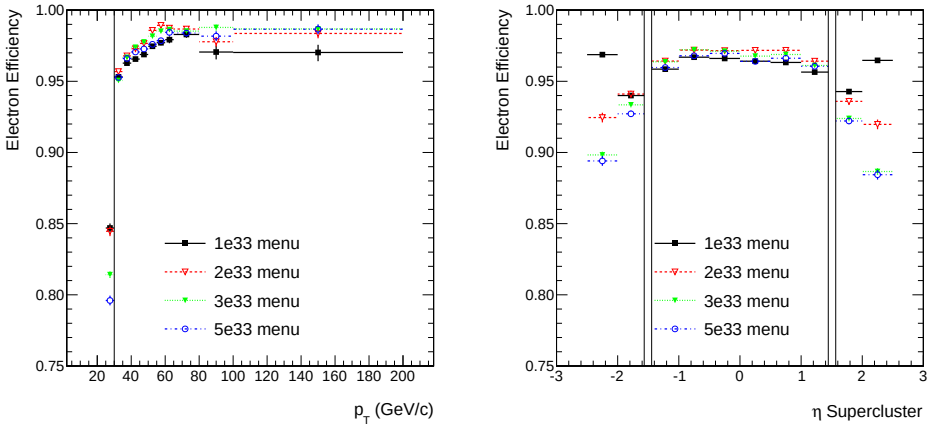
The trigger efficiency is measured separately for the hadronic and for the leptonic part of the electron + jets trigger as described in Section 5.1.1. In Section 5.1.1 the luminosity weighted scale factors are derived for the analysis. In this section the efficiencies in the individual trigger menus are compared.

B.1. Electron Trigger Efficiency in Various Trigger Menus

The electron leg trigger efficiency is measured using the Tag & Probe method. The efficiency is measured for the full HLT trigger path, for only the Level-1 trigger seed and for the pure High-Level Trigger efficiency with respect to the Level-1 trigger, the latter denoted as *HLT/L1*.

Figure B.1(a) shows that there are small changes on the order of 1% in the efficiency observed in the various trigger menus in the intermediate transverse momentum regime from $50\,\mathrm{GeV}/c$ to $70\,\mathrm{GeV}/c$. Nevertheless, in all trigger menus after the "1e33" menu the same plateau efficiency of 98.5% is reached. In the "1e33" menu an efficiency of 97% is measured. From Figure B.1(b) it becomes clear that the efficiency in the barrel region of the detector was very stable while a reduction of the efficiency in the endcap region is observed. This is due to stricter selection criteria that have been applied in the course of the 2011 data taking period to reduce the trigger rate. The contribution of the trigger rate from the endcap region can become large due to the larger background from the beam and proton remnants. Figure B.2 shows that the trigger efficiency was stable for all trigger menus with respect to the number of pile-up interactions in the event.

Details on the trigger efficiency, separated in the pure Level-1 trigger efficiency and the High-Level Trigger efficiency with respect to Level-1 are helpful to understand the composition of the overall trigger efficiency. Figure B.3 shows the Level-1 trigger efficiency for different trigger menus in the 2011 data taking period. Figure B.3(a) shows the turn-on behaviour of the Level-1 trigger as a function of the transverse momentum p_T of the offline reconstructed electron. In the "1e33" menu the Level-1 seed for the High-Level Trigger was *SingleEG12*, that means imposing a p_T threshold of $12\,\mathrm{GeV}/c$. Starting from the "2e33" menu and throughout the "3e33" and "5e33" menu the Level-1 seed was changed to *SingleEG20* with a p_T



(a) Electron trigger efficiency as a function of transverse momentum of the electron.

(b) Electron trigger efficiency as a function of the pseudorapidity η of the supercluster associated to the electron.

Figure B.1.: Electron trigger efficiency as a function of p_{T} and η_{SC} of electron for different trigger menus. The efficiency shown is the product of Level-1 Trigger efficiency and High-Level Trigger efficiency.

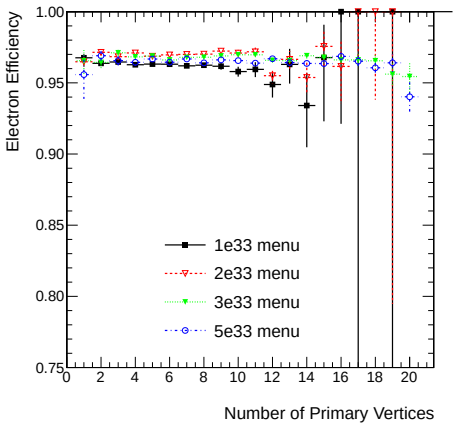
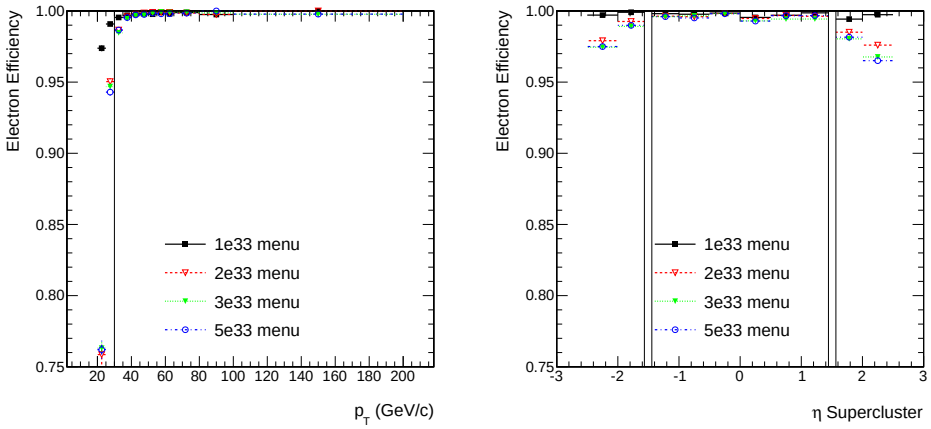


Figure B.2.: Electron trigger efficiency as a function of the number of primary vertices in the event for different trigger menus. The efficiency shown is the product of Level-1 Trigger efficiency and High-Level Trigger efficiency.

threshold of 20 GeV/c. This explains the earlier turn-on behaviour of the Level-1 trigger in the "1e33" menu. In the remaining menus the turn-on behaviour was stable and fully efficient for offline reconstructed electrons with a transverse momentum of $p_{\mathrm{T}} > 35\,\mathrm{GeV}/c$. Figure B.3(b) shows the Level-1 trigger efficiency as a function of the pseudorapidity η_{SC} of the supercluster associated to the offline reconstructed electron. In the barrel region the trigger is fully efficient throughout all trigger menus within an uncertainty of 1%. In the endcap region the trigger is fully efficient in the "1e33" menu but after the Level-1 seed p_{T} threshold change in later menus from 12 GeV/c to 20 GeV/c the efficiency reduces to 97%.



