Measurement of the $t$-Channel Single-Top-Quark-Production Cross Section and the CKM-Matrix Element $V_{tb}$ with the CMS Experiment

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Physiker
Dennis Klingebiel
aus Bielefeld

Berichter: Universitätsprofessor Dr. Martin Erdmann
Universitätsprofessor Dr. Thomas Hebbeker


Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek online verfügbar.
Contents

1 Introduction

2 The Top Quark within the Standard Model of Particle Physics
   2.1 The Standard Model of Particle Physics
      2.1.1 Matter
      2.1.2 Interactions
      2.1.3 Mass Generation
   2.2 The Top Quark
      2.2.1 Modeling of Hadron Collisions
      2.2.2 Top-Quark-Pair Production
      2.2.3 Single-Top-Quark Production
      2.2.4 Top-Quark Decay

3 Experimental Setup
   3.1 The Large Hadron Collider
   3.2 The Compact Muon Solenoid Detector
      3.2.1 The Tracker
      3.2.2 The Calorimeter
      3.2.3 Superconducting Solenoid
      3.2.4 The Muon System
      3.2.5 The Trigger System
      3.2.6 Detector Simulation
   3.3 Reconstruction of Physics Objects
      3.3.1 Vertex Reconstruction
      3.3.2 Particle-Flow Reconstruction
      3.3.3 Muons
      3.3.4 Electrons
      3.3.5 Jets
      3.3.6 Identification of Jets from b-Quark Fragmentation
      3.3.7 Missing Transverse Energy
   3.4 Collision Data and Luminosity Determination
   3.5 Simulation of Signal and Background Processes
      3.5.1 Signal-Event Generation and Normalization
      3.5.2 Background-Event Generation and Normalization
      3.5.3 Summary of Simulated Events and Process Normalization
      3.5.4 Corrections to Simulated Events
   3.6 Software and Libraries

4 Analysis Strategy
   4.1 Blind Analysis and Categorization
   4.2 Phenomenology of t-Channel Single-Top-Quark Production
   4.3 Event Selection
   4.4 Single-Top-Quark Event Reconstruction
### Contents

4.4.1 "b-tagged-jet" Hypothesis .................................................. 100  
4.4.2 Neutrino and W-Boson Reconstruction ................................. 101  
4.4.3 Top-Quark Reconstruction ................................................ 106  
4.5 Boosted Decision Trees ....................................................... 107  
4.5.1 The Main Principle .......................................................... 108  
4.5.2 TMVA Setup And Simulated-Sample Splitting ......................... 110  
4.5.3 Discriminating Variables .................................................. 111  
4.5.4 Boosted-Decision-Tree Classifier Distribution ..................... 121  
4.6 Statistical Inference ........................................................... 124  
4.6.1 Bayesian Evaluation ....................................................... 124  
4.6.2 Construction of the Likelihood Function .............................. 127  
5 Background Estimation and Evaluation of Simulated Event Modeling  133  
5.1 QCD-Multijet-Background Estimation ..................................... 134  
5.1.1 QCD-Multijet Events with Muons ....................................... 136  
5.1.2 QCD-Multijet Events with Electrons .................................. 138  
5.2 Event Yield and Jet-Multiplicity Distribution ......................... 141  
5.3 Top-Quark-Pair- and W-Boson-plus-Jets-Enriched Phase Space ... 144  
5.3.1 Basic Kinematic Quantities .............................................. 146  
5.3.2 Top-Quark Reconstruction .............................................. 155  
5.3.3 BDT-Training Variables .................................................. 159  
5.3.4 BDT-Discriminator Distribution ....................................... 164  
5.4 Signal Region ................................................................. 167  
5.4.1 Basic Kinematic Quantities .............................................. 167  
5.4.2 Top-Quark Reconstruction .............................................. 172  
5.4.3 BDT-Training Variables .................................................. 174  
5.4.4 BDT-Discriminator Distribution ....................................... 177  
6 Sources of Systematic Uncertainties and Expected Sensitivity .... 179  
6.1 Experimental Uncertainties ................................................ 180  
6.1.1 Statistical Precision of Templates for Simulated Events and QCD-Multijet Events ................................................................. 180  
6.1.2 Jet-Flavor-Tagging Efficiency ......................................... 181  
6.1.3 Jet-Transverse-Momentum Resolution ................................. 184  
6.1.4 Jet-Energy Scale ........................................................... 184  
6.1.5 Luminosity Determination .............................................. 187  
6.1.6 Missing Transverse Energy ............................................. 187  
6.1.7 Pile-Up ........................................................................ 187  
6.1.8 Identification, Reconstruction, and Triggering of Charged Leptons ................................................................. 190  
6.1.9 Hadronic-Trigger Efficiencies ........................................... 190  
6.1.10 Charge Ratio for W-Boson-plus-Jets Events ....................... 190  
6.1.11 Uncertainties of Background Normalization ....................... 192  
6.2 Generator-related Uncertainties ............................................ 194  
6.2.1 Modeling of Additional Partons ........................................ 194  
6.2.2 Parton-Distribution Functions .......................................... 196  
6.2.3 Renormalization and Factorization Scales ........................... 197  
6.2.4 Signal Modeling ........................................................... 201  
6.3 Impact of Systematic Uncertainties on the Measurement ......... 203  
6.3.1 Experimental and Normalization Uncertainties .................... 205  
6.3.2 Generator-Related Uncertainties ...................................... 206  
6.4 Expected Sensitivity ............................................................ 210
Contents

6.5 Charge-Ratio Measurement .................................................. 213
   6.5.1 Statistical Model ................................................................. 213
   6.5.2 Expected Sensitivity ................................................................. 215

7 Measurement Results ......................................................... 217
   7.1 Measurement of the Inclusive $t$-Channel Cross Section ................. 217
   7.2 Measurement of the CKM-Matrix Element $|V_{tb}|$ ......................... 226
   7.3 Measurement of the Ratio of Top-Quark-Production and Top-Antiquark-Production Cross Sections $R = \sigma_t / \sigma_{\bar{t}}$ .................................................. 229

8 Conclusion ........................................................................... 233

Bibliography ........................................................................... 235

List of Acronyms ..................................................................... 249

Contributions to Previous Publications (Vorveröffentlichungen) ............ 253
1 Introduction

The top quark was observed at the Tevatron collider in 1995 [1,2]. The existence of top quarks was established by measuring their production in pairs with their antiquarks (ttbar), which occurs via strong-interaction processes. The Tevatron delivered proton-antiproton (pbarp) collisions at center-of-mass energies of 1.8 TeV and 1.96 TeV. These data were used to precisely measure many properties of the top quark and to extensively probe top-quark production via strong interactions [3–5].

Single top quarks are produced by electroweak interactions. Single-top-quark production was observed in pbarp collisions at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV in 2009 [6,7]. The experimental confirmation of single-top-quark production was demanding due to its small production cross section, a challenging experimental signature, and a large number of background processes.

In 2010, the Large Hadron Collider (LHC) began to deliver proton-proton (pp) collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV. The recorded data allow to probe elementary particles and their interactions in a new energy regime. For the first time, precision measurements of the single-top-quark-production mechanisms are feasible.

The total cross section of single-top-quark production is predicted to be 85 pb in pp collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV [8–10]. The dominant production mode is the t-channel with a cross section of 65 pb. The t-channel mode involves the exchange of a space-like W boson.

The motivations to precisely measure the single-top-quark-production cross section in the t-channel are the following. The elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix determine the coupling strengths of weak-charged-current interactions. They are “fundamental parameters” of the Standard Model (SM) of particle physics and have to be determined experimentally [11]. The t-channel-production cross section $\sigma$ is proportional to the squared coupling strength at the Wtb vertex, i.e. $\sigma \propto |V_{tb}|^2$. The precise measurement of the inclusive t-channel-production cross section $\sigma$ offers a unique access to the CKM-matrix element $V_{tb}$ without assuming CKM-matrix unitarity. Any deviation from $|V_{tb}| \approx 1$ hints at new physics phenomena such as additional quarks or modified top-quark interactions.

Furthermore, single-top-quark production in the t-channel is sensitive to the structure of the electroweak interactions at the Wtb vertex [12–15]. The measurement of the inclusive t-channel-production cross section provides constraints on anomalous couplings at the Wtb vertex [16]. In particular, constraints from single-top-quark production are complementary to those obtained from top-quark decays through measurements of the W-boson helicity in top-quark-pair production.

Moreover, in proton-proton collisions, top quarks are expected to be produced more often than top antiquarks by a factor of $\approx 1.85$ in the t-channel production mode [8]. The relative production rates of top quarks and top antiquarks originates from the parton composition of the proton. The measurement of the ratio of top-quark-production and top-antiquark-production cross sections is sensitive to the description of the proton parton-distribution functions [17] [18].
This thesis reports on precise measurements of the inclusive $t$-channel single-top-quark-production cross section, the $|V_{tb}|$ matrix element, and the ratio of $t$-channel top-quark-production and top-antiquark-production cross sections. Proton-proton collisions with a center-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ have been analyzed. The analyzed data were recorded by the Compact Muon Solenoid (CMS) experiment at the LHC and correspond to an integrated luminosity of $1.6\,\text{fb}^{-1}$. In this analysis, signal and background processes are discriminated by exploiting their characteristic signatures with Boosted Decision Trees (BDTs). The signal-production cross section is inferred from data by using the discriminator-output distribution.

Previous measurements of single-top-quark production are in agreement with SM predictions. However, tensions between measurements at the two Tevatron experiments are observed. The D0 experiment measured the $t$-channel cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,\text{TeV}$ to $3.07^{+0.54}_{-0.50}\,\text{pb}$ [19]. The Collider Detector at Fermilab (CDF) experiment measured $1.49^{+0.47}_{-0.42}\,\text{pb}$ [20]. The SM prediction is $2.08 \pm 0.12\,\text{pb}$ at $\sqrt{s} = 1.96\,\text{TeV}$ [8]. The most precise measurement of the $t$-channel-production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,\text{TeV}$ has an uncertainty of 16% [19]. Single-top-quark production was re-discovered with the CMS experiment using $pp$ collisions at $\sqrt{s} = 7\,\text{TeV}$ recorded in 2010. In a previous iteration of the analysis that is described in this thesis, the inclusive $t$-channel cross section was measured with a relative uncertainty of 37% using $36\,\text{pb}^{-1}$ of data [21]. The ATLAS experiment measured the $t$-channel cross section in $1.04\,\text{fb}^{-1}$ of $pp$ collisions at $\sqrt{s} = 7\,\text{TeV}$ with a relative uncertainty of 23% [22]. The most stringent limit on $|V_{tb}|$ without assuming $|V_{tb}|$ matrix unitarity results from a combination of Tevatron measurements and yields $|V_{tb}| = 0.88 \pm 8.0\%$ [23]. The ratio of top-quark-production and top-antiquark-production cross sections was measured in $pp$ collisions at $\sqrt{s} = 7\,\text{TeV}$ with an uncertainty of 12% using data corresponding to $4.7\,\text{fb}^{-1}$ [17] and at $\sqrt{s} = 8\,\text{TeV}$ with an uncertainty of 15% using data corresponding to $12.2\,\text{fb}^{-1}$ [18]. The results of the inclusive cross-section measurement that is described in this thesis were pre-published in [24]. This measurement was combined with two other measurements, which are described in ref. [25–27].

This thesis is organized as follows. Chapter 2 introduces the SM of particle physics, which is the theory that describes the elementary particles, the fundamental interactions among them, and the mass-generation mechanism. Particular emphasis is put on the description of the characteristics, production mechanisms, and decay properties of top quarks. An overview of previous measurements of the single-top-quark-production cross section and the $|V_{tb}|$ matrix element is given. Chapter 3 introduces the experimental framework in which this analysis is performed. The particle-accelerator complex LHC and the measurement of particle collisions with the CMS detector are described. The reconstruction of physics objects, e.g. electrons, muons, or jets of hadrons, is detailed. The determination of the recorded luminosity and the simulation of particle collisions with regard to this analysis are described.

The strategy of this analysis is given in chapter 4. The event selection and top-quark reconstruction are described. The principle of BDTs and the discriminating variables that are used for the classifier training are detailed. The statistical inference, i.e. the methodology to infer the signal-production cross section from measured data, is described in the last part of chapter 4. Chapter 5 is devoted to the description of background estimation and the evaluation of the compatibility of predicted and observed kinematics. Both background-enriched control regions and signal regions are considered. Chapter 6 discusses sources of systematic uncertainties and their impact on the measurements. The expected sensitivities of the measurements are presented. Finally, chapter 7 discusses the results of the measurements and chapter 8 presents the conclusions.
2 The Top Quark within the Standard Model of Particle Physics

The Standard Model (SM) of particle physics is the theory in which the fundamental particles, the fundamental interactions, and the mass-generation mechanism are described. The phenomenology of the SM and its mathematical formulation are discussed in the first part of this chapter.

The elementary particles, which include quarks, leptons, and their antiparticles, are presented in section 2.1.1. Four fundamental forces are observed in nature. These are the electromagnetic, strong, weak, and gravitational forces. The first three fundamental forces constitute the SM of particle physics and are described in section 2.1.2. The gravitational force is described by the general theory of relativity and not detailed in this thesis. It is important on macroscopic scales, whereas it is considered as insignificant on microscopic scales, such as for particle collisions at present colliders.

The “Higgs mechanism” is the mass-generation mechanism of the SM [28]. It describes why some of the force carriers are massive and other are massless. Moreover, elementary particles acquire mass due to interactions with the “Higgs” field. This mechanism is described by Yukawa interactions. An important consequence is that weak-charged currents mix mass states and interaction states of quarks. The strength of these weak-charged currents are described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The mass-generation mechanism and the CKM matrix are discussed in section 2.1.3.

The SM is described in detail e.g. in ref. [11, 28–34]. In particular, the mathematical descriptions and their interpretations given in the first part of this chapter (sec. 2.1.2 and 2.1.3) follow those given in ref. [33, 34].

The second part of this chapter focuses on the characteristics and production mechanisms of top quarks, which are the heaviest of all known elementary particles. The modeling of hadron collisions is addressed on a phenomenological level (sec. 2.2.1). The dominant top-quark-production mode is pair production ($t\bar{t}$) via strong interactions (sec. 2.2.2). Single top (anti)quarks are produced by electroweak interactions. An overview of the production mechanisms of single top quarks and present measurement results is given in section 2.2.3. The motivations for the measurements that are performed in this thesis are detailed, and an overview of physics phenomena beyond the SM is given. Finally, the decay of top quarks is addressed in section 2.2.4.