(a) Electron trigger efficiency as a function of transverse momentum of the electron.

(b) Electron trigger efficiency as a function of the pseudorapidity η of the supercluster associated to the electron.

Figure B.3.: Electron Level-1 trigger efficiency as a function of p_{T} and η_{SC} of the offline reconstructed electron for different trigger menus.

Figure B.4 shows the Level-1 trigger efficiency as a function of the number of primary vertices as an indicator for the number of pile-up interactions in the event. The trigger efficiency is independent of pile-up for all trigger menus.

The pure High-Level Trigger efficiency with respect to the Level-1 trigger is shown in Figure B.5. The High-Level Trigger can be identified as the source for the difference of the order of 1% in the turn-on behaviour of the different trigger menus for the full trigger path. This becomes clear in Figure B.5(a) and due to the fact that the Level-1 trigger was measured to be fully efficient. There is no significant change in the pseudorapidity dependent behaviour of the High-Level Trigger in the barrel region. However, in the endcap region the High-Level Trigger efficiency reduced throughout the 2011 data taking period from 97% to 92% in the pseudorapidity range $2 < |\eta_{\mathrm{SC}}| < 2.5$.

The pile-up dependence of the High-Level Trigger in the barrel region is below 1% in all trigger menus as can be seen from Figure B.6.

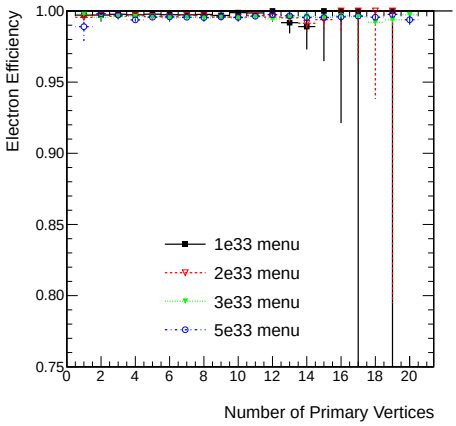
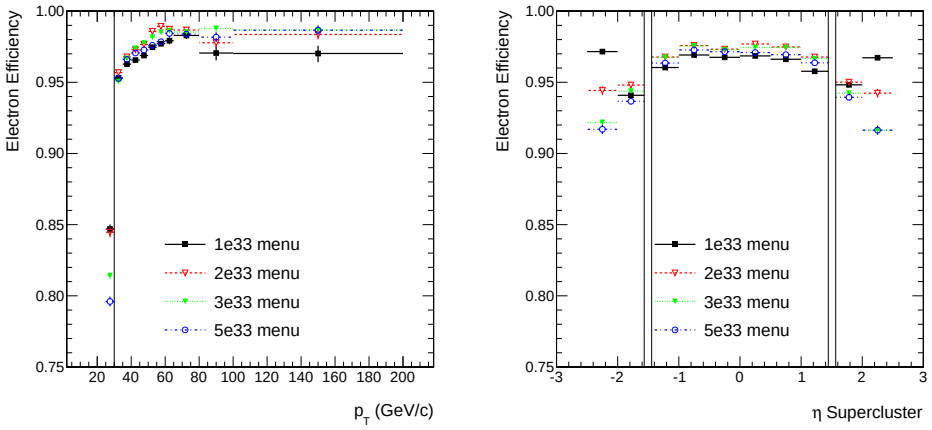


Figure B.4.: Electron Level-1 trigger efficiency as a function of the number of primary vertices in the event for different trigger menus.



(a) Electron trigger efficiency as a function of transverse momentum of the electron.

(b) Electron trigger efficiency as a function of the pseudorapidity η of the supercluster associated to the electron.

Figure B.5.: Electron High-Level Trigger efficiency as a function of p_{T} and η_{SC} of the offline reconstructed electron for different trigger menus.

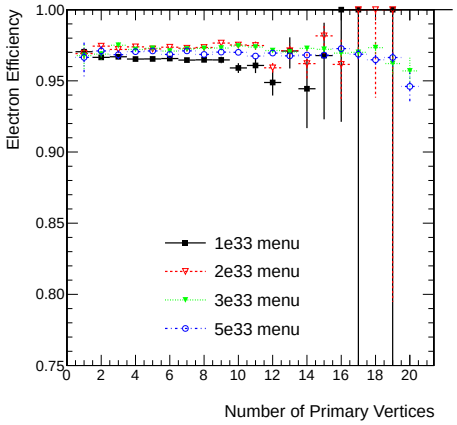


Figure B.6.: Electron High-Level Trigger efficiency as a function of number of primary vertices in the event for different trigger menus.

Figure B.7 shows the trigger efficiency for electrons in the barrel region throughout the 2011 data taking period as a function of the CMS run number. The different data points represent changes of the trigger menu throughout the year. It becomes clear that the efficiency of the electron part of the electron-and-three-jets trigger is very stable throughout the 2011 data taking period. The Level-1 efficiency is very stable on the sub-percent level while the High-Level Trigger efficiency is stable on the 1% level.

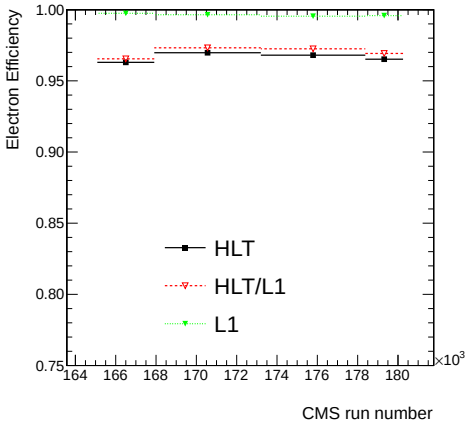


Figure B.7.: Electron trigger efficiency for the full electron part of the trigger path, for only the Level-1 seed of the trigger and for the High-Level Trigger part with respect to the Level-1 trigger. The efficiency is shown as a function of the CMS run number. The individual points represent the average efficiency for a given trigger menu in the order "1e33", "2e33", "3e33" and "5e33".

B.2. Jet Trigger Efficiency in Various Trigger Menus

Figure B.8 shows the trigger efficiency as a function of the transverse momentum p_{T} of the fourth-leading jet in p_{T} in the event. For all trigger menus the jet part of the trigger path is fully efficient within an uncertainty of 1%. This behaviour is a consequence of the offline event selection of four jets with at least two jets above $p_{\mathrm{T}} > 50\,\mathrm{GeV}/c$, one additional jet with $p_{\mathrm{T}} > 40\,\mathrm{GeV}/c$ and at least a fourth jet with $p_{\mathrm{T}} > 30\,\mathrm{GeV}/c$.

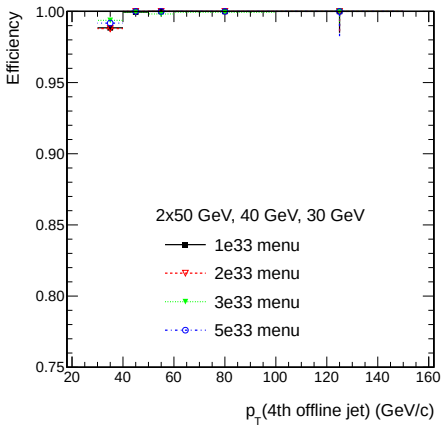


Figure B.8.: Jet trigger efficiency for the different trigger menus of the 2011 data taking period.

List of Figures

1.1. Proton-proton collision in an accelerator. . . . 9
1.2. CTEQ 6.6 parton distribution functions. . . . 11
1.3. Feynman diagram top quark pair production. . . . 12
1.4. Standard Model cross sections. . . . 13
1.5. Feynman diagrams single top quark production. . . . 14
1.6. Top quark pair decay in the semi-leptonic channel. . . . 15

2.1. Acceleration chain for protons. . . . 18
2.2. Layout of the LHC. . . . 19
2.3. Schematic view of the CMS detector. . . . 21
2.4. Tracker layout. . . . 23
2.5. Tracker resolution in collision data. . . . 23
2.6. Tracker performance in 2011. . . . 24
2.7. Schematic view of the CMS ECAL. . . . 25
2.8. ECAL resolutions in 2011. . . . 25
2.9. Dielectron mass in 2011. . . . 26
2.10. HCAL Tower map. . . . 27
2.11. HCAL response as a function of tracker track p_{T}. . . . 28
2.12. E_T^{miss} resolution in 2011. . . . 28
2.13. Schematic view of the CMS muon system. . . . 29
2.14. DT and CSC performance in 2011. . . . 30
2.15. RPC performance in 2011. . . . 30
2.16. Dimuon mass in 2011. . . . 31
2.17. Level-1 trigger electron reconstruction. . . . 34
2.18. Level-1 trigger jet reconstruction. . . . 36
2.19. Slice through the CMS detector. . . . 38
2.20. Number of primary vertices versus pile-up. . . . 39
2.21. Pile-up in 2011 data. . . . 40
2.22. Comparison jet clustering algorithms. . . . 44
2.23. Jet correction scheme. . . . 45
2.24. Jet energy scale uncertainties. . . . 46
2.25. Recorded luminosity in 2011. . . . 47

3.1. Structure of a particle physics process. . . . 50
3.2. Workflow in the CMSSW framework. . . . 55
3.3. Hierarchical structure of Grid centres. . . . 57
3.4. Grid job workflow. . . . 58

4.1. Top quark event display. 65
4.2. Lepton isolation cone. 68
4.3. 2011 pile-up in data and simulation. 75

5.1. Electron trigger efficiency versus p_{T}. 81
5.2. Electron trigger efficiency versus pseudorapidity η. 82
5.3. Electron trigger efficiency versus primary vertices. 83
5.4. Electron trigger efficiency versus number of jets. 84
5.5. Electron trigger efficiency versus number of jets. 84
5.6. 2-dimensional scale factor for electron trigger. 85
5.7. Jet trigger efficiency with loose jet selection. 87
5.8. Jet trigger efficiency variation due to jet energy scale uncertainty. . . 88
5.9. Jet trigger efficiency with loose jet selection. 89
5.10. Jet trigger efficiency pile-up dependence. 90
5.11. Electron identification efficiency versus p_{T}. 92
5.12. Electron identification efficiency versus pseudorapidity η. 93
5.13. Electron identification efficiency versus primary vertices. 93
5.14. Electron identification efficiency versus number of jets. 94
5.15. Electron efficiency after conversion rejection versus p_{T}. 94
5.16. Electron efficiency after conversion rejection versus pseudorapidity. . 95
5.17. Electron efficiency after conversion rejection versus primary vertices. 96
5.18. Electron efficiency after conversion rejection versus number of jets. . 96
5.19. Electron isolation efficiency versus p_{T}. 97
5.20. Electron isolation efficiency versus pseudorapidity η. 98
5.21. Electron isolation efficiency versus primary vertices. 99
5.22. Electron isolation efficiency versus number of jets. 99

6.1. M3 distribution on data. 106
6.2. Anti-Relative-Isolation definition. 108
6.3. Anti-Relative-Isolation regions. 111
6.4. Template shapes. 112
6.5. Results of template fit to pseudo-data. 114
6.6. Pull distributions for template fit parameter. 115
6.7. Template fit dependence on pseudo-Data. 116
6.8. M3 distribution after fit to data. 118
6.9. Template fit dependence on collision data. 120
6.10. Electron transverse momentum and impact parameter. 121
6.11. Electron pseudorapidity and azimuthal angle. 122
6.12. Electron pseudorapidity with Standard Model background. 122
6.13. E_T^{miss} and H_{T}. 123
6.14. First jet transverse momentum and pseudorapidity. 124
6.15. Second jet transverse momentum and pseudorapidity. 124
6.16. Third jet transverse momentum and pseudorapidity. 125
6.17. Fourth jet transverse momentum and pseudorapidity. 125