2 The Top Quark within the Standard Model of Particle Physics

2.1 The Standard Model of Particle Physics

2.1.1 Matter

Matter is composed of point-like elementary fermions, which are quarks (u, d, c, s, t, b), leptons ($\nu_e$, $e$, $\nu_\mu$, $\mu$, $\nu_\tau$, $\tau$), and their corresponding antiparticles. The up-type quarks (u, c, and t) have an electric charge $Q = \frac{2}{3}e$, down-type quarks (d, s, b) have $Q = -\frac{1}{3}e$, charged leptons have $Q = -e$, and neutrinos are electrically neutral. Fermions are particles with spin $\frac{1}{2}\hbar$.

The elementary fermions of the SM are grouped into three generations. An overview of the elementary fermions is given in table 2.1. Antiparticles are not explicitly listed in table 2.1, but have the same mass (within the current experimental resolution) as the corresponding particle and opposite electric charge.

<table>
<thead>
<tr>
<th>Generation</th>
<th>Name (Symbol)</th>
<th>Mass [GeV/c^2]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quarks</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>up (u)</td>
<td>$(2.3^{+0.7}_{-0.5}) \times 10^{-3}$</td>
</tr>
<tr>
<td></td>
<td>down (d)</td>
<td>$(4.8^{+0.3}_{-0.3}) \times 10^{-3}$</td>
</tr>
<tr>
<td>2</td>
<td>charm (c)</td>
<td>1.275 ± 0.025</td>
</tr>
<tr>
<td></td>
<td>strange (s)</td>
<td>$(95 \pm 5) \times 10^{-3}$</td>
</tr>
<tr>
<td>3</td>
<td>top (t)</td>
<td>173.20 ± 0.87</td>
</tr>
<tr>
<td></td>
<td>bottom (b)</td>
<td>4.18 ± 0.03</td>
</tr>
<tr>
<td><strong>Leptons</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>electron neutrino ($\nu_e$)</td>
<td>$&lt; 2 \times 10^{-9}$ (95% CL)</td>
</tr>
<tr>
<td></td>
<td>electron ($e$)</td>
<td>$(0.510998928 \pm 0.000000011) \times 10^{-3}$</td>
</tr>
<tr>
<td>2</td>
<td>muon neutrino ($\nu_\mu$)</td>
<td>$&lt; 0.19 \times 10^{-3}$ (90% CL)</td>
</tr>
<tr>
<td></td>
<td>muon ($\mu$)</td>
<td>$0.1056583715 \pm 0.0000000035$</td>
</tr>
<tr>
<td>3</td>
<td>tau neutrino ($\nu_\tau$)</td>
<td>$&lt; 18.2 \times 10^{-3}$ (95% CL)</td>
</tr>
<tr>
<td></td>
<td>tau ($\tau$)</td>
<td>1.77682 ± 0.00016</td>
</tr>
</tbody>
</table>

Table 2.1: Elementary fermions in the SM. Each particle has a corresponding antiparticle with equal mass. All mass values are determined in the modified minimal-subtraction scheme (MS) [11], except for the top-quark mass, which is a direct measurement (cf. sec. 2.2). Values are taken from ref. [11] unless another reference is quoted.

The lightest charged lepton is the electron with a mass of $m_e \approx 0.51\text{MeV}/c^2$ [11]. The $\tau$ lepton is as heavy as $m_\tau \approx 1.78\text{GeV}/c^2$. While electrons are stable particles, the lifetime of charged leptons of the second generation ($\tau_\mu \approx 10^{-6} \text{s}$) and third generation ($\tau_\tau \approx 10^{-13} \text{s}$) significantly decreases. It was found that neutrinos oscillate among different lepton flavors, which implies that the mass differences of neutrinos among the three generations are nonzero [28]. Upper bounds on neutrino masses have been set in direct measurements, whereas the most stringent limit is for the electron-neutrino mass ($m_{\nu_e} < 2 \text{eV}/c^2$ at 95% confidence level (CL)) [11]. Therefore, the electron-neutrino mass is at least five orders of magnitudes smaller than the mass of the lightest charged lepton. While the mass-generation mechanism for quarks and charged leptons is part of the SM (and discussed in section 2.1.3), the mass-generation mechanism for neutrinos is subject of current research [11, 36, 37].

The top quark is the by far heaviest known elementary particle in the SM. The top quark has a mass that is two orders of magnitudes larger than the second heaviest fermion. Quarks
and antiquarks do not exist freely, but form hadrons. Top quarks decay before they can form hadrons (cf. sec. 2.2). Hadrons are bound states of (anti)quarks and gluons. Hadrons are either baryons, which are pairs of three quarks (qqq) or three antiquarks (¯¯qqq), or mesons, which are quark-antiquark (q¯¯q) pairs. The lightest baryons are protons (uud) and neutrons (udd) with masses of \( m_p \approx 0.938 \text{ GeV}/c^2 \) and \( m_n \approx 0.940 \text{ GeV}/c^2 \). The lightest mesons are pions, which consist of up- and down quarks. Neutral pions (\( \pi^0 \)) and charged pions (\( \pi^\pm \)) have masses of \( m_{\pi^0} \approx 0.135 \text{ GeV}/c^2 \) and \( m_{\pi^\pm} \approx 0.140 \text{ GeV}/c^2 \), which are between the masses of the lightest charged lepton and baryon [11].

While electrons, up quarks, and down quarks are constituents of “ordinary matter”, other fermions are produced e.g. in cosmic-ray-induced showers, nuclear reactions, and particle colliders (cf. [28]). The interactions between fermions, as well as the mechanism that generates their masses, are briefly discussed in the next sections.

### 2.1.2 Interactions

The electromagnetic, strong, and weak forces constitute the fundamental interactions of the SM of particle physics. The SM mathematically describes the interactions between elementary particles by gauge theories within the framework of relativistic quantum-field theories. The SM is based on the \( SU(3) \otimes SU(2) \otimes U(1) \) symmetry groups [33]. The group \( SU(3) \) corresponds to strong interactions. The groups \( SU(2) \otimes U(1) \) describe electroweak interactions, which are based on a unified description of both electromagnetic and weak interactions. When demanding local gauge invariance to the Lagrangian density, gauge fields need to be introduced. Their excitation modes are particles (“gauge bosons”) with spin \( \hbar \). The gauge bosons “mediate” the forces [28]. A summary of the gauge bosons and their mediated forces is given in table 2.2. The gauge bosons of the three interactions are supposed to be massless within their mathematical description. However, the three bosons corresponding to the weak force are found to be quite massive with \( m_{W^\pm} = 80.4 \text{ GeV}/c^2 \) and \( m_Z = 91.2 \text{ GeV}/c^2 \) [11]. The generation of masses is explained by spontaneous symmetry breaking of the electroweak gauge symmetry, which is described by the Higgs mechanism as discussed in the next section (2.1.3).

In this section, the strong force and its mathematical foundations are discussed first. The second part of this section refers to the electroweak interactions. The mathematical descriptions and their interpretations follow those given in ref. [33, 34].

| Mediated force | Name (Symbol) | Mass [GeV/c²] | Spin \( |\hbar| \) | Electric charge \( e \) | Weak isospin \( T_3 \) | Color charge |
|----------------|---------------|---------------|----------------|----------------|----------------|------------|
| Electromagnetic | Photon (\( \gamma \)) | \(< 1 \times 10^{-27}\) | 1 | 0 | 0 | - |
| Strong | Gluon (\( g \)) | 0 | 1 | 0 | 0 | r,g,b |
| Weak | W bosons (\( W^\pm \)) | 80.385 ± 0.015 | 1 | ±1 | ±1 | - |
| Weak | Z boson (\( Z^0 \)) | 91.1876 ± 0.0021 | 1 | 0 | 0 | - |

Table 2.2: Gauge bosons in the SM. Values are taken from ref. [11].

**Strong force** The strong force is described by Quantum Chromodynamics (QCD). The strong force has a range that corresponds to the size of a nucleus (\( \mathcal{O}(10^{-15} \text{ m}) \)), and it is mediated by the exchange of gluons. From a phenomenological point of view, the strong force “glues” quarks and/or antiquarks to hadronic bound states.
The quantum number that is related to strong interactions is the color charge, which is conserved. Quarks exist in three different color charges, which are usually referred to as “red”, “green”, and “blue”. Instead, the total color charge of all observed hadrons is zero, i.e. hadrons form color-singlet states. A direct consequence of the color charge is that it extends the degrees of freedom for the quantum states of quarks. Therefore, also baryons ($qqq$ or $\bar{q}\bar{q}\bar{q}$) with same-flavored quarks and equal spins exist, e.g. the $\Delta^{++}$ (uuu) and $\Delta^{-}$ (ddd) resonances. The “mediators” of the strong force, the gluons, carry color charges that are combinations of color and anticolor [28]. In total, eight independent color states, i.e. combinations of color and anticolor, exist. Gluons that form color singlets have not been observed yet.

Furthermore, a self-coupling of gluons exists, i.e. gluons couple to gluons, due to the non-abelian $SU(3)$ structure of the QCD. An important consequence is that the coupling strength of the strong interactions ($\alpha_s$) has a characteristic energy dependence. The coupling strength is expressed as a function of the transferred four-momentum squared ($Q^2$). In particular, the strong coupling strength is independent of flavor and color of a parton $^1$.

The coupling strength (asymptotically) decreases at large energy scales and small separation distances between the interacting partons, but diverges at low energy scales and large separation distances. The peculiarity of strong interactions is that quarks become quasi-free particles at large energy scales, which is referred to as “asymptotic freedom”. At low energy scales, the coupling strength among quarks is so large that quarks always form color singlets, i.e. they are bound in hadrons and cannot be observed freely $^2$. This is referred to as “color confinement”.

This “running” of the strong coupling strength is described as (cf. [31])

$$\alpha_s(\mu_R^2) \propto \frac{1}{\ln \frac{\mu_R^2}{\Lambda_{QCD}^2}}. \quad (2.1)$$

Here, $\Lambda_{QCD} \approx 200$ MeV is a constant and $\mu_R$ is the renormalization scale. The strong coupling strength has been measured at various values of $Q^2$, its value at an energy scale equivalent to the Z-boson mass is determined to be $\alpha_s(m_Z^2) = 0.1184 \pm 0.0007$ [11].

At large energy scales $\mu_R \gg \Lambda_{QCD}$, i.e. $\alpha_s(\mu_R^2) \ll 1$, strong interactions are described by perturbative QCD. Instead, the color confinement at low energy scales is non-perturbative. The hadronization of partons into color singlets then is subject to phenomenological models (cf. also the discussion about event modeling in section 2.2.1).

Mathematically, the dynamics of quarks and gluons are described by the Lagrangian density of QCD, which is given by [33]

$$L_{QCD} = -\frac{1}{4} \sum_{A=1}^{8} G_{\mu\nu}^A G^{A\mu\nu} + \sum_{j=1}^{n_f} \bar{q}_j(iD_\mu \gamma^\mu - m_j)q_j. \quad (2.2)$$

Here, the Dirac matrices are referred to as $\gamma^\mu$ [38], and the quark fields of $n_f = 6$ quark flavors with mass $m_f$ are referred to as $q_j$. The quark fields are color triplets of fermionic fields $q = (q_r, q_b, q_g)^T$ [34].

The covariant derivative $D_\mu$ is given by [33]

$$D_\mu = \partial_\mu + i g s A^A G_{\mu}^A. \quad (2.3)$$

$^1$The term “parton” refers to quarks and gluons.
$^2$In particular, the lifetime of a top quark is so small that it decays before it can hadronize (cf. sec. 2.2).
The coupling strength of the strong interactions ($g_s$) is usually referred to as $\alpha_s = \frac{g_s^4}{4\pi}$. The eight massless gauge fields, i.e. the gluon fields, are denoted by $g_A^\mu$. The $T^A$ are the generators of the $SU(3)$ group, e.g. represented by the Gell-Mann matrices [28].

The gluon-field-strength tensor is given by

\[ G^A_{\mu\nu} = \partial_\mu G^A_\nu - \partial_\nu G^A_\mu - g_s f_{ABC} G^B_\mu G^C_\nu, \tag{2.4} \]

in which the $f_{ABC}$ are the antisymmetric structure constants [28]. In particular, the terms $G^A_{\mu\nu} G^A_{\mu\nu}$ involve gluon-self couplings with three-gluon and four-gluon vertices [34].

**Electroweak interactions** The electromagnetic force explains the interactions among electrically charged particles and includes phenomena like the emission of light from excited atoms. The electromagnetic force is mathematically described by Quantum Electrodynamics (QED). It is mediated by massless photons and has an infinite range. Since photons are electrically neutral, no photon-self couplings exist.

The weak force has a range of $O(10^{-18} \text{ m})$ [33] and is mediated by $W^\pm$ bosons and $Z^0$ bosons. The beta decay of atomic nuclei probably is the most prominent example for a process that is mediated by a weak interaction.

The unified description of electromagnetic and weak interaction is described by electroweak theory based on the $SU(2) \otimes U(1)$ symmetry groups. Two quantum numbers are related to the electroweak interactions. The weak isospin $T$ corresponds to the $SU(2)$ symmetry group. The weak hypercharge $Y$ corresponds to the $U(1)$ symmetry group. Both quantum numbers are related by the electric charge $Q = T_3 + Y/2$ [34].