6.18. All jets transverse momentum and pseudorapidity. 126
6.19. All jets azimuthal angle and number of jets. 127
6.20. Number of primary vertices and minimal distance $\Delta R(\text{jet}, e)$. 128
6.21. ΔR(jet, electron) versus relative isolation. 128

7.1. Statistical uncertainty on template shapes. 131
7.2. Template changes due to pile-up interaction uncertainty. 132
7.3. Template changes due to factorisation scale uncertainty. 134
7.4. Template changes due to jet-parton matching uncertainty. 135
7.5. Top quark selection efficiency and shape versus top quark mass. . . . 136
7.6. Uncertainty due to top quark mass. 137
7.7. Efficiency and shape uncertainty due to top quark mass. 138
7.8. Template changes due to uncertainty on W/Z cross section ratio. . . 139
7.9. Template changes due to jet energy resolution uncertainty. 140
7.10. Template changes due to jet energy scale uncertainty. 141
7.11. Cross section in different relative isolation slices. 144
7.12. Uncertainty on multijet contamination removal. 144
7.13. Template changes due to parton distribution function uncertainty. . 147

8.1. Template shapes in electron/muon + jets channels. 154
8.2. Pull distributions for fit in combined channels. 156
8.3. M3 distribution after fit to data in combined channel. 159
8.4. Uncertainty due to top quark mass in combined channels. 160
8.5. Mass efficiency and shape uncertainty in combined channels. 161
8.6. Cross section in different relative isolation slices. 163

9.1. Mass from cross section in electron + jets channel. 170
9.2. Mass from cross section in combined channels. 171

10.1. Comparison cross section results. 177
10.2. Cross section as a function of centre-of-mass energy. 178

B.1. Electron trigger menu efficiency versus p_{T} and η_{SC}. 196
B.2. Electron trigger menu efficiency versus primary vertices. 196
B.3. Electron Level-1 menu efficiency versus p_{T} and η_{SC}. 197
B.4. Electron Level-1 menu efficiency versus primary vertices. 198
B.5. Electron HLT menu efficiency versus p_{T} and η_{SC}. 198
B.6. Electron HLT menu efficiency versus primary vertices. 199
B.7. Electron menu efficiency versus CMS run number. 200
B.8. Jet trigger menu efficiency. 201

List of Tables

1.1. Quantum numbers of fundamental particles. 5
1.2. Gauge bosons of the Standard Model. 6

2.1. Trigger menus in 2011. 33

4.1. Standard Model predictions of cross sections. 64
4.2. Scaled event numbers after selection steps. 71
4.3. Theory prediction of number of events at each selection step. 72
4.4. Event weights W branching fraction. 75
4.5. Jet energy resolution scale factors. 76

5.1. Identification and isolation criteria for trigger electrons 78
5.2. Sequence of trigger modules. 79
5.3. Summary trigger scale factors 91
5.4. Summary electron identification scale factors 100

6.1. Isolation slices definition. 109
6.2. Contamination multijet slices. 110
6.3. Result of template fits to pseudo-experiments. 115
6.4. Correlations between fit parameters in data. 118
6.5. Result of template fit to 2011 collision data. 119
6.6. Selection efficiencies fit contributions. 119

7.1. Uncertainty due to jet energy scale and resolution. 142
7.2. Uncertainty due to lepton/trigger efficiencies. 143
7.3. Comparison shape versus efficiency. 149
7.4. Comparison systematic efficiencies. 150
7.5. Summary of systematic uncertainties. 151

8.1. Result on pseudo-experiments in combined channels. 156
8.2. Correlations between fit parameters in data. 158
8.3. Result of fit to 2011 collision data in combined channels. 159
8.4. Comparison shape versus efficiency. 164
8.5. Summary of systematic uncertainties in combined channels. 165

10.1. Comparison measured cross sections. 176

A.1. Primary datasets for collision data. 186

A.2. Data quality definition. 187
A.3. Nominal simulated datasets. 188
A.4. Simulated datasets for systematic Q^2 matching studies. 189
A.5. Simulated datasets for systematic factorisation scale studies. 190
A.6. Simulated datasets with varying top quark mass. 191
A.7. Raw number of selected events. 193

Bibliography

[1] M. Planck, "Über irreversible Strahlungsvorgänge," *Annalen der Physik* **306** (1900), 1, 69–122. doi:10.1002/andp.19003060105.

[2] W. Yao et al., "Review of Particle Physics," *Journal of Physics G: Nuclear and Particle Physics* **33** (2006) 1–1232.

[3] S. Walsh, "Measurement of the Top Quark Pair Production Cross Section in the Muon + Jets Channel with the CMS Detector at the LHC,". Ph.D. thesis, Ghent University, 2012.

[4] S. L. Glashow, "Towards a unified theory: Threads in a tapestry," *Rev. Mod. Phys.* **52** (1980) 539–543. doi:10.1103/RevModPhys.52.539.

[5] A. Salam, "Gauge unification of fundamental forces," *Rev. Mod. Phys.* **52** (1980) 525–538. doi:10.1103/RevModPhys.52.525.

[6] S. Weinberg, "Conceptual foundations of the unified theory of weak and electromagnetic interactions," *Rev. Mod. Phys.* **52** (1980) 515–523. doi:10.1103/RevModPhys.52.515.

[7] D. Perkins, "Introduction to High Energy Physics". Cambridge University Press, 2000.

[8] A. Quadt, "Top quark physics at hadron colliders," *The European Physical Journal C - Particles and Fields* **48** (2006), 3, 835–1000. doi:10.1140/epjc/s2006-02631-6.

[9] B. Klein, "Application of Virtualisation Techniques to Grid Resources and Reconstruction of Heavy Resonances Decaying to Quark and Gluon Pairs with the CMS Detector at the LHC," *diploma thesis, University of Karlsruhe* (2008). IEKP-KA/2008-23.