The electroweak theory is constructed such that it distinguishes between “left-handed” and “right-handed” fermions. Left-handed fermions (and right-handed antifermions) correspond to weak-isospin doublets ($T = 1/2$, third component $T_3 = \pm 1/2$). Right-handed fermions (and left-handed antifermions) correspond to weak-isospin singlets ($T = 0$) [34].

\[
\begin{pmatrix}
\nu_e \\
e_L
\end{pmatrix}, \quad
\begin{pmatrix}
\nu_\mu \\
\mu_L \\
\nu_\tau \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
\epsilon_R, \mu_R, \tau_R
\end{pmatrix}, \quad
\begin{pmatrix}
u_L \\
e_L \\
\mu_L \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_\mu \\
\mu_L \\
\nu_\tau \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_R \\
\epsilon_R, \mu_R, \tau_R
\end{pmatrix}, \quad
\begin{pmatrix}
u_e \\
e_L \\
\mu_L \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_\mu \\
\mu_L \\
\nu_\tau \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_R \\
\epsilon_R, \mu_R, \tau_R
\end{pmatrix}, \quad
\begin{pmatrix}
u_e \\
e_L \\
\mu_L \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_\mu \\
\mu_L \\
\nu_\tau \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_R \\
\epsilon_R, \mu_R, \tau_R
\end{pmatrix}, \quad
\begin{pmatrix}
u_e \\
e_L \\
\mu_L \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_\mu \\
\mu_L \\
\nu_\tau \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_R \\
\epsilon_R, \mu_R, \tau_R
\end{pmatrix}, \quad
\begin{pmatrix}
u_e \\
e_L \\
\mu_L \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_\mu \\
\mu_L \\
\nu_\tau \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_R \\
\epsilon_R, \mu_R, \tau_R
\end{pmatrix}, \quad
\begin{pmatrix}
u_e \\
e_L \\
\mu_L \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_\mu \\
\mu_L \\
\nu_\tau \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_R \\
\epsilon_R, \mu_R, \tau_R
\end{pmatrix}, \quad
\begin{pmatrix}
u_e \\
e_L \\
\mu_L \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_\mu \\
\mu_L \\
\nu_\tau \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_R \\
\epsilon_R, \mu_R, \tau_R
\end{pmatrix}, \quad
\begin{pmatrix}
u_e \\
e_L \\
\mu_L \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_\mu \\
\mu_L \\
\nu_\tau \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_R \\
\epsilon_R, \mu_R, \tau_R
\end{pmatrix}, \quad
\begin{pmatrix}
u_e \\
e_L \\
\mu_L \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_\mu \\
\mu_L \\
\nu_\tau \\
\tau_L
\end{pmatrix}, \quad
\begin{pmatrix}
u_R \\
\epsilon_R, \mu_R, \tau_R
\end{pmatrix}.
\tag{2.5} \]

In the following paragraphs, the left-handed $\psi_L$ and right-handed $\psi_R$ fermionic fields are referred to as [33]

\[ \psi_{L,R} = [(1 \mp \gamma^5)/2] \psi, \quad \bar{\psi}_{L,R} = \bar{\psi}[(1 \pm \gamma^5)/2], \tag{2.6} \]

in which $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \gamma_5$ [38].

The Lagrangian density of the electroweak interactions is separated into four parts [34]

\[ \mathcal{L}_{\text{EW}} = \mathcal{L}_{\text{Gauge fields}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Fermion fields}} + \mathcal{L}_{\text{Yukawa}}. \tag{2.7} \]

In the following paragraphs, the Lagrangian densities for the gauge fields ($\mathcal{L}_{\text{Gauge fields}}$), as well as for the fermionic fields and fermion-gauge interactions ($\mathcal{L}_{\text{Fermion fields}}$) are discussed. The Lagrangian densities for the Higgs mechanism ($\mathcal{L}_{\text{Higgs}}$) and Yukawa coupling ($\mathcal{L}_{\text{Yukawa}}$) are

---

7The left-handed fields form weak-isospin doublets $\psi^L_\alpha = (\psi_{L,T_3=+1/2}, \psi_{L,T_3=-1/2})^T$, while the right-handed fields form singlets $\psi_R \equiv \psi_{R,T=0}$ (cf. eq. 2.5). Defined by the eigenvalues ($\pm 1$) of the $\gamma^5$ operator, the left-handed component refers to “chirality” $-1$, and the right-handed component to “chirality” $+1$ (cf. [33]).
discussed in the next section (2.1.3). These explain why \( W^\pm \) bosons, \( Z^0 \) boson, and fermions are massive, but photons massless.

The Lagrangian density for the \( SU(2) \otimes U(1) \) gauge fields is given by \[ \mathcal{L}_{\text{Gauge fields}} = -\frac{1}{4} \sum_{A=1}^{3} W^A_{\mu\nu} W^{A\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \] (2.8)

The field-strength tensors \( B_{\mu\nu} \) and \( W^A_{\mu\nu} \) \((A = 1, 2, 3)\) are given by
\[
B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \\
W^A_{\mu\nu} = \partial_\mu W^A_\nu - \partial_\nu W^A_\mu - g \epsilon_{ABC} W^B_\mu W^C_\nu,
\]
in which the three vector fields \( W^A_\mu \) are related to the \( SU(2) \) symmetry group, and the singlet field \( B_\mu \) is related to the \( U(1) \) symmetry group. The symbol \( \epsilon_{ABC} \) refers to the Levi-Civita tensor [33]. The Lagrangian density involves cubic and quartic gauge couplings of the \( W^A_\mu \) field, similar to QCD. The coupling strength of the \( W^A_\mu \) fields is determined by \( g \). Here, the gauge fields are massless to preserve gauge invariance.

The Lagrangian density for the fermionic fields and fermion-gauge interactions is given by
\[
\mathcal{L}_{\text{Fermion fields}} = \sum_j \bar{\psi}_j^L i \gamma^\mu D^L_\mu \psi^L_j + \sum_k \bar{\psi}_k^R i \gamma^\mu D^R_\mu \psi^R_k.
\]
(2.10)

The first summation in eq. 2.10 takes the weak-isospin doublets of the three generations for both leptons and quarks into account, i.e. \( j = 1 \ldots 2n_g \). The index \( k \) runs over all weak-isospin singlets, i.e. three generations of up-type quarks, down-type quarks, and charged leptons.

The covariant derivatives are given by
\[
D^L,R_\mu = \partial_\mu + ig \sum_{A=1}^{3} t^A_{L,R} W^A_\mu + ig Y^R_2 B_\mu,
\]
(2.11)
in which \( t^A_{L,R} \) are the generators of the \( SU(2) \) symmetry group and \( Y^R_2 \) are the generators of the \( U(1) \) symmetry group. The generators \( t^A_L \) are given e.g. by \( t^A_L = \sigma_A/2 \) with the Pauli matrices \( \sigma_A \) [28]. Since the right-handed fermions form weak-isospin singlets, \( t^A_R \psi^R_\mu \) is zero in the SM (cf. [33]).

The Lagrangian density for the fermion fields (eq. 2.10) does not include fermion masses so far. However, fermions are observed to be massive. Furthermore, massless gauge fields have been assumed, but the W and Z bosons are quite heavy, while the photon is massless (table 2.2). The next section discusses the mass-generating mechanism.

### 2.1.3 Mass Generation

The following paragraphs briefly discuss the mechanism that introduces masses of \( W^\pm \) and \( Z^0 \) bosons while preserving local gauge invariance. This mass-generating mechanism is referred to as “Higgs mechanism” in the following. The second part of this section discusses

---

4The gauge-boson masses are introduced by interaction with the Higgs field.

5Various naming conventions, which also give credits to other contributing authors, are used in the literature (cf. e.g. [11] [28] [33] [39]).
the Yukawa coupling, which explains the masses of fermions by introducing an interaction between the Higgs and fermionic fields. Both the Higgs mechanism and the Yukawa couplings are discussed in detail in ref. [33, 34], which are summarized in the following paragraphs. Finally, the mixing between mass eigenstates and weak eigenstates of quarks is discussed. The CKM matrix, which relates both eigenstates, is introduced. The measurement of a particular element of the CKM matrix is the motivation for the measurement, which is described in the document at hand.

**Higgs mechanism** The Lagrangian density for the Higgs mechanism is given by [33]

\[ L_{\text{Higgs}} = (D_{\mu} \phi)\dagger (D^{\mu} \phi) - V(\phi\dagger \phi). \]  

(2.12)

Here, \( \phi \) is a weak-isospin doublet [34]

\[ \phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \]  

(2.13)

with weak hypercharge \( Y = 1 \). The fields \( \phi^+ \) and \( \phi^0 \) are complex scalar fields.

The covariant derivative introduces the couplings of the Higgs field to the \( W^A_\mu \) and \( B_\mu \) gauge fields and is given by [33]

\[ D_{\mu} = \partial_{\mu} + ig \sum_{A=1}^{3} t^A W^A_{\mu} + ig' Y B_{\mu}. \]  

(2.14)

The Higgs potential is given by [33, 34]

\[ V(\phi\dagger \phi) = -\mu^2 \phi\dagger \phi + \frac{\lambda}{2} (\phi\dagger \phi)^2, \]  

(2.15)

and defines the self couplings of the Higgs field. Here, the constants \( \mu \) and \( \lambda \) are introduced. All configurations that satisfy \( \phi\dagger \phi = \mu^2/\lambda \) minimize the potential, and the vacuum-expectation value (VEV) is given by [33]

\[ \langle 0 | \phi | 0 \rangle = \begin{pmatrix} 0 \\ v \end{pmatrix} \neq 0 \quad \text{with} \quad v = \sqrt{\frac{\mu^2}{\lambda}}. \]  

(2.16)

In particular, if both \( \mu \) and \( \lambda \) are larger than zero, the minimum of the potential is not \( \phi = 0 \). The symmetry of the Lagrangian density is “spontaneously broken” for non-zero VEVs under gauge transformations of the \( SU(2) \otimes U(1) \) group [34]. However, the Lagrangian density is symmetric under gauge transformations of the \( U(1)_{\text{em}} \) subgroup also for non-zero VEVs (cf. [34]).

The field \( \phi \) (eq. 2.13) is now expanded around the VEV and can be written as [33, 34]

\[ \phi = \begin{pmatrix} H + i\chi \\ v \end{pmatrix} \]  

(2.17)

in which the real part \( H \) defines the physical Higgs field. Gauge transformations can be used to fix the two \( \phi^+ \) fields \( \phi^+ \) and the field \( \chi \) to zero (cf. [34]). These scalar fields correspond to massless

\( ^6 \phi^+ \) includes two fields since it is complex.
“would-be” Goldstone bosons, which become the longitudinal modes of the (physical) $W^\pm$ and $Z^0$ bosons \[33, 34\], which are discussed in eq. 2.20 and eq. 2.21. Then, $\phi$ becomes $\phi = \begin{pmatrix} 0 \\ v + (H/\sqrt{2}) \end{pmatrix}$. \[2.18\]

The potential $V$ (eq. 2.15) becomes $V = -\frac{\mu^2 v^2}{2} + \mu^2 H^2 + \frac{\mu^2}{\sqrt{2}} v H^3 + \frac{\mu^2}{8v^2} H^4$, \[2.19\]
in which the self couplings (up to quartic vertices) are proportional to $m_H^2 = 2\mu^2$. A new scalar boson with mass $m_H = \sqrt{2}\mu$ is introduced. The mass of the Higgs boson is a free parameter in the theoretical description of the SM and has to be determined experimentally. Recently, a new boson with a mass of $(125.9 \pm 0.4)\text{ GeV}/c^2$ was observed at the LHC \[40, 41\], which could be the Higgs boson as predicted by the (minimal) SM (table 2.3).

<table>
<thead>
<tr>
<th>Name (Symbol)</th>
<th>Mass [GeV$/c^2$]</th>
<th>Spin [$\hbar$]</th>
<th>Electric charge [$e$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higgs mechanism</td>
<td>Higgs ($H^0$)</td>
<td>125.9 $\pm$ 0.4</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 2.3: Higgs boson in the SM. Values are taken from ref. \[11\]. The Higgs-boson-mass value corresponds to the mass of the recently observed boson \[40, 41\] even if it is not yet experimentally confirmed that this boson corresponds to the SM Higgs boson.

The Lagrangian density $\mathcal{L}_{\text{Higgs}}$ (eq. 2.12) involves couplings of the Higgs field $H$ to the gauge fields $W_\mu^3$ and $B_\mu$. The gauge fields ($W_\mu^1, W_\mu^2, W_\mu^3, B_\mu$) of the electroweak interaction can be combined into four “physical fields” (cf. \[33, 34\]). The combination of the gauge fields $W_\mu^1$ and $W_\mu^2$ to

$$W_\mu^\pm = \frac{W_\mu^1 \pm iW_\mu^2}{\sqrt{2}} \tag{2.20}$$

yield charged-current interactions (cf. \[33\]). The charged-current interactions are mediated by the $W^\pm$ bosons. The (mixed) fields $B_\mu$ and $W_\mu^3$ are transformed to their mass eigenstates (cf. \[33\])

$$A_\mu = \cos \theta_W B_\mu + \sin \theta_W W_\mu^3$$
$$Z_\mu = -\sin \theta_W B_\mu + \cos \theta_W W_\mu^3 \tag{2.21}$$
in which $\theta_W$ is referred to as the weak-mixing angle. The field $A_\mu$ gives rise to the massless photon, the field $Z_\mu$ gives rise to the massive $Z^0$ boson. These bosons mediate neutral-current interactions. The couplings between Higgs field and gauge fields (eq 2.12) then include $HW^\pm W^\mp$, $HZ^0 Z^0$, $HHW^\pm W^\mp$, and $HHZ^0 Z^0$ vertices \[34\].

By demanding that photons equally couple to left-handed and right-handed fermions, as well as demanding that photons couple to electrons with strength $e$, the two coupling strengths $g$ and $g'$ are given by \[33\]

$$g = \frac{e}{\sin \theta_W}$$
$$g' = \frac{e}{\cos \theta_W} \tag{2.22}$$

10
2.1 The Standard Model of Particle Physics

Both couplings are related via $\tan \theta_W = g'/g$. They can be described by one common coupling $e$. Thus, the electromagnetic and weak interactions are “unified”. The electromagnetic-coupling strength is usually expressed as the fine-structure constant $\alpha_{em} = e^2/(4\pi)$ [34]. The weak-mixing angle is given by $\theta_W$. Both parameters $\alpha_{em}$ and $\theta_W$ have to be determined experimentally [11].

The masses of the $W^\pm$ and $Z^0$ bosons can be related to the VEV of the Higgs field, while the photon remains massless [33],

\[ m_{W^\pm} = \frac{g v}{\sqrt{2}}, \quad m_{Z^0} = \frac{g v}{\sqrt{2} \cos \theta_W}, \quad \text{and} \quad m_\gamma = 0. \quad (2.23) \]

**Yukawa coupling** Fermion masses can be explained by interactions between the fermion fields and the scalar Higgs field. These interactions are introduced by the Yukawa coupling [33]

\[ \mathcal{L}_{\text{Yukawa}} = -\bar{\psi}_L \Gamma \psi_R \phi - \bar{\psi}_R \Gamma^\dagger \psi_L \phi^\dagger. \quad (2.24) \]

Here, the $\Gamma$ matrices include the couplings constants [33]. The fermionic fields $\psi_{L,R}$ are weak-isospin doublets as discussed in eq. 2.10.

When considering $\mathcal{L}_{\text{Yukawa}}$ at the VEV (eq. 2.18), the fermion-mass matrix $\mathcal{M}$ is obtained, and the Lagrangian density becomes (cf. [33, 34])

\[ \mathcal{L}_{\text{Yukawa}} = -\bar{\psi}_L \mathcal{M} \psi_R - \bar{\psi}_R \mathcal{M}^\dagger \psi_L - \bar{\psi}_L \frac{\mathcal{M}}{\sqrt{2} v} \psi_R H - \bar{\psi}_R \frac{\mathcal{M}^\dagger}{\sqrt{2} v} \psi_L H. \quad (2.25) \]

The fermion-mass matrix $\mathcal{M} = \Gamma \cdot \nu$ is generated by the Higgs field, and relates the fermion masses with the Yukawa couplings $\Gamma$ and the Higgs VEV $v$ (cf. [33]). If the fermions interact with the Higgs field, they become massive and they couple to the Higgs field with a strength that is proportional to their masses. Thus, in order to proof the (postulated) mass-generation mechanism for fermions, it is important to experimentally determine the coupling strengths between the fermionic fields and the Higgs field.

When applying the unitary transformations $\psi'_L = U_L \psi_L$ and $\psi'_R = U_R \psi_R$, the fermion-mass matrix can be diagonalized (cf. [33])

\[ \mathcal{M} \rightarrow \mathcal{M}' = U_L^\dagger \mathcal{M} U_R. \quad (2.26) \]

In total, four unitary matrices ($U^u_L$, $U^u_R$, $U^d_L$, and $U^d_R$) are obtained that operate on the fields for up-type ($u$) and down-type ($d$) quarks (cf. [33]). The matrices $U^u_L$ and $U^d_L$ operate on the left-handed weak-isospin doublets ($\psi_L$). In consequence, the Yukawa couplings introduce weak-charged-current interactions due fermion-gauge interactions (eq. 2.10). The couplings of these weak-charged currents ($W^\pm$) among quarks are expressed in terms of the CKM matrix [33]

\[ V_{\text{CKM}} = (U^u_L)^\dagger U^d_L, \quad (2.27) \]

which is discussed in the next paragraphs.
CKM matrix  The weak-charged-current interactions alter the flavor of quarks by $W^\pm$ emission. However, the mass eigenstates $D$ and weak eigenstates $D'$ of quarks do not coincide (cf. [33]). The two eigenstates are related by the unitary CKM matrix $V_{\text{CKM}}$ [28]

$$
\begin{pmatrix}
  d' \\
  s' \\
  b'
\end{pmatrix}
= V_{\text{CKM}}
\begin{pmatrix}
  d \\
  s \\
  b
\end{pmatrix}.
$$

The squared absolute elements of the CKM matrix $|V_{ij}|^2$ determine the transition probability of $q_i \rightarrow q'_j + W^+$ (resp. $q'_j \rightarrow q_i + W^-$), in which $q_i$ refers to an up-type quark of flavor $i$ and $q'_j$ refers to a down-type quark of flavor $j$ (cf. [28, 33]).

The CKM matrix has nine (complex) parameters, which can be reduced to a set of three independent mixing angles ($\theta_{12}, \theta_{23}, \theta_{13}$) and one CP-violating phase factor $\delta$ (cf. [11, 28]). The mixing angles $\theta_{ij}$, which determine the mixing among the quark generations $i$ and $j$, and the (nonzero) phase $\delta$ are fundamental parameters of the SM, which have to be determined experimentally. All individual elements $|V_{ij}|$ of the CKM matrix (eq. 2.28) are measured independently.

The peculiarity of the measurement, which is described in the document at hand, is the following. The measurement of the single-top-quark-production cross section allows $|V_{tb}|$ to be determined without assuming unitarity of $V_{\text{CKM}}$ and without assuming three quark generations. Hence, the result of this measurement is a crucial input to the description of the SM (cf. sec. 2.2.3 for more details).

A global fit, which assumes unitarity of the $V_{\text{CKM}}$ with three quark generations, is used to determine $\theta_{ij}$ and $\delta$, as well as to further constrain the individual elements $|V_{ij}|$. The CKM matrix elements are determined to be [11]

$$
|V_{ij}| =
\begin{pmatrix}
  0.97427 & 0.22534 & 0.00351 \\
  0.22520 & 0.97344 & 0.0412 \\
  0.00867 & 0.0404 & 0.999146
\end{pmatrix}.
$$

In particular, the global fit includes results from a previous iteration of this analysis [21].

2.2 The Top Quark

The top quark is the fermion with spin $\frac{1}{2}$ and weak isospin $T_3 = 1/2$ that belongs to a weak-isospin doublet with the bottom quark in the SM of particle physics (cf. [11, 15]). The top quark has an electric charge of $Q_t = +2/3$ e. A recent direct measurement in $t\bar{t}$ events yields $Q_t = (0.64 \pm 0.02 \pm 0.08)$ e [42]. Indirect measurements disfavor an exotic charge of $|Q_t| = 4/3$ e (cf. [43]).

Direct measurements of the top-quark mass are obtained from kinematics in $t\bar{t}$ events and yield $m_t = (173.20 \pm 0.87)$ GeV/c$^2$ [35]. The top-quark mass is most precisely measured in $p\bar{p}$ collisions. These measurements are confirmed by recent measurements in $pp$ collisions at the LHC, which yield $m_t = (173.29 \pm 0.95)$ GeV/c$^2$ [44].

The top-quark-pole mass can be indirectly inferred from measurements of the $t\bar{t}$ production cross section, since top-quark mass and $t\bar{t}$ production cross section are related with each

$^7$Unitarity is given by $(V_{\text{CKM}} V_{\text{CKM}}^\dagger)_{ij} = (V_{\text{CKM}}^\dagger V_{\text{CKM}})_{ij} = \delta_{ij}$ [11].
2.2 The Top Quark

The top-quark-pole mass is determined to be $m_{t\text{pole}} = 176.7^{+3.8}_{-3.4} \text{GeV}/c^2$ using the NNPDF2.3 Parton-Distribution Function (PDF) set and Next-To-Next-To-Leading Order (NNLO)\textsuperscript{[15]} Next-To-Next-To-Leading-Logarithmic Order (NNLL) \textit{tt} cross-section prediction \textsuperscript{[45]} . Measurements of the top-quark mass in the MS-renormalization scheme give $m_{t\text{MS}} = 160.0^{+4.8}_{-4.4} \text{GeV}/c^2$ with the MSTW08 PDF set and approximate NNLO \textit{tt} cross section prediction (cf. \textsuperscript{[46]} for further details, as well as \textsuperscript{[15, 43]} ).

The top-quark width is predicted to be $\Gamma_t \approx 1.3 \text{GeV}$ \textsuperscript{[47]} , which corresponds to a lifetime of about $\tau_t \approx 0.5 \times 10^{-24} \text{s}$ \textsuperscript{[11, 15]} . The lifetime of the top quark is rather short \textsuperscript{[11, 15]} . In particular, it is smaller than the typical QCD hadronization scale, which is $\approx 3.0 \times 10^{-24} \text{s}$ \textsuperscript{[15]} . Top quarks decay before they are bound in top-flavored hadrons, which offers the unique opportunity to study a “bare” quark through the properties of its decay products. As an example, the top quark is not depolarized and information about the top-quark spin is encoded in typical angular correlations of its decay products (cf. ref. \textsuperscript{[3, 4, 15, 43]} for more details).

This section discusses the top-quark production at hadron colliders. It is divided into four parts. The first part of this section (2.2.1) briefly discusses the modeling of hadron collisions.

2.2.1 Modeling of Hadron Collisions

The squared center-of-mass energy of the hadron collisions at the \textbf{LHC} is given by $s = 4E_AE_B = 4E_{\text{beam}}^2$ . In the context of this analysis, the two colliding hadrons $A$ and $B$ are protons ($A \equiv B = p$).

Protons are composite objects, whose constituents are referred to as “partons”. Protons are composed of three valence quarks (two up quarks and one down quark) and a parton sea, which consists of gluons and quark-antiquark pairs. Hence, not the protons themselves but their partons interact in the hard-scattering process.

Partons carry different fractions $x_i$ of the total hadron momentum ($i = u, d, c, s, b, g$). Their probability-density distributions are described by the proton
Parton-Distribution Functions (PDFs) $f_{i\bar{A}}(x_i, Q^2)$, which are parametrized in momentum fractions $x$ of the hadron and a more general energy scale $Q$, at which the PDFs are evaluated. In particular, the PDFs are independent of the hard-scattering process.

Parton-distribution function As an example, figure 2.1 shows the distribution $x \times f(x, Q^2)$ for the CT10 Next-To-Leading Order (NLO) PDFs at $Q^2 = (173.2 \text{ GeV})^2$ [51]. Figure 2.1 (left) shows the distributions for up and down valence quarks (“upv” and “downv”) in the proton, as well as for the total contributions of up- and down quarks. The valence quarks carry most of the high-fraction momenta, while up- and down quarks from the proton sea carry mostly lower fractions of momenta. Contributions from up-valence quarks and down-valence quarks roughly are two to one at large $x$.

Figure 2.1 (right) shows the distributions for bottom quarks and gluons, and the distribution for up quarks is shown as a reference. Figure 2.1 (right) has a different scale of the $y$ axis. The gluon contributions are by far dominant at low $x$, while the valence quarks are preferred when carrying large fractions of the hadron momentum. At low $x$, the contributions due to up quarks are preferred over that from bottom (anti)quarks.

![Figure 2.1: Distributions of $x \times f(x, Q^2)$ for the CT10 NLO PDFs at $Q^2 = (173.2 \text{ GeV})^2$ [51]. The left figure shows the distributions for total contributions of up- and down quarks, as well as for the valence quarks (“upv” and “downv”) separately. The right figure shows the distributions for up quarks, bottom quarks, and gluons. The $y$-axis has different scales for the left and right figures.](image)

The PDFs are derived from global fits to the structure functions of hadrons [11], which sum the weighted contributions of all individual PDFs. The structure functions are measured in deep-inelastic-scattering processes, e.g. in neutral- and charged-current processes with the Hadron-Elektron-Ring-Anlage (HERA) $ep$ collider or fixed-target experiments with proton, neutron, or deuteron targets. Fixed target experiments with neutrino-nucleon-scattering processes provide further information. Moreover, the rapidity distributions of the Z boson, rapidity distributions of charged leptons from W-boson production, and jet-cross-section measurements are utilized in hadron collisions. More details are given e.g. in ref. [11] [51] [52]. In particular, these experiments are sensitive to different PDFs and cover different kinematic domains in the $x$–$Q^2$ plane. Experimental data from the LHC is expected to further tighten experimental constraints on the PDFs [11].
Several groups extract the PDFs. The derived PDFs differ in their parametrization, the analyzed experimental data, the used (central) value for the strong-coupling constant $\alpha_s(m_Z)$, the assumed bottom- and charm-quark mass, as well as in the choice of the QCD-evolution equation \[11, 53\]. The fitting groups provide different prescriptions to calculate an uncertainty of the parametrization of a particular $x \times f(x_i, Q^2)$ distribution. An overview of these prescriptions is given e.g. in ref. \[54\]. The resulting uncertainties usually are of the order of a few percent. However, as the assumptions to determine the PDFs itself differ among the fitting groups, it is recommended to carefully consider PDFs of different groups for uncertainty estimation, e.g. for a cross-section measurement of a physics process \[55\].

![Figure 2.2: Factorization of the hard-scattering process at hadron colliders (cf. \[56\).](image)

**Hard-scattering process** Figure 2.2 illustrates the hard-scattering process. The total hadronic cross section, e.g. for $t\bar{t}$ production $\sigma(AB \to t\bar{t})$, of the two colliding hadrons $A$ and $B$ is given by the convolution of the partonic cross section $\hat{\sigma}$ with the PDFs $f(x, Q^2)$. This is referred to as the QCD factorization theorem \[4, 43, 56\].

$$\sigma(AB \to t\bar{t}, s, m_t) = \sum_{i,j=q,\bar{q},g} \int d\hat{s} f_i|_A(x_i, \mu_F^2) f_j|_B(x_j, \mu_F^2) \times \hat{\sigma}(ij \to t\bar{t}; \hat{s}, \mu_R, \alpha_s). \quad (2.30)$$

The partonic center-of-mass energy $\hat{s}$ depends on the momentum of the initial partons ($x_i$ and $x_j$) and is given by $\hat{s} = x_i x_j s$.

Two scales $\mu_F$ and $\mu_R$ are introduced in eq. 2.30. These can be understood as follows (cf. \[4, 43, 56\]). When calculating the total hadronic cross section $\sigma$ without considering any scale $\mu_R$ and $\mu_F$, large logarithms arise if gluon emissions collinear to the incoming partons are calculated. The factorization scale $\mu_F$ determines the scale up to which these collinear divergences are absorbed in the definition of the PDFs. The factorization scale $\mu_F$ “separates long- and short-distance physics” \[56\]. $\alpha_s$ refers to the (running) strong coupling, and $\mu_R$ is the renormalization scale for $\alpha_s$.

Finite corrections in the perturbative calculations remain. These are compensated by additional corrections of $O(\alpha_s^2)$ to the leading logarithm \[56\]

$$\hat{\sigma}(ij \to t\bar{t}; \hat{s}, \mu_R^2, \alpha_s) = \left[ \hat{\sigma}^{(0)} + \alpha_s(\mu_R^2)\hat{\sigma}^{(1)} + \alpha_s^2(\mu_R^2)\hat{\sigma}^{(2)} + \ldots \right]_{ij \to t\bar{t}}. \quad (2.31)$$
Typically, both renormalization and factorization scale are set to a common scale \( Q \equiv \mu_F = \mu_R \), which corresponds to the momentum scale of the hard-scattering process, e.g. \( Q = m_t \) in case of top-quark-pair production.