[10] P. J. Mohr, B. N. Taylor, and D. B. Newell, "CODATA Recommended Values of the Fundamental Physical Constants: 2010," arXiv:1203.5425.

[11] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton fragmentation and string dynamics,". doi:10.1016/0370-1573(83)90080-7.

[12] D. T. Goldman and R. J. Bell, "International system of units (SI)," *NASA STI/Recon Technical Report N* **87** (July, 1986) 20444.

[13] R. Field, "The Underlying Event in Hard Scattering Processes," *Arxiv preprint hep-ph/0201192* (2002).

[14] W. Heisenberg, "The Physical Principles of the Quantum Theory". Dover.

[15] Fermi National Accelearator Laboratory, "Design report Tevatron 1 Project,". FERMILAB-DESIGN-1982-01.

[16] P. Schmueser, "The electron-proton colliding beam facility HERA," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **235** (1985), 2, 201 – 208. doi:10.1016/0168-9002(85)90554-6.

[17] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al., "Implications of CTEQ global analysis for collider observables," *Phys.Rev.* **D78** (2008) 013004, arXiv:0802.0007. doi:10.1103/PhysRevD.78.013004.

[18] C. Coll, "Physics Technical Design Report Vol. 1: Detector Performance and Software," *CERN/LHCC* **1** (2006) 2006.

[19] **D0** Collaboration, S. Abachi et al., "Observation of the Top Quark," *Phys. Rev. Lett.* **74** (Apr, 1995) 2632–2637. doi:10.1103/PhysRevLett.74.2632.

[20] Fermilab, http://www.fnal.gov/.

[21] S. W. Herb, D. C. Hom, L. M. Lederman, J. C. Sens, H. D. Snyder, J. K. Yoh, J. A. Appel, B. C. Brown, C. N. Brown, W. R. Innes, K. Ueno, T. Yamanouchi, A. S. Ito, H. Jöstlein, D. M. Kaplan, and R. D. Kephart, "Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions," *Phys. Rev. Lett.* **39** (Aug, 1977) 252–255. doi:10.1103/PhysRevLett.39.252.

[22] Martin-Stirling-Watts-Thorne collaboration , http://www.hep.phy.cam.ac.uk/~wjs/plots/plots.html.

[23] A. V. Feschenko, C. E. Hill, A. M. Lombardi, A. Liou, R. Scrivens, and M. Vretenar, "Tests of the CERN proton linac performance for LHC-type beams. oai:cds.cern.ch:455800,".

[24] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, "LHC Design Report". CERN, Geneva, 2004.

[25] Cern Bulletin, BUL-NA-2010-077.

[26] O. Bruning, H. Burkhardt, and S. Myers, "The Large Hadron Collider. oai:cds.cern.ch:1443022," *Prog. Part. Nucl. Phys.* **67** (Apr, 2012) 705–734. mult. p. Published in Progress in Particle and Nuclear Physics (2012), 10.1016/j.ppnp.2012.03.001.

[27] **ATLAS** Collaboration, "ATLAS: Detector and physics performance technical design report. Volume 1,". CERN-LHCC-99-14, ATLAS-TDR-14.

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

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

[30] **ALICE** Collaboration, "ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC". LHC Tech. Proposal. CERN, Geneva, 1995.

[31] **LHCb** Collaboration, "LHCb : Technical Proposal". Tech. Proposal. CERN, Geneva, 1998.

[32] O. Adriani et al., "LHCf experiment: Technical Design Report". Technical Design Report LHCf. CERN, Geneva, 2006.

[33] **TOTEM** Collaboration, "The TOTEM Experiment at the CERN Large Hadron Collider," *Journal of Instrumentation* **3** (2008), 08, S08007.

[34] V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, M. Friedl, R. Frühwirth, and et al., "CMS tracking performance results from early LHC operation," *European Physical Journal C* **70** (December, 2010) 1165–1192, arXiv:1007.1988. doi:10.1140/epjc/s10052-010-1491-3.

[35] V. Karimaki, "The CMS tracker system project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.

[36] CMS Tracker Performance, https://twiki.cern.ch/twiki/bin/view/CMSPublic/DPGResultsTRK.

[37] **CMS** Collaboration, "ECAL Detector Performance, 2011 Data,". CMS-DP-2012-007.

[38] Electron/Gamma Phyics Results, https://twiki.cern.ch/twiki/pub/CMSPublic/PhysicsResultsEGM.

[39] **CMS** Collaboration, "HCAL performance from first collisions data,". CMS-DP-2010-025.

[40] **CMS** Collaboration, "MET Performance in 2011 CMS Data,". CMS-DP-2012-003.

[41] **CMS** Collaboration, "The CMS muon project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 1997.

[42] CMS Muon System Performance, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/MuonDPGResults`.

[43] CMS CSC Approved Performance Plots, `https://twiki.cern.ch/twiki/bin/view/CMS/CSCDPGApprovedPerformancePlots`.

[44] **CMS** Collaboration, "Performance of the CMS Resistive Plate Chambers (RPC) in 2011,". CMS-DP-2012-001.

[45] Muon Phyics Results, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsMUO`.

[46] G. L. Bayatyan et al., "CMS TriDAS project: Technical Design Report; 1, the trigger systems". Technical Design Report CMS. 2000. CMS-TDR-006-1.

[47] S. Cittolin, A. Racz, and P. Sphicas, "CMS trigger and data-acquisition project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 2002. CMS-TDR-006-add-2.

[48] P. R. Chumney, S. Dasu, J. Lackey, M. Jaworski, P. Robl, et al., "Level-1 regional calorimeter trigger system for CMS," *eConf* **C0303241** (2003) THHT003, `arXiv:hep-ex/0305047`.

[49] W. Badgett, S. Dasu, W. H. Smith, M. Jaworski, and J. Lackey, "CMS Calorimeter Level 1 Regional Trigger Electron Identification," Technical Report CMS-NOTE-1999-026, CERN, Geneva, Apr, 1999.

[50] **CMS** Collaboration, "Electron reconstruction and identification at sqrt(s) = 7 TeV,". CMS-PAS-EGM-10-004.

[51] ECAL Clustering, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideEcalRecoClustering`.

[52] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm," *Journal of High Energy Physics* **2008** (2008), 04, 063.

[53] **CMS** Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET,". CMS-PAS-PFT-09-001.

[54] CMS Outreach, `http://cmsinfo.cern.ch/outreach/`.

[55] **CMS** Collaboration, "Description and performance of the CMS track and primary vertex reconstruction,". CMS-TRK-11-001.

[56] K. Rose, "Deterministic Annealing for Clustering, Compression, Classification, Regression, and Related Optimization Problems," in *Proceedings of the IEEE*, pp. 2210–2239. 1998.

[57] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, "Adaptive Vertex Fitting," Technical Report CMS-NOTE-2007-008, CERN, Geneva, Mar, 2007.

[58] R. E. KALMAN, "A new approach to linear filtering and prediction problem," *Journal of Basic Engineering Transactions* **82** (1960), 1, 34–45.

[59] G. Welch and G. Bishop, "An Introduction to the Kalman Filter,". http://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf.

[60] **CMS** Collaboration, "Performance of CMS muon reconstruction in pp collision events at sqrt(s) = 7 TeV," *ArXiv e-prints* (June, 2012) arXiv:1206.4071.

[61] W. Adam, R. Fruhwirth, A. Strandlie, and T. Todorov, "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC," *Journal of Physics G: Nuclear and Particle Physics* **31** (2005), 9, N9.

[62] **CMS** Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS," *J. Instrum.* **6** (Jul, 2011) P11002. 67 p.

[63] **CMS** Collaboration, "Jet Energy Scale performance in 2011,". CMS-DP-2012-006.

[64] **CMS** Collaboration, "Jet and MET Commissioning Results from 7 TeV Collision Data,". CMS-DP-2010-014.

[65] Jet Energy Scale Uncertainties, https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources.

[66] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas," *Phys.Lett.* **B659** (2008) 119–126, arXiv:0707.1378. doi:10.1016/j.physletb.2007.09.077.

[67] M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets," *JHEP* **0804** (2008) 005, arXiv:0802.1188. doi:10.1088/1126-6708/2008/04/005.

[68] **CMS** Collaboration, "Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update,". CMS-PAS-SMP-12-008.

[69] CMS Public Luminosity Results, https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults2011.

[70] S. V. D. Meer, "ISR-PO/68-31 - KEK68-64,".

[71] T. Sjostrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP* **0605** (2006) 026.

[72] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, et al., "New generation of parton distributions with uncertainties from global QCD analysis," *JHEP* **0207** (2002) 012, arXiv:hep-ph/0201195.

[73] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms," *JHEP* **0411** (2004) 040, arXiv:hep-ph/0409146. doi:10.1088/1126-6708/2004/11/040.

[74] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," *JHEP* **0711** (2007) 070, arXiv:0709.2092. doi:10.1088/1126-6708/2007/11/070.

[75] 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* **1006** (2010) 043, arXiv:1002.2581. doi:10.1007/JHEP06(2010)043.

[76] S. Frixione, F. Stoeckli, P. Torrielli, B. R. Webber, and C. D. White, "The MCaNLO 4.0 Event Generator," arXiv:1010.0819.

[77] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput.Phys.Commun.* **178** (2008) 852–867, arXiv:0710.3820. doi:10.1016/j.cpc.2008.01.036.

[78] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, et al., "Herwig++ Physics and Manual," *Eur.Phys.J.* **C58** (2008) 639–707, arXiv:0803.0883. doi:10.1140/epjc/s10052-008-0798-9.

[79] Z. Was, "TAUOLA the library for tau lepton decay, and KKMC / KORALB / KORALZ /... status report," *Nucl.Phys.Proc.Suppl.* **98** (2001) 96–102, arXiv:hep-ph/0011305. doi:10.1016/S0920-5632(01)01200-2.

[80] J. von Neumann and S. Ulam, "Bulletin A.M.S., Abstract 51-9-165 (1945),".

[81] G. COWAN, "Statistical data analysis". Clarendon Press, Oxford, 1998.

[82] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond," *JHEP* **1106** (2011) 128, arXiv:1106.0522. doi:10.1007/JHEP06(2011)128.

[83] PYTHON, www.python.org/.

[84] N. D. Christensen and C. Duhr, "FeynRules - Feynman rules made easy," *Comput.Phys.Commun.* **180** (2009) 1614–1641, arXiv:0806.4194. doi:10.1016/j.cpc.2009.02.018.

[85] International Organization for Standardization, "ISO/IEC 14882:1998: Programming languages — C++". September, 1998.

[86] International Organization for Standardization, "Draft International Standard ISO/IEC 1539-1:2004(E): Information technology — Programming languages — Fortran Part 1: Base Language". May, 2004.

[87] T. Stelzer and W. Long, "Automatic generation of tree level helicity amplitudes," *Comput.Phys.Commun.* **81** (1994) 357–371, arXiv:hep-ph/9401258. doi:10.1016/0010-4655(94)90084-1.

[88] T. Sjostrand, "High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4," *Comput.Phys.Commun.* **82** (1994) 74–90. doi:10.1016/0010-4655(94)90132-5.

[89] S. Agostinelli et al., "Geant4 – a simulation toolkit," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), 3, 250 – 303. doi:10.1016/S0168-9002(03)01368-8.

[90] ROOT, http://root.cern.ch/.

[91] NA49, http://na49info.web.cern.ch.

[92] PAW, http://paw.web.cern.ch/paw/.

[93] CINT, http://root.cern.ch/twiki/bin/view/ROOT/CINT.

[94] PROOF, http://root.cern.ch/twiki/bin/view/ROOT/PROOF.

[95] W. Verkerke and D. Kirkby, "The RooFit toolkit for data modeling," *ArXiv Physics e-prints* (June, 2003) arXiv:arXiv:physics/0306116.

[96] "CMS Computing Model : The CMS Computing Model RTAG," *CMS NOTE* **2004-031** (December, 2004).

[97] CMS global tags, https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideCalAli.