The total hadronic cross section depends on both \( \mu_F \) and \( \mu_R \). The resulting variations of the total hadronic cross section due to variations of \( \mu_F \) or \( \mu_R \) become smaller the more corrections are included into the calculation. If the complete perturbative series is known, the dependence on \( \mu_F \) and \( \mu_R \) vanishes. With finite order of the perturbative series, the scale dependence of the cross section \( \sigma \) is taken into account as an additional uncertainty that is estimated by varying \( \mu_F \) and \( \mu_R \). Typical choices to estimate the uncertainty due to the choice of the scale \( Q \) are (simultaneous) variations of both \( \mu_F \) and \( \mu_R \) by factors 2 and 1/2.

The calculations of the partonic cross sections (eq. 2.31) are known up to a fixed order of perturbative QCD and depend on the particular processes, e.g. NNLO for top-quark-pair production (cf. sec. 2.2.2). In particular, these calculations become rather complex the higher the accuracy in perturbative QCD is. While the inclusive cross section of a particular process is calculated with high precision, distributions of differential cross-section calculations are usually available only for a limited number of observables.

**Event modeling and Simulation of Hadron Collisions** High-energy experiments rely on a large number of observables. Thus, a simulation of particle collisions is used. Simulations provide a “generic” output in terms of high-energy physics collisions, e.g. a list of all generated particles with id, four vector, spin, parent-child history, ... [57]. The simulation output can be used to calculate any observable offline. The simulation of high-energy particle collisions is factorized into several steps (fig. 2.3), which are briefly discussed in the following paragraphs.

![Diagram of hadron collision simulation](image)

**Figure 2.3:** Illustration of the simulation of hadron collisions, which is factorized into several steps [58]. The time axis is given by the vertical axis.

Matrix-element event generators are used to sample configurations of final-state particles over the full kinematic phase space according to the matrix element of a particular process. The matrix element defines the transition from the “initial state” to the “final state”. These configurations are weighted with the PDFs. A detailed description of the matrix-element generators and parton-shower simulation that are used to generate single-top-quark-\( t \)-channel and...
background processes as used in this analysis is given in section 3.5. Typical matrix-element generators, which are used in this analysis, are MADGRAPH [59] for Leading order (LO) accuracy and POWHEG BOX [60–62] for NLO accuracy.

The output of the matrix-element event generators are particles at a rather high energy scale. In a next step, parton-shower programs are used to evolve the “final-state” partons down to a low energy scale, which is of \( O(\Lambda_{\text{QCD}}) \) [56]. The parton showering effectively approximates all higher-order corrections that hardly can be calculated in perturbative QCD, mainly due to large logarithmic corrections that arise due to soft and collinear emissions (cf. [11]). The evolution of partons involves gluon radiation and subsequent splitting of these gluons into pairs of gluons \((gg)\) or quarks \((q\bar{q})\). Subsequent gluon splitting and gluon radiation generate entire parton cascades until a certain cut-off scale is reached. Moreover, the parton-shower models account for initial- and final-state gluon radiation.

All individual, colored partons are “combined” into color-singlet hadrons, which are observed in an experiment. This process is referred to as “hadronization” and described by phenomenological models, e.g. the “Lund string model” or “cluster model” (details are given e.g. in ref. [11]). Subsequently, all hadrons are decayed until only stable hadrons remain.

Typical programs used for parton-shower simulation and hadronization are PYTHIA [63] or HERWIG [64]. Additional libraries can be used to correct for various simplifications, e.g. for QED, radiative corrections to particle decays. An example related to top-quark physics is the use of the TAUOLA library, which simulates \( \tau \)-lepton decays [65].

Parton-shower models are mainly suited to describe soft and collinear emissions, while matrix-element generators can be used to simulate hard and wide-angle emissions (cf. also [11]). Some event generators are able to perform exact matrix-element calculations in LO accuracy with up to \( n \) partons in addition to the “elementary” process, e.g. MADGRAPH [59]. These generators are also referred to as “multi-leg generators”.

Additional partons are generated at tree level beyond a certain cut-off scale, which is also referred to as “matching threshold”. A combination of parton-shower modeling and tree-level calculations of matrix-element generators improves the description of additional partons. The conceptual challenge is to avoid a double- or under counting of phase space between the matrix-element calculation and parton-shower simulation. Several schemes were developed to handle such a “matching” of parton-shower and matrix-element generators [66]. In this analysis, the MLM-matching prescription with \( k_T \)-jets is used [66]. Furthermore, dedicated frameworks were developed to match NLO calculations with parton-shower simulations, including the POWHEG BOX [60–62] and MC@NLO [67] formalisms.

**Underlying event** A peculiarity of hadron collisions is that energy in addition to that from the primary interaction is found experimentally (“minimum bias collisions” in fig. 2.3). This additional activity in an event is referred to as “underlying event”, and is supposed to stem from the color-connected constituents of the hadron remnants. Moreover, in the context of top-quark physics, the decay products of the (colored) top quark, i.e. the \( b \) quark, are color-connected with the proton remnants and other jets in the event. In general, both effects are small\(^8\).

\(^8\) However, these effects become significant (up to 0.5 GeV/c\(^2\)) for a determination of the top-quark mass at a precision of \( \approx 1 \) GeV/c\(^2\) [44, 68].
The underlying event can be described by a combination of phenomenological models for soft interactions with perturbative QCD (cf. [11, 69]), e.g. multiple parton interactions [69]. Typically, these models have several degrees of freedom. The underlying event is experimentally characterized with a few observable, e.g. with charged-particle production [69]. A parameter set for a particular model is also referred to as “tune”. The PYTHIA 6 tune Z2 [70] is used for underlying-event modeling in this analysis.

In addition to the primary interaction, other interactions may also occur in parallel. These additional interactions mainly are soft. They are referred to as Pile-Up interactions. The occurrence of Pile-Up interactions is a statistical process that linearly scales with the total inelastic proton-proton cross section. Pile-Up is simulated with parton-shower programs and discussed in detail in section 3.5.4.

Recent focus of the LHC experiments was not only on the development of a dedicated tune to adjust the underlying-event models to the experimental data. In order to mitigate contributions due to Pile-Up, the CMS experiment utilizes its Particle Flow (PF) algorithm to subtract charged particles that are not compatible with the primary vertex in an event (cf. sec. 3.3.2). The CMS experiment also uses algorithms to estimate the underlying-event activity on an event-per-event basis and to correct individual jets for it (cf. sec. 3.3.5 and ref. [71]). First studies arrived that characterize the underlying event in the context of top-quark physics to improve the understanding of color-reconnection effects as well as their simulation [72].

2.2.2 Top-Quark-Pair Production

![Feynman diagrams at leading-order QCD accuracy for top-quark-pair production](image)

Figure 2.4: Feynman diagrams at leading-order QCD accuracy for top-quark-pair production (cf. e.g. [3, 4, 73]). The left figure (2.4a) refers to top-quark-pair production by quark-antiquark annihilation ($q\bar{q}$-channel). The other three figures (2.4b) describe the top-quark-pair production by gluon-gluon fusion ($g\bar{g}$, $t\bar{t}$, and $u\bar{u}$-channels).

At hadron colliders, pairs of a top-quark and a top-antiquark are produced via strong interactions. Figure 2.4 shows the LO QCD Feynman diagrams for $t\bar{t}$ production, which consist of either quark-antiquark ($q\bar{q}$ → $t\bar{t}$) (fig. 2.4a) or gluon-gluon initial states ($gg$ → $t\bar{t}$) (fig. 2.4b). In higher orders of perturbative QCD also $qg$ initial states contribute to the production cross section.

Today, the fixed-order-partonic cross section calculation for $t\bar{t}$ production is known up to NNLO accuracy ($O(\alpha_s^4)$) and corrections due to emissions of soft gluons are included up to NNLL accuracy [74]. The top-quark-pair-production cross section in $pp$ collisions at $\sqrt{s} = 7$ TeV is predicted to be

$$\sigma_{t\bar{t}} = (172.0^{+4.4+4.7}_{-5.8-4.8}) \text{ pb} \quad (2.32)$$
The Top Quark

in NNLO+NNLL accuracy [74]. The first uncertainty refers to uncertainties due to the choice of the renormalization and factorization scales. The second uncertainty refers to the uncertainty due to the parametrization of the PDFs. For this calculation, a top-quark mass of 173.3 GeV/c² and the MSTW2008 NNLO PDF set at 68% CL [75] are used.

Table 2.4 summarizes the cross-section predictions for top-quark-pair production. Calculations of different groups, as well as different orders of accuracy in perturbative QCD are quoted. The approximate NNLO calculations differ in the methods used to re-sum corrections from soft gluon emissions (more details are given e.g. in [43, 73]). All predictions are compatible with each other within uncertainties. The variations that result from the choice of the top-quark mass are smaller than uncertainties due to scale or pdf variations.

<table>
<thead>
<tr>
<th>Accuracy</th>
<th>Cross section [pb]</th>
<th>PDF set, m_t [GeV/c²]</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NLO</td>
<td>158^{+18}_{-19}</td>
<td>CTEQ6M NLO (68% CL)</td>
<td>[52], 172.5</td>
</tr>
<tr>
<td>approx. NNLO</td>
<td>155^{+8}_{-9}</td>
<td>MSTW2008 NNLO (90% CL)</td>
<td>[73], 173.1</td>
</tr>
<tr>
<td></td>
<td>163^{+7}_{-8}</td>
<td>MSTW2008 NNLO (90% CL)</td>
<td>[73], 173.0</td>
</tr>
<tr>
<td></td>
<td>163^{+7+15}_{-8-15}</td>
<td>MSTW2008 NNLO (90% CL)</td>
<td>[73], 173.4</td>
</tr>
<tr>
<td></td>
<td>175^{+10+5}_{-13-5}</td>
<td>MSTW2008 NNLO (68% CL)</td>
<td>[73], 173.0</td>
</tr>
<tr>
<td>NNLO+NNLL</td>
<td>172^{+4+5}_{-6-5}</td>
<td>MSTW2008 NNLO (68% CL)</td>
<td>[73], 173.3</td>
</tr>
</tbody>
</table>

Table 2.4: Summary of cross-section predictions for top-quark-pair production. The first uncertainty is due to scale variations, and the second uncertainty due to the PDF parametrization (PDF⊕α_s for ref. [80]).

In pp collisions at \( \sqrt{s} = 7 \) TeV, the top-quark-pair production cross section for quark-antiquark initial states (\( q\bar{q} \rightarrow t\bar{t} \)) (fig. 2.4a) contributes to about 16% to the total \( t\bar{t} \) cross section. The dominant production mechanism is through gluon-gluon fusion (gg \( \rightarrow t\bar{t} \)) with about 84% (fig. 2.4b). Both values are calculated with top++ v2.0 [74, 82].

The top quark was observed in \( t\bar{t} \) production in pp collisions at a center-of-mass energy of \( \sqrt{s} = 1.8 \) TeV at the Tevatron in 1995 [1, 2]. \( t\bar{t} \) production was rediscovered in pp collisions at a center-of-mass energy of \( \sqrt{s} = 7 \) TeV shortly after the start of the LHC.

Recent measurements of the \( t\bar{t} \) cross section are so accurate that they challenge the theoretical predictions. A comparison of measurements of the \( t\bar{t} \) production cross section in pp and pp collisions to the theoretical prediction in NNLO+NNLL QCD accuracy is given in figure 2.5. The theoretical prediction is calculated with top++ v2.0 [74, 82]. A top-quark mass of 173.3 GeV/c² and the MSTW2008 NNLO PDF set at 68% CL [75] are used (cf. eq. 2.32).

An important feature of top-quark-pair production is its charge asymmetry. Measurements of the charge asymmetry at the Tevatron showed deviations with the SM expectation that might hint to new physics phenomena. This topic is not discussed in this document, but is discussed in detail e.g. in ref. [5, 43].
Figure 2.5: Measurements [83-87] of the \( t\bar{t} \) production cross section in \( pp \) and \( pp \) collisions. The theoretical predictions are calculated with \( \text{top++ v2.0 [74, 82]} \) in \( \text{NNLO+NNLL QCD} \) accuracy.
2.2.3 Single-Top-Quark Production

Single top quarks are produced via electroweak interactions. Hence, single-top-quark production is often also referred to as “electroweak top-quark production”.

This section is organized as follows. First, the production mechanisms of single top quarks are discussed. Then, experimental evidence of single-top-quark production is presented. In the next paragraphs, the motivation for a measurement of single-top-quark production is detailed. Single-top-quark production offers the possibility to determine both the “strength and structure” of the electroweak interactions at the \( Wtb \) vertex (cf. \[15\]). The coupling strength at the \( Wtb \) vertex, i.e. the CKM-matrix element \(|V_{tb}|\), can be directly inferred from the cross-section measurement without assuming CKM-matrix unitarity. An overview of the current \(|V_{tb}|\) measurements is given. Moreover, single-top-quark production enables the determination of the structure of the electroweak interactions in terms of anomalous couplings. Finally, physics phenomena beyond the SM are briefly discussed.

Production mechanisms In total, three production modes are distinguished by means of the virtuality \( Q^2 = -q^2 \) of the exchanged W boson, whose four-momentum is referred to as \( q \) in the following (cf. \[4, 48\]). All processes also include charge conjugates, which are not explicitly stated here.

- The \( t \)-channel production mechanism (fig. 2.6) involves a space-like W boson (\( q^2 \leq 0 \)). If the b quark is considered massive within the matrix-element calculation and not part of the proton PDFs, the initial state is \( qg \) and the final state is \( qt\bar{b} \) (“4-flavor scheme”). If the initial-b quark is assumed to be massless and part of the proton PDFs, the initial state is \( qb \) and the final state is \( q't \) (“5-flavor scheme”). A detailed description of the modeling of the \( t \)-channel process with event generators is given in section 3.5.

![Feynman diagrams at leading-order QCD accuracy](image)

Figure 2.6: Feynman diagrams at leading-order QCD accuracy for \( t \)-channel single-top-quark production in the 4-flavor scheme (left) and 5-flavor scheme (right).

- The \( s \)-channel production mechanism (fig. 2.7a) involves a time-like W boson (\( q^2 \geq (m_t + m_b)^2 \). The initial state contains an up-type quark and a down-type antiquark (\( q\bar{q}' \)). The final state is \( t\bar{b} \).

- In the \( tW \)-channel production mechanism (fig. 2.7b), a real W boson (\( q^2 = m_W^2 \)) is produced in association with a top quark (\( tW \)). The initial state of the \( tW \)-channel (fig. 2.7b) includes a gluon and a b quark (\( gb \)). NLO corrections to the \( tW \)-channel result in an interference between \( tW \)-channel and \( tt \) production \[88, 89\]. Two definitions exist to define the NLO corrections, namely diagram removal and diagram subtraction.
2 The Top Quark within the Standard Model of Particle Physics

Figure 2.7: Feynman diagrams at leading-order QCD accuracy for s-channel (fig. 2.7a) and tW-channel (fig. 2.7b) single-top-quark production.

An overview of the predicted production cross sections at the LHC at $\sqrt{s} = 7$ TeV is given in table 2.5. The dominant production mechanism of single top quarks in pp collisions at $\sqrt{s} = 7$ TeV is via the t-channel with a total cross section of about 64.6 pb. The tW-channel is the second-dominant production mechanism with a total cross section of about 15.7 pb. Instead, the s-channel production with its $q\bar{q}'$-initial state has a relatively small cross section of 4.6 pb. At the LHC pp collider, the production of top quarks is preferred to top-antiquark production in the s- and t-channel mechanisms due to their initial states ($q\bar{q}'$ and $qg$) and the valence-quark content of the proton (uud). Instead, the production in the tW-channel with its $gb$-initial state is charge-symmetric.

<table>
<thead>
<tr>
<th>Process</th>
<th>Cross section [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t$-channel ($t$)</td>
<td>$41.92 \pm 1.59 \pm 0.83$ (approx. NNLO [8])</td>
</tr>
<tr>
<td>$t$-channel ($\bar{t}$)</td>
<td>$22.65 \pm 0.50 \pm 0.68$ (approx. NNLO [8])</td>
</tr>
<tr>
<td>$s$-channel ($t$)</td>
<td>$3.19 \pm 0.06 \pm 0.13$ (approx. NNLO [9])</td>
</tr>
<tr>
<td>$s$-channel ($\bar{t}$)</td>
<td>$1.44 \pm 0.01 \pm 0.06$ (approx. NNLO [9])</td>
</tr>
<tr>
<td>$tW$-channel ($t$, Diagram Removal)</td>
<td>$7.87 \pm 0.20 \pm 0.55$ (approx. NNLO [10])</td>
</tr>
<tr>
<td>$tW$-channel ($\bar{t}$, Diagram Removal)</td>
<td>$7.87 \pm 0.20 \pm 0.55$ (approx. NNLO [10])</td>
</tr>
</tbody>
</table>

Table 2.5: Summary of production cross sections for electroweak top-quark production in pp collisions at $\sqrt{s} = 7$ TeV. The first uncertainty is due to scale variations, and the second uncertainty due to the PDF parametrization. The MSTW2008 NNLO PDF set at 90% CL [75] is used.

Experimental evidence Single-top-quark production was observed in pp collisions at the Tevatron in 2009 [6, 7]. The dominant production at the Tevatron in pp collisions at $\sqrt{s} = 1.96$ TeV were due to both s-channel and t-channel production, while tW-channel production was negligible due to the $gb$ initial state and its “massive particles in the final state” [90].

The observation of (individual) s-channel production was recently reported by the Tevatron experiments [91]. Measurements performed by the CDF and D0 experiments in pp collisions at $\sqrt{s} = 1.96$ TeV were combined. The relative uncertainty of the s-channel-production cross section is 19%. A direct search for s-channel production in pp collisions at $\sqrt{s} = 7$ TeV yielded
an upper limit of 5.7 times the SM prediction \((95\% \text{ CL})\) \cite{92}. In \(pp\) collisions at \(\sqrt{s} = 8\) TeV, an upper limit of 2.1 times the SM prediction \((95\% \text{ CL})\) was measured \cite{93}.

Evidence for the associated production of a single top quark and a W boson \((tW\)-channel) was reported for \(pp\) collisions at \(\sqrt{s} = 7\) TeV \cite{94,95} and \(\sqrt{s} = 8\) TeV \cite{96}. Finally, single-top-quark production in the \(tW\)-channel was observed in \(pp\) collisions at \(\sqrt{s} = 8\) TeV \cite{97}. Here, the cross-section measurement has a relative uncertainty of 24%.

A previous iteration of the analysis at hand was published in \cite{21}. Based on a much smaller dataset, which corresponds to an integrated luminosity of 36 pb\(^{-1}\), the inclusive \(t\)-channel cross section was measured with a relative uncertainty of 36%. The analysis that is described in the document at hand is referred to as “this analysis” in the following paragraphs.

The results of this inclusive \(t\)-channel cross-section measurement are based on \(pp\) collisions at \(\sqrt{s} = 7\) TeV corresponding to an integrated luminosity of 1.6 fb\(^{-1}\) and have been pre-published in \cite{24}. The results from this measurement were combined with two other measurements. One measurement is based on a fit to a neural-network discriminator \cite{25}. The other measurement is based on a fit to the pseudo-rapidity of the forward jet (ref. \cite{26} used events with electrons and ref. \cite{27} used events with muons). The inclusive \(t\)-channel-cross-section measurement based on ref. \cite{25} has an (expected) precision that is similar to the precision that is obtained in this analysis. The measurements from ref. \cite{26,27}, as well as their combination, are less precise.

Before the results of this analysis were published in \cite{24}, the most precise \(t\)-channel-cross-section measurement in \(pp\) collisions at \(\sqrt{s} = 7\) TeV had a relative uncertainty of 23\% \cite{22}. Here, data corresponding to an integrated luminosity of 1.04 fb\(^{-1}\) were analyzed.

After the results of this analysis were published in \cite{24}, measurements of the \(t\)-channel cross section in \(pp\) collisions at \(\sqrt{s} = 8\) TeV with relative uncertainties of 16\% \cite{98} and 19\% \cite{99} arrived. Their combination yields a relative uncertainty of 14\% \cite{100}. The analyzed datasets correspond to an integrated luminosity of 5.0–5.8 fb\(^{-1}\). Furthermore, updated measurements from the Tevatron arrived. The most precise \(t\)-channel-cross-section measurement in \(p\bar{p}\) collisions at \(\sqrt{s} = 1.96\) TeV has an uncertainty of 16\% \cite{19}. The combined \(s\)- and \(t\)-channel cross-section was measured with a relative uncertainty of 14\% using the full dataset obtained with the D0 detector corresponding to an integrated luminosity of 9.7 fb\(^{-1}\) \cite{19}.

All mentioned measurements of electroweak top-quark production are compatible with the SM prediction.

**Coupling strength and determination of \(|V_{tb}|\)** The measurement of the inclusive production cross section \((\sigma)\) of single top (anti-)quarks offers a unique access to the \(\text{CKM}\)-matrix element \(|V_{tb}|\) without assuming \(\text{CKM}\)-matrix unitarity. The production cross section is proportional to the squared coupling strength at the \(Wtb\) vertex

\[
\sigma \propto |f_{VL} \times V_{tb}|^2. \tag{2.33}
\]

Here, \(f_{VL}\) is a form factor that effectively modifies the strength of the V-A interaction in terms of beyond-the-SM physics scenarios, and which is \(f_{VL} = 1\) in terms of SM interactions. The more precise the cross-section measurement is, the more precise is the determination of the fundamental SM parameter \(|V_{tb}|\).

Figure 2.8 summarizes the measurements of the \(\text{CKM}\)-matrix element \(|V_{tb}|\) in single-top-quark production. Before the analysis at hand was published, the most precise measurement of \(|V_{tb}|\) without assuming \(\text{CKM}\)-matrix unitarity was a combination of CDF and D0 results in \(s\)-
and \( t \)-channel production with \( p\bar{p} \) collisions at \( \sqrt{s} = 1.96 \) TeV. The result is \( |V_{tb}| = 0.88 \pm 0.07 \) with a relative uncertainty of 8% [23].

\( |V_{tb}| \) can also be inferred from electroweak loop corrections to the partial decay width \( \Gamma_{bb} \) of Z bosons decaying into bottom-quark pairs (\( Z \to b\bar{b} \)) [101]. In particular, this measurement does not assume CKM-matrix unitarity, and results in \( |V_{tb}| = 0.77^{+0.18}_{-0.24} \), which is consistent with the SM prediction (as obtained with CKM-matrix unitarity) at the level of one standard deviation.

![Figure 2.8: Direct measurements of the CKM-matrix element \( |V_{tb}| \) without assuming CKM-matrix unitarity [19, 20, 22, 23, 94–99].](image)

In more detail, the \( Wtb \) vertex is defined in a “most-general way” by

\[
\mathcal{L}_{Wtb} = - \frac{g}{\sqrt{2}} \bar{b} \gamma^\mu (V_L P_L + V_R P_R) W^- \mu
- \frac{g}{\sqrt{2}} \bar{b} \sigma^{\mu\nu} q_{\nu} \left( g_L P_L + g_R P_R \right) W^-_{\mu} + h.c.,
\]

(2.34)

in which \( m_W \) refers to the W-boson mass, \( q \) to the four-momentum of the W-boson, and \( P_L = \frac{1 - \gamma^5}{2} \) and \( P_R = \frac{1 + \gamma^5}{2} \) to the left-handed and right-handed projection operators [12–14, 102]. \( V_L = |f_{V_L} \times V_{tb}| \) refers to the strength of the left-handed vector coupling. \( V_R, g_L, \) and \( g_R \) refer to anomalous couplings. \( V_R \) refers to right-handed vector couplings, \( g_L \) to left-handed tensor couplings, and \( g_R \) to right-handed tensor couplings.
All couplings $V_L$, $V_R$, $g_L$, and $g_R$ are real contributions if conservation of Charge conjugation parity (CP) symmetry is assumed [13]. In the SM, $V_R$, $g_L$, and $g_R$ are zero at tree level [13]. In particular, $V_L$ is positive, and $V_L \equiv |V_{tb}| \approx 1$ in the SM [11].

The measurement of the single-top-quark cross section allows for two possible interpretations. First, the measurement can be interpreted in terms of SM couplings with $f_{V_L} \equiv 1$, in which $|V_{tb}|$ is constrained to the interval [0, 1]. Second, the absolute coupling strength can be inferred without any assumption on $f_{V_L}$, which yields the “unconstrained” $|f_{V_L} \times V_{tb}|$ measurement. Both interpretations will be discussed in more detail in the “Results” section (7.2).

Structure of the electroweak interactions at the $Wtb$ vertex Top quarks from electroweak production are highly polarized w.r.t. a certain spin basis due to the V–A structure of the interaction (cf. e.g. [15, 50, 103–108]).

The top-quark decay proceeds on a shorter timescale than the typical QCD-hadronization timescale, i.e. the timescale which would be necessary to form hadrons consisting of at least one top quark (cf. also the next section 2.2.4). Therefore, the top quark decays before it can be depolarized and characteristic angular correlations among its decay products occur.

The V–A structure of the weak interaction (eq. 2.34) can be probed e.g. by studying the angular correlations between the direction of top-quark spin and its decay products. As an example, the direction of the jet arising from fragmentation of the final-state light quark ($j$) may serve as a spin basis in which the top quark is maximally polarized. The expected polarization $p = \frac{N_j - N_\bar{j}}{N_j + N_\bar{j}}$ of the top (anti)quark in this basis is $p \approx 100\%$ [15, 50, 103–108]. A recent measurement in $t$-channel events gives $0.82 \pm 0.34$ [108]. The charged lepton ($l^\pm$) from the top-quark decay is used as the (most powerful) spin analyzer [109]. It has a spin-analyzing power of $|c_{l\pm}| = 1$ in the SM [109].

The cosine of the angle between the momentum vectors of the charged lepton $l^\pm$ and the direction of light quark $j$, both boosted to the (reconstructed) top-(anti)quark rest frame, yields a characteristic distribution for $t$-channel single-top-quark events. This angle is referred to as $\theta^* \equiv \theta(j, l^\pm)$, and its differential distribution is given by [15]

$$
\frac{1}{\sigma} \frac{d\sigma}{d \cos \theta^*} = \frac{1}{2} (1 + pc_{l\pm} \cos \theta^*).
$$

(2.35)

This angular correlation is accessible even on reconstruction level (cf. e.g. [24, 108, 110]), and can be utilized to explore the structure of the charged-current interaction. Contributions due to anomalous couplings at the $Wtb$ vertex, i.e. $V_R \neq 0$, $g_L \neq 0$, or $g_R \neq 0$, would result in deviations to the expected angular correlations (cf. e.g. [102, 111]). A recent measurement of the $\cos \theta^*$ distribution shows that the SM prediction is in agreement with data [108].

In particular, the information about $V_L$, $V_R$, $g_L$, and $g_R$ as obtained from single-top-quark production is complementary to measurements of the W-boson-helicity fractions\(^5\) in top-quark-pair production [16, 112]. Moreover, dedicated searches for $CP$ violation in $t$-channel events provide constraints on the complex phase of right-handed tensor couplings $g_R$ [113].

Already the measurement of the inclusive $t$-channel production cross section provides additional constraints to anomalous couplings. In figure 2.9a, the constraints on $V_L$ and $V_R$ under the assumptions $g_L = g_R = 0$ (left plot), as well as $g_L$ and $g_R$ under the assumptions $V_L = 1$ and

---

\(^5\)The helicity of the W bosons from the top-quark decay is determined by the V–A structure of the weak interaction at the decay vertex and can be either left-handed ($F_-$), right-handed ($F_+$), or longitudinal ($F_0$).
$V_H = 0$ (right plot) are shown. The measurement of the inclusive $t$-channel production cross section \cite{21} is labeled as “$\sigma_t$ (CMS 2010)”. This measurement has an uncertainty of 36%, driven by a previous iteration of the analysis at hand. Information about the measurement of the W-boson-helicity fractions is encoded in the measurement of angular asymmetries “$A_L$ (ATLAS 2010)” \cite{114}. The uncertainties of these measurements are (5–14)%. Both measurements are performed in $pp$ collisions at $\sqrt{s} = 7$ TeV with data corresponding to integrated luminosities of 35–36 pb$^{-1}$. A precise $t$-channel-cross-section measurement is expected to significantly constrain the phase space for anomalous couplings at the $Wtb$ vertex.

![Graph showing constraints on anomalous couplings](image)

**Figure 2.9:** Constraints on anomalous couplings as obtained from the measurements of the $t$-channel cross section (\cite{21}, with results from a previous iteration of this analysis) and W-boson-helicity fractions \cite{114} in $pp$ collisions at $\sqrt{s} = 7$ TeV with data corresponding to an integrated luminosity of 36 pb$^{-1}$. Both figures are from ref. \cite{16}.