[98] CMSSW, https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideFrameWork.

[99] M. Stavrianakou et al., "An object-oriented simulation program for CMS,". CHEP '04, C04-09-27.

[100] S. Abdoulline et al., "The CMS object-oriented simulation ,". doi:10.1109/NSSMIC.2005.1596421.

[101] J. Allison et al., "Geant4 developments and applications," *IEEE Transactions on Nuclear Science* **53** (February, 2006) 270–278. doi:10.1109/TNS.2006.869826.

[102] **CMS** Collaboration, "The CMS experiment at the CERN LHC," *Journal of Instrumentation* **3** (2008), 08, S08004.

[103] CMS Data Aggregation Service, https://cmsweb.cern.ch/das/.

[104] A. Scheurer, "Algorithms for the Identification of b-Quark Jets with First Data at CMS,". IEKP-KA/2008-19.

[105] SiteDB, https://cmsweb.cern.ch/sitedb/.

[106] PBS Works, http://www.pbsworks.com/.

[107] dCache, http://www.dcache.org/.

[108] CentOS, https://www.centos.org/.

[109] MyriNet, http://www.myricom.com/.

[110] Infiniband, http://www.infinibandta.org/.

[111] Sun Microsystems, "NFS: Network File System Protocol specification,". RFC 1094.

[112] B. List. Why and when to optimise efficiency times purity, https://www.desy.de/~blist/notes/whyeffpur.ps.gz.

[113] N. Kidonakis, "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production," *Phys. Rev. D* **83** (2011), 091503, arXiv:1103.2972.

[114] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-," *Phys. Rev. D* **82** (2010), 054018, arXiv:1005.4451.

[115] N. Kidonakis, "NNLL resummation for s-channel single top quark production," *Phys. Rev. D* **81** (2010), 054028, arXiv:1001.5034.

[116] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order," *Comput. Phys. Commun.* **82** (2011) 2388–2403, arXiv:1011.3540.

[117] N. Kidonakis, "Higher-order corrections to top-antitop pair and single top quark production," arXiv:0909.0037.

[118] Electron/Gamma Phyics Results, https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOP.

[119] **CMS** Collaboration, "Jet Performance in pp Collisions at 7 TeV,". CMS-PAS-JME-10-003.

[120] Cuts in Categories Electron ID, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideCategoryBasedElectronID`.

[121] CMS Shower Shape Variables, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideEgammaShowerShape`.

[122] Electron Effective Area Corrections, `https://twiki.cern.ch/twiki/bin/view/CMS/EgammaEARhoCorrection`.

[123] **CMS** Collaboration, "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV,". CMS-PAS-PFT-10-002.

[124] Pile-Up Reweighting Procedures, `https://twiki.cern.ch/twiki/bin/view/CMS/PileupMCReweightingUtilities`.

[125] Jet Energy Resoultions and Uncertainties, `https://twiki.cern.ch/twiki/bin/view/CMS/JetResolution`.

[126] **CMS** Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS," *Journal of Instrumentation* **6** (November, 2011) 11002, `arXiv:1107.4277`. `doi:10.1088/1748-0221/6/11/P11002`.

[127] V. Abazov, I. Bertram, G. Borissov, H. Fox, P. Ratoff, and A. Sopczak, "Measurement of the ttbar production cross section in ppbar collisions at sqrt(s)=1.96 TeV using kinematic characteristics of lepton+jets events," *Physical Review D* **76** (2007) 092007.

[128] **CMS** Collaboration, "Measurement of the single-top-quark t-channel cross section in pp collisions at sqrt(s) = 7 TeV," `arXiv:1209.4533`.

[129] **CMS** Collaboration, "Search for single top tW associated production in the dilepton decay channel in pp collisions at sqrt s =7 TeV,". CMS-PAS-TOP-11-022.

[130] F. James and M. Roos, "Minuit - a system for function minimization and analysis of the parameter errors and correlations," *Computer Physics Communications* **10** (1975), 6, 343 – 367. `doi:10.1016/0010-4655(75)90039-9`.

[131] L. Demortier and L. Lyons, "Everything you always wanted to know about pulls,". CDF/ANAL/PUBLIC/5776.

[132] Pile-Up Systematic Uncertainties, `https://twiki.cern.ch/twiki/bin/view/CMS/PileupSystematicErrors`.

[133] 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, arXiv:hep-ph/0611129. doi:10.1088/1126-6708/2007/01/013.

[134] MadGraph Summer11 Production Page, https://twiki.cern.ch/twiki/bin/view/CMS/MadGraphSummer11Production.

[135] J. Pumplin et al., "Uncertainties of predictions from parton distribution functions. 2. The Hessian method," *Phys. Rev.* **D65** (2001) 014013, arXiv:hep-ph/0101032. doi:10.1103/PhysRevD.65.014013.

[136] D. Bourilkov and M. R. Whalley, "LHAPDF : PDF Use from the Tevatron to the LHC," Technical Report hep-ph/0605240, FERMILAB, Batavia, IL, May, 2006.

[137] L. Lyons, D. Gibaut, and P. Clifford, "How to combine correlated estimates of a single physical quantity," *Nuclear Instruments and Methods in Physics Research A* **270** (July, 1988) 110–117. doi:10.1016/0168-9002(88)90018-6.

[138] **CDF and D0** Collaborations, T. Aaltonen et al., "Combination of the top-quark mass measurements from the Tevatron collider," arXiv:1207.1069.

[139] **ATLAS and CMS** Collaborations, "Combination of ATLAS and CMS results on the mass of the top quark using up to 4.9 fb-1 of data,". ATLAS-CONF-2012-095.

[140] **D0** Collaboration, V. M. Abazov et al., "Determination of the pole and MSbar masses of the top quark from the $t\bar{t}$ cross section," *Phys.Lett.* **B703** (2011) 422–427, arXiv:1104.2887. doi:10.1016/j.physletb.2011.08.015.

[141] **ATLAS** Collaboration, "Determination of the Top-Quark Mass from the ttbar Cross Section Measurement in pp Collisions at sqrt(s)=7 TeV with the ATLAS detector," Technical Report ATLAS-CONF-2011-054, CERN, Geneva, Apr, 2011.