Furthermore, angular correlations due to the V–A structure of the weak interaction are implicitly assumed in SM measurements of single-top-quark production. Multivariate analyses exploit the prior knowledge of electroweak top-quark production and probe the data of SM single-top-quark event topology.

**Charge-production asymmetry and constraints to the proton PDFs**

$t$-channel single-top-quark production produces top quarks more often than top antiquarks by a factor of $\approx 1.85$ in $pp$ collisions (cf. table 2.5). The charge of the virtual W boson is determined by the charge of the initial-state quark (cf. fig. 2.6). An up-type quark leads to the production of a top quark, while a down-type quark leads to the production of a top antiquark. Instead, the initial-state gluon (or initial-state b parton in case of the 5-flavor scheme) stems from the proton sea and is independent of the proton charge.

Thus, the relative production rate of $t$-channel top-quarks and top antiquarks originates from the parton composition of the proton, i.e. the proton PDFs because of the $qg$ (or $q\bar{b}$) initial state of the $t$-channel. The ratio of top-quark and top-antiquark-production cross sections

$$R_{\sigma_t/\sigma_{\bar{t}}} = \sigma_t/\sigma_{\bar{t}}$$

(2.36)

is sensitive to the description of the proton PDFs. The measurement of $R_{\sigma_t/\sigma_{\bar{t}}}$ then provides constraints to the proton PDFs.

The integrated charge ratio $R_{\sigma_t/\sigma_{\bar{t}}}$ was measured in $pp$ collisions at $\sqrt{s} = 7$ TeV by the ATLAS experiment with an uncertainty of 12% \cite{17} and in $pp$ collisions at $\sqrt{s} = 8$ TeV by the CMS...
experiment with an uncertainty of 15% [18]. In these measurements, data corresponding to an integrated luminosity of 4.7 fb$^{-1}$ [17] and 12.2 fb$^{-1}$ [18] were analyzed.

Figure 2.10 compares the $R_{\sigma_{t}/\sigma_{t'}}$ measurements with various NLO PDF sets. Both measurements start constraining the proton PDFs but are not yet competitive even with a large amount of analyzed data. In the left figure, the yellow band corresponds to the statistical uncertainty of the measurement, and the green band corresponds to the total uncertainty. The theoretical prediction is obtained with MCFM [76, 115]. The uncertainties of the theoretical prediction for the PDF sets correspond to variations of renormalization and factorization scales. In the right figure, the theoretical prediction is obtained by reweighting the simulated events. The uncertainties of the theoretical predictions combine contributions due to statistical uncertainties, top-quark-mass variations, and variations of factorization and renormalization scales. The smaller error bar of the measurement refers to the statistical uncertainty, and the larger error bar corresponds to the total uncertainty.

![Figure 2.10: Measurements of the ratio of t-channel top-quark and top-antiquark production in pp collisions at $\sqrt{s} = 7$ TeV [17] (left) and $\sqrt{s} = 8$ TeV [18] (right). Definitions of the error bars are given in the text.](image)

**Beyond the SM** Single-top-quark production in terms of new physics phenomena is discussed e.g. in ref. [15 50 90 116 117]. Different types of new physics phenomena manifest themselves in different single-top-quark production modes. Therefore, precise measurements of $s$, $t$, and $tW$-channel production cross sections will disentangle different types of new physics phenomena.

Additional non-SM particles that couple to top quarks can be grouped into i) new gauge or scalar bosons, which couple to SM top and bottom quarks, or ii) additional fermions, which couple to SM bosons and fermions. The $s$-channel is most sensitive to new (charged) resonances, which may significantly enhance the $s$-channel cross section. However, it has not been individually observed yet and the SM cross section is relatively low. Instead, the production cross section of a $t$-channel exchange of new bosons $X$ is suppressed by $1/m_X^2$ [15 90]. The $tW$-channel requires by definition a real W boson in the final state, which can be (mostly) distinguished from new particles. Hence, the $tW$-channel is rather insensitive to additional non-SM particles.

**Extra gauge bosons** include charged bosons that couple the top quark to down-type fermions, e.g. $W'^{\pm}$ bosons with charge $Q = \pm e$ in case the electromagnetic symmetry is conserved. A representative for this process is $q\bar{q}' \rightarrow W' \rightarrow t\bar{b}$ as shown in figure 2.11a $W'$ bosons
may exclusively couple to left-handed or right-handed fermions, as well as to combinations of both (cf. [4]). Moreover, neutral bosons may couple the top quark to up-type fermions. If the collision energy is lower than the production threshold of a new resonance, interference with the SM contributions can occur. In case of interference, the production cross section of SM processes, as well as their kinematics, is altered (cf. [50]).

Extra scalar bosons include charged or neutral top-pions $\pi^{\pm,0}$. These are predicted by top-color models [15,90]. Representative processes are e.g. $c\bar{b} \to \pi^+ \to \bar{t}b$, in which the $\pi^+$ couples to right-handed top and charm quarks (fig. 2.11b), and $gg \to \pi^0 \to t\bar{c}$ [90]. Top-pions are predominantly produced by an $s$-channel exchange, since they are expected to be as heavy as $\mathcal{O}(100 \text{ GeV}/c^2)$ [15,90]. Moreover, a charged Higgs boson $H^\pm$ that is produced in association with a single top (anti)quark is proposed in some models, which extend the SM by a “non-minimal Higgs sector” (cf. [90,118,119]).

Additional fermions, e.g. vector-like up-type quarks ($t'$) or a fourth generation of quarks ($t'$ and $b'$), may be produced in association with single top quarks. If additional fermions exist, but have too high masses to be produced directly, they can be indirectly inferred with single-top-quark production. Additional fermions can be indirectly inferred if they mix with the third quark generation via the weak interaction. Since unitarity of the $3 \times 3$ CKM matrix would be broken, $|V_{tb}|$, $|V_{ts}|$, and $|V_{td}|$ can significantly differ from the expected values [10](cf. [116]). As an example, $|V_{tb}|$ could be much smaller than one, while $|V_{ts}|$ and $|V_{td}|$ could be enhanced.

The measurement of the production cross sections of all three single-top-quark production mechanisms ($s$, $t$, and $tW$-channel) will provide important information not only about the magnitude of $|V_{tb}|$, but also about the structure of the CKM matrix: The $t$- and $tW$-channel production cross sections are directly sensitive to the magnitudes of $|V_{ts}|$ and $|V_{td}|$ due to their initial states, while the $s$-channel cross section is not. $|V_{ts}| \gg 0$ and $|V_{td}| \gg 0$ would manifest themselves in significantly larger $t$- and $tW$-channel production cross sections due to the high proton down and strange PDFs while $s$-channel production would be decreased (cf. [90,117]).

The $t$-channel is most sensitive to modified top-quark interactions. These include anomalous couplings at the $Wtb$ vertex, which are discussed in the previous paragraphs, and

\[10\] Here, the “expected values” refer to the values as derived when assuming CKM-matrix unitarity.
2.2 The Top Quark

**flavor-changing neutral currents (FCNCs)** FCNCs are forbidden at tree level in the SM and suppressed in higher orders, but can be significantly enhanced due to new physics phenomena [14, 15, 120]. A mixing of the top quark and lighter quarks (up and charm) can be e.g. mediated by $Z$, $\gamma$, $g$, and $h$ bosons ($t \rightarrow qZ$, $t \rightarrow q\gamma$, $t \rightarrow qg$, and $t \rightarrow qh$ with $q = u, c$) [90, 120].

FCNCs can be studied in top-(anti)quark decays as well as at the top-quark-production vertex. While top-quark decays are naturally explored in top-quark-pair production, FCNCs at the production vertex can be probed in single-top-quark processes. The most interesting coupling in terms of single-top-quark production is the production of a top quark via the $tqg$ vertex, since the process $t\bar{t} \rightarrow tqg$ has a huge background due to QCD-multijet production. Figure 2.12 shows examples of Feynman diagrams in LO-QCD accuracy for the $tqg$ interactions in the $s$- and $t$-channel, as well as for the associated production.

![Feynman diagrams](image)

Figure 2.12: Feynman diagrams in LO-QCD accuracy of FCNCs $tqg$ interactions in the $s$- and $t$-channel, as well as in associated production (cf. [50, 90]).

Even if the $tqg$-coupling strength would be small, the FCNC $t$-channel cross section (fig. 2.12b) profits from the fact that its initial state has a relatively large density in the proton PDFs. From figure 2.12a it can be seen that there is no interference with the SM $s$-channel due to the different initial and final states. The associated production (fig. 2.12c) has a different final state than the SM $tW$-channel. Hence, the SM $tW$-channel is insensitive to FCNCs.

The “effective” Lagrangian density that describes FCNCs involving a gluon at a scale $\Lambda$ is given by

$$L_{\text{FCNC}} = \frac{\kappa_{tqg}}{\Lambda} g_s \bar{q} \sigma^{\mu\nu} T^a t G^{a\mu\nu}$$

(2.37)

in which $\kappa_{tqg}$ defines the strength of the $tqg$ couplings, $q = u, c$ and $t$ refer to the quark fields, $\sigma^{\mu\nu} = \frac{i}{2}(\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu)$, $T^a$ are the generators of the SU(3), and $G^{a\mu\nu}$ refers to the gluon-field-strength tensor [50].

Single top (anti)quarks can also be produced in association with particles that leave undetected, but which experimentally manifest themselves as (large) missing transverse momenta, e.g. as discussed in $R$-parity-violating Supersymmetry (SUSY) (cf. [121]). Signatures include both $s$- and $t$-channel diagrams. In a model-independent approach, this signature is called “monotop” and discussed e.g. in ref. [117, 121].

The $tW$-channel is rather insensitive to new physics phenomena, except for (anomalous) couplings at the $Wtb$ vertex and the magnitude of the CKM-matrix elements $|V_{tb}|$. Since the $tW$-channel has a rather large production cross section at the LHC it can provide important complementary information to disentangle new physics phenomena if observed in one of the other two channels (cf. [90]).

So far, no evidence for new physics phenomena have been found in single-top-quark production. An overview of current searches for new physics phenomena is given e.g. in ref. [5, 43].
2.2.4 Top-Quark Decay

The top quark decays via the weak interaction. It almost exclusively decays into a W boson and a b quark ($t \rightarrow Wb$, fig. 2.13). Top-quark decays into $W_s$ and $W_d$ final states are suppressed by the squared magnitude of the CKM matrix elements $V_{ts} = 0.0404^{+0.0011}_{-0.0005}$ and $V_{td} = 0.00867^{+0.00029}_{-0.00031}$ [11]. $V_{ts}$ and $V_{td}$ are indirectly determined utilizing information from $B_s$-$\bar{B}_s$ and $B_d$-$\bar{B}_d$ oscillations, as well as CKM-matrix unitarity with three generations (cf. sec. 2.1.3 and [11, 15] for more details). If assuming CKM-matrix unitarity with three families, the CKM matrix element $V_{tb}$ is determined to be $V_{tb} = 0.999146^{+0.000021}_{-0.000046}$ [11].

The top-quark-branching ratio

$$\mathcal{R} = \frac{B(t \rightarrow Wb)}{\sum_{q=d,s,b} B(t \rightarrow Wq)} = |V_{tb}|^2$$

Figure 2.13: Top-quark decay into a b quark and a W boson, which subsequently decays into either leptons or quarks.

The $W$-boson-decay modes and their relative fractions are given in table 2.7. In this analysis, $W$-boson decays into electrons or muons are analyzed, which have a partial branching ratio of approximately 21%.

<table>
<thead>
<tr>
<th>$\mathcal{R}$</th>
<th>$\sqrt{s}$ [TeV]</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90 ± 0.04</td>
<td>1.96</td>
<td>[124]</td>
</tr>
<tr>
<td>0.94 ± 0.09</td>
<td>1.96</td>
<td>[122]</td>
</tr>
<tr>
<td>0.98 ± 0.04</td>
<td>7</td>
<td>[125]</td>
</tr>
<tr>
<td>1.023^{+0.036}_{-0.034}</td>
<td>8</td>
<td>[126]</td>
</tr>
<tr>
<td>0.97^{+0.09}_{-0.08}</td>
<td>1.96</td>
<td>[123]</td>
</tr>
</tbody>
</table>

Table 2.6: Measurements of the top-quark-branching ratio.
2.2 The Top Quark

<table>
<thead>
<tr>
<th>Decay mode</th>
<th>Fraction ($\Gamma_i/\Gamma_{\text{total}}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Gamma_1$</td>
<td>$e^+ \nu$ $(10.75 \pm 0.13)%$</td>
</tr>
<tr>
<td>$\Gamma_2$</td>
<td>$\mu^+ \nu$ $(10.57 \pm 0.15)%$</td>
</tr>
<tr>
<td>$\Gamma_3$</td>
<td>$\tau^+ \nu$ $(11.25 \pm 0.20)%$</td>
</tr>
<tr>
<td>$\Gamma_4$</td>
<td>hadrons $(67.60 \pm 0.27)%$</td>
</tr>
</tbody>
</table>

Table 2.7: $W^+$-boson decay modes. $W^-$ boson decay modes are charge conjugates. The table is an excerpt from ref. [11].

The **top-quark-decay width** in LO-QCD accuracy is given by

$$\Gamma_t^{\text{LO}} \approx \Gamma_t^{\text{LO}}(t \to Wb) = \frac{G_F m_t^3}{8 \pi \sqrt{2}} |V_{tb}|^2 \left( 1 - \frac{m_W^2}{m_t^2} \right)^2 \left( 1 + 2 \frac{m_W^2}{m_t^2} \right)$$

(2.39)

with the CKM-matrix element $|V_{tb}|$, Fermi constant $G_F$, W-boson mass $m_W$, and top-quark mass $m_t$ [15, 43]. Furthermore, the b-quark mass is neglected in eq. 2.39. The prediction of the top-quark width in NNLO-QCD accuracy is [47]

$$\Gamma_t^{\text{NNLO}} = (1.32 \pm 0.8\%) \text{ GeV}.$$  

(2.40)

Here, the uncertainty is due to the choice of the renormalization scale. A top-quark mass of 172.5 GeV/$c^2$ is assumed, and a finite b-quark mass is used in this calculation.