[142] **CMS** Collaboration, "Determination of the Top Quark Mass from the ttbar Cross Section at sqrt(s) = 7 TeV,". CMS-PAS-TOP-11-008.

[143] V. Ahrens, A. Ferroglia, M. Neubert, B. D. Pecjak, and L. L. Yang, "Precision predictions for the production cross section at hadron colliders," *Physics Letters B* **703** (2011), 2, 135 – 141. doi:10.1016/j.physletb.2011.07.058.

[144] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann, "HATHOR - HAdronic Top and Heavy quarks crOss section calculatoR," *Computer Physics Communications* **182** (2011), 4, 1034 – 1046. doi:10.1016/j.cpc.2010.12.040.

[145] **CMS** Collaboration, "Measurement of ttbar Pair Production Cross Section at sqrt(s)=7 TeV using b-quark Jet Identification Techniques in Lepton + Jet Events,". CMS-PAS-TOP-11-003.

[146] **CMS** Collaboration, "Combination of top pair production cross section measurements,". CMS-PAS-TOP-11-024.

[147] **ATLAS** Collaboration, "Measurement of the ttbar production cross-section in pp collisions at sqrts = 7 TeV using kinematic information of lepton+jets events," Technical Report ATLAS-CONF-2011-121, CERN, Geneva, Aug, 2011.

[148] **ATLAS** Collaboration, "Statistical combination of top quark pair production cross-section measurements using dilepton, single-lepton, and all-hadronic final states at sqrt(s) = 7 TeV with the ATLAS detector,". ATLAS-CONF-2012-024.

[149] **ATLAS and CMS** Collaborations, "Combination of ATLAS and CMS top-quark pair cross section measurements using up to 1.1 fb-1 of data at 7 TeV," Technical Report ATLAS-CONF-2012-134, CERN, Geneva, Sep, 2012.

[150] **CDF** Collaboration, "Combination of CDF top quark pair production cross section measurements with up to 4.6 fb-1,". CDF note 9913.

[151] **D0** Collaboration, V. M. Abazov et al., "Measurement of the top quark pair production cross section in the lepton+jets channel in proton-antiproton collisions at $\sqrt{s}$=1.96 TeV," *Phys.Rev.* **D84** (2011) 012008, arXiv:1101.0124. doi:10.1103/PhysRevD.84.012008.

[152] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., "HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)," *JHEP* **0101** (2001) 010, arXiv:hep-ph/0011363.

[153] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., "HERWIG 6.5 release note," arXiv:hep-ph/0210213.

[154] A. Martin, W. Stirling, R. Thorne, and G. Watt, "Parton distributions for the LHC," *Eur.Phys.J.* **C63** (2009) 189–285, arXiv:0901.0002. doi:10.1140/epjc/s10052-009-1072-5.

[155] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, et al., "A first unbiased global NLO determination of parton distributions and their uncertainties," *Nucl.Phys.* **B838** (2010) 136–206, arXiv:1002.4407. doi:10.1016/j.nuclphysb.2010.05.008.

[156] **H1 and ZEUS** Collaborations, "Combined measurement and QCD analysis of the inclusive $e^{\pm}p$ scattering cross sections at HERA," *Journal of High Energy Physics* **2010** (2010) 1–63. doi:10.1007/JHEP01(2010)109.

[157] **CMS** Collaboration, "Top pair cross section in e/mu+jets at 8 TeV,". CMS-PAS-TOP-11-003.

[158] D. Crockford, "RFC 4627 - The application/json Media Type for JavaScript Object Notation (JSON)," technical report.

[159] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions," *JHEP* **0909** (2009) 111, arXiv:0907.4076. doi:10.1007/JHEP02(2010)011, 10.1088/1126-6708/2009/09/111.

[160] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method," *Eur.Phys.J.* **C71** (2011) 1547, arXiv:1009.2450. doi:10.1140/epjc/s10052-011-1547-z.

[161] **CMS** Collaboration, S. Chatrchyan et al., "Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7$ TeV and Comparison with $\sqrt{s} = 0.9$ TeV," *JHEP* **1109** (2011) 109, arXiv:1107.0330. doi:10.1007/JHEP09(2011)109.

Acknowledgements

I would like to thank Prof. Dr. Martin Grünewald for offering me the possibility to work towards a Ph.D. degree in his group. Throughout the four years of my Ph.D. he provided excellent and prompt support in both physics analysis and organisational matters at all times. I would also like to thank Prof. Dr. Dirk Ryckbosch for his help in many organisational aspects of work and for the excellent work atmosphere he contributed to and established at the Institute of Physics and Astronomy.

I am thankful to Dr. Volker Adler for the numerous discussions about physics and technical aspects of the analysis and his continuous practical help resolving technical problems.

I thank Kelly Beernaert, Dr. Leonardo Benucci, Guillaume Garcia, Dr. Jérémie Lellouch, Nadja Strobbe and Dr. Efe Yazgan for their comments and suggestions for improvements to the text of this thesis. They, together with the other members of the institute in Ghent, created an excellent work atmosphere.

Dr. Silvia Goy Lopez, Dr. Ivan Mikulec and Dr. Slava Valuev provided many very interesting opportunities for studies and were always available for the discussion of results. I consider their help and dedication exemplary.

I thank Dr. Manuel Zeise for his valuable feedback concerning my thesis and the countless fruitful discussions throughout the years. Dr. Fred Stober provided technical help and solutions in various situations. Encounters with current or former members of the Institut für Experimentelle Kernphysik of the KIT in Karlsruhe are always very positive experiences.

Dr. Marta Felcini was of great help in getting introduced to top quark physics and organising work efficiently at the beginning of my Ph.D. studies. This was very much appreciated.

The administrators of the computing resources at the IIHE institute in Brussels were at all times very helpful and patient. The numerous occasions when they solved urgent technical problems outside of the normal working hours, be it late in the evening or on week ends, shall not go unmentioned.

During my studies towards a Ph. D. degree, I enjoyed the time with Dr. Sinéad Walsh very much and I am glad to have had the opportunity to work together.

I thank my parents Josef and Elisabeth, my brother Patrik and my friends for their support.