A direct measurement of $\Gamma_t$ in $t\bar{t}$ events confirms the SM prediction. The measurement yields

$$1.10 < \Gamma_t < 4.05 \text{ GeV}$$  

(2.41)

at 68% CL when assuming a top-quark mass of $m_t = 172.5 \text{ GeV}/c^2$ [127].

$\Gamma_t$ is also indirectly determined by combining the measurements of the partial decay width $\Gamma_t(t \to Wb)$, which is obtained using the measured $t$-channel-single-top-quark-production cross section, and the measurement of the branching ratio $B(t \to Wb)$, which is obtained from $t\bar{t}$ production (cf. [3, 124]). $\Gamma_t$ is then given by [128]

$$\Gamma_t = \frac{\Gamma_t(t \to Wb)}{B(t \to Wb)}$$  

(2.42)

with $\Gamma_t(t \to Wb) = \sigma_{t\text{-channel}} \frac{\Gamma_t^{\text{SM}}(t \to Wb)}{\sigma_{t\text{-channel}}^{\text{SM}}}$. This measurement assumes CKM-matrix unitarity in the determination of $B(t \to Wb)$, and yields [128]

$$\Gamma_t = 2.00^{+0.47}_{-0.43} \text{ GeV}.$$  

(2.43)

The **lifetime of the top-quark** is predicted to be [11, 15]

$$\tau_t = 1/\Gamma_t \approx 0.5 \times 10^{-24} \text{ s}.$$  

(2.44)

This lifetime is smaller than the typical QCD hadronization timescale, which is of order (cf. [15])

$$\tau_{\text{hadr}} = 1/\Lambda_{\text{QCD}} \approx 3.0 \times 10^{-24} \text{ s}.$$  

(2.45)
The top-quark lifetime is measured to be in the interval
\[ 0.16 \times 10^{-24} \text{ s} < \tau_t < 0.6 \times 10^{-24} \text{ s} \] (2.46)
at 68\% CL \[ \text{[127]} \]. The measurement in ref. \[ \text{[128]} \] yields \( \tau_t = 0.329^{+0.090}_{-0.063} \times 10^{-24} \text{ s} \).

Exotic top-quark decays have not been observed yet. Flavor-changing-neutral currents (e.g. \( t \rightarrow Zq \), \( t \rightarrow gq \), and \( t \rightarrow \gamma q \) with \( q = u, c \)) are loop-suppressed in the SM \[ \text{[3, 15]} \]. Therefore, the branching ratio of these decays is low. Examples for top-quark decays beyond the SM are searches for decays into a charged Higgs boson \( t \rightarrow H^\pm b \), searches for modifications to the top-quark decay (which include anomalous couplings in the \( tbW \) vertex), and decays into supersymmetric particles. More details are given e.g. in ref. \[ \text{[3, 4, 15, 43]} \].
3 Experimental Setup

This chapter introduces the experimental framework in which this analysis is performed. This chapter initially starts (section 3.1) with a description of the particle-accelerator complex that provides the proton-proton collisions, namely the Large Hadron Collider (LHC). The protons are extracted from a hydrogen source and successively accelerated through an accelerator chain consisting of linear accelerators and synchrotrons. The main storage ring has a complex magnet system consisting of about 9600 magnets which focus and bend the proton beams. The LHC is designed to accelerate protons up to an energy of $7\text{ TeV}$ per beam with a peak luminosity of $10^{34}\text{ cm}^{-2}\text{ s}^{-1}$. The general scientific goals of the six large LHC experiments are described at the end of the first section.

The second part of this chapter (section 3.2) is devoted to the description of one of the main experiments at the LHC, the Compact Muon Solenoid (CMS) detector. The CMS detector is a general-purpose detector, which is constructed following the “classical principles” of modern collider-detector design, in which the beamline is consecutively enclosed by several layers of subdetectors. Closest to the beamline is the inner-tracking system, followed by the electromagnetic and hadronic calorimeters, the superconducting solenoid, and the muon chambers, which are located outside the solenoid coil. Furthermore, the trigger system and detector simulation are briefly described.

The third part of this chapter (section 3.3) discusses the reconstruction of physics objects, as well as their performance in terms of response, resolution, and uncertainties. Estimated and simulated performance are compared with measurements in data. The vertex reconstruction and Particle-Flow-event reconstruction are described first. The reconstruction of muons, electrons, jets, and missing transverse energy, as well as jet-flavor tagging are discussed afterwards.

The fourth part (section 3.4) discusses the recorded collision data and the luminosity determination. An overview of various techniques to measure luminosity is given. In this analysis, the luminosity is calculated offline with a technique based on counting silicon-pixel-tracker clusters. An absolute normalization is determined with a Van der Meer (VdM) scan.

The fifth part (section 3.5) discusses the simulation of physics events. Signal and background modeling and their normalization are described. Corrections to simulated events include muon- and electron-trigger reweighting, Pile-Up reweighting, corrections to the jet-transverse-momentum resolution, and jet-flavor tagging.

The last part of this chapter, section 3.6, is devoted to the description of the software packages and libraries that are used in this analysis.
3 Experimental Setup

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a superconducting particle accelerator and particle collider that is located at the Franco-Swiss border near Geneva (cf. [129]). The LHC is operated at the Conseil Européen pour la Recherche Nucléaire (CERN). It is installed in the 26.7 km-long tunnel of the former Large Electron-Positron Collider (LEP) and uses parts of the former LEP injection chain [129].

The LHC can be operated in three modes with proton-proton (pp) collisions, lead-lead (PbPb) collisions, or proton-lead (pPb) collisions. Lead is accelerated completely ionized ($^{208}\text{Pb}^{82+}$).

The proton mode is designed [129] for collisions at a peak instantaneous luminosity of $L = 10^{34} \text{cm}^{-2} \text{s}^{-1}$ at a center-of-mass energy of $\sqrt{s} = 2E_{\text{beam}} = 14 \text{TeV}$. (3.1)

Proton-proton collisions were recorded at center-of-mass energies of 900 GeV in the end of 2009. These collisions are mostly used for commissioning of the detector hardware and physics-object reconstruction. Collisions for physics analyses were delivered to the experiments at a center-of-mass energy of 7 TeV with a total integrated luminosity of $L = 6.1 \text{fb}^{-1}$ in 2010 and 2011, and at 8 TeV with $L = 23.3 \text{fb}^{-1}$ in 2012. These runs are extensively analyzed in the context of SM measurements and searches for new physics phenomena.

The heavy-ion mode is designed for collisions at a peak instantaneous luminosity of $L = 10^{27} \text{cm}^{-2} \text{s}^{-1}$ at an energy of $\sqrt{s} = 2.76 \text{TeV per nucleon}$ with a total center-of-mass energy of 1.1 PeV [129]. Data corresponding to an integrated luminosity of $L = 166 \mu \text{b}^{-1}$ were delivered to the experiments at $\sqrt{s} = 2.76 \text{TeV per nucleon}$ in 2011.

The asymmetric pPb mode helps to decouple effects in PbPb collisions that originate from having heavy ions in the initial state. During the pPb run in the beginning of 2013, data corresponding to $L = 32 \text{nb}^{-1}$ at $\sqrt{s} = 5.02 \text{TeV per nucleon}$ were delivered by the LHC.

The LHC magnet system bends the particle beams on their circular path within the storage ring. A complex injector chain is required before filling the beam into the main LHC storage ring. Both systems are explained in the following, and the definition of luminosity is explained. Afterwards, the scientific goals that motivated the construction of the LHC are described, as well as the experiments built to reach those goals.

Magnet System

Same-charged hadrons are accelerated in two counter-rotating beams in order to achieve the high luminosity. These hadrons are bend by a complex system with superconducting magnets. The superconducting magnets are cooled with superfluid helium down to temperatures of below 2 K, and they are made of NbTi Rutherford wires. A total of 1232 dipole magnets with magnetic fields up to 8.33 T are used to bend hadrons onto their circular path [130]. Quadrupole magnets are used to focus or defocus the beams. Beams are focused to a size of $\approx 16 \mu \text{m}$ at the interaction point in order to achieve a maximal number of interactions per crossing, and the beams have sizes of 1 mm between the interaction points [131]. Magnets with higher orders are used for residual focusing of the beam and adjustments to the beam trajectory. In total, about 9600 magnets are used to bend and focus the beams [130].

Electromagnetic resonators are used to accelerate or decelerate the particles such that their energy is synchronized within a bunch and close to the target energy. Eight superconducting accelerating cavities per beam are installed at the LHC. Each cavity provides an accelerating field of 5.5 MV/m [129]. The radio frequency of the LHC is 400.79 MHz. Maximally 2808
bunches are possible at a design bunch spacing of 24.95 ns \[129\] and an LHC-orbit frequency (or “revolution frequency”) of \( f = 11.245 \text{ kHz} \[130\]. At the design-bunch spacing, two consecutive bunches are separated by a distance of 8 m \[131\].

**Injector chain and Beam Energy** The LHC is designed to deliver particle beams with energies up to 7 TeV per proton to the experiments, i.e. 574 TeV per lead ion. The CERN accelerator complex with the LHC-injector chain, the LHC storage ring, and its experiments at the four interaction regions is shown in fig. 3.1. The following, brief introduction focuses on the acceleration of protons (cf. \[129, 132\]). The lead-injector chain is described in detail elsewhere \[132\].

Protons are extracted from a hydrogen source using a so-called duoplasmatron \[133\]. The proton beam is fed into the Linear accelerator 2 (Linac2) in which it is accelerated to an energy of 50 MeV. Afterwards, the beam is sequentially ramped up with a few synchrotrons. The Proton Synchrotron Booster (PSB) accelerates the beam to 1.4 GeV, the Proton Synchrotron (PS) to 25 GeV, and the Super Proton Synchrotron (SPS) up to an energy of 450 GeV. Finally, the proton beam is injected into the main LHC ring with an energy of 450 GeV, a transverse normalized emittance of 3.5 \( \mu \text{m} \), and a bunch spacing of 24.95 ns (all LHC design criteria). The LHC storage ring accelerates the protons to their final energy of up to 7 TeV.

Figure 3.1: The CERN accelerator complex with the injector chain and its experiments at the four interaction regions \[134\].
Experimental Setup

### Luminosity

The number of interactions per time interval for a particular process is given by

\[
\frac{dN_{\text{events}}}{dt} = L \times \sigma_{\text{process}},
\]

(3.2)

in which \(L\) refers to the instantaneous luminosity, and \(\sigma_{\text{process}}\) refers to the production cross section. The instantaneous luminosity in a certain data-taking period is referred to as “integrated luminosity”,

\[
\mathcal{L} = \int L \, dt,
\]

(3.3)

and the number of produced events in this period is given by

\[
N_{\text{events}} = \mathcal{L} \times \sigma_{\text{process}}.
\]

(3.4)

The instantaneous luminosity (cf. ref. [129, 131]) is defined as

\[
L = \frac{N_b^2 n_b f}{4\pi \sigma_x^* \sigma_y^*} F = \frac{N_b^2 n_b f \gamma_r}{4\pi \epsilon^* \beta^*} F
\]

(3.5)

if one assumes the particles intensities to be Gaussian distributed within both beams. Here, \(N_b\) refers to the number of particles per bunch, \(n_b\) to the number of bunches per beam, \(f\) to the LHC-orbit frequency, \(\gamma_r\) to the relativistic gamma factor, \(\epsilon^*\) to the normalized transverse emittance, \(\beta^*\) to the beta function evaluated at the interaction point, and \(F\) to the geometric luminosity-reduction factor. The luminosity increases with the number of colliding protons and bunches, but decreases with larger beam sizes.

In order to achieve the design luminosity of \(L = 10^{34} \text{cm}^{-2} \text{s}^{-1}\) for proton-proton collisions at \(\sqrt{s} = 14\) TeV, \(n_b = 2808\) bunches per beam with \(N_b = 1.15 \times 10^{11}\) protons per bunch are required. Furthermore, \(\gamma_r\) is 7461 for protons that are accelerated to an energy of 7 TeV [130]. The normalized (transverse) emittance is designed to be \(\epsilon^* = 3.75\) µm. A \(\beta^* = 0.55\) m is assumed at the interaction point [130].

The geometric luminosity reduction factor \(F\) is 1 for head-on collisions and \(F \leq 1\) for non-parallel beams. The crossing angle between both beams at the LHC is very small (\(\approx 200\) µrad [129]) and \(F\) is close to 1.

The beam sizes in transverse directions \(x\) and \(y\) are expressed in terms of the beta function \(\beta_{x,y}\) and the emittance \(\epsilon_{x,y}\) (cf. [131]). The transverse beam sizes are given by \(\sigma_{x,y} = \sqrt{\beta_{x,y} \epsilon_{x,y}}\). The beams are (approximately) round at the interaction point. The emittance \(\epsilon = \epsilon^* / \gamma_r\), with \(\gamma_r = E/(m_0 c^2)\) becomes smaller with increasing energy \(E > E_0\). Here, \(E_0\) refers to the injection energy, and the normalized (transverse) emittance \(\epsilon^*\) is defined by the machine parameters and remains constant. If the emittance \(\epsilon\) becomes smaller, the transverse beam sizes decrease. The transverse beam sizes at an energy \(E > E_0\) are then given by \(\sigma_{x,y}(E) = \sigma_{x,y}(E_0) \cdot \sqrt{\frac{E_0}{E}}\) (cf. [131]).

### Scientific Goals and Experiments

Both beams are being collided at four separate interaction points in huge caverns. The interaction points are surrounded by large detectors which observe the interactions.

Two of these experiments are general-purpose and high-luminosity experiments, namely the ATLAS experiment [135] and the CMS experiment [136]. The
3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) experiments study collisions with both protons and heavy ions, and they cover a diversified physics program (cf. [135, 136]). They perform precision measurements of SM processes and SM parameters in terms of QCD or electroweak interactions, as well as flavor physics. Furthermore, the LHC enables a precise determination of the top-quark couplings and properties. One of the main motivations of the two large LHC experiments is to understand the mechanism that breaks electroweak symmetry. Both ATLAS and CMS experiments are dedicated to observe or exclude the Higgs boson [11] over a wide range of masses with diversified couplings to bosons and fermions. Searches for new physics phenomena include supersymmetric extensions to the SM, “exotic” physics signatures (e.g. heavy charged gauge bosons, extra dimensions, or black holes), and additional quark generations among others. A big step forward has been recently made when the two large experiments, ATLAS and CMS, announced the discovery of a new boson with a mass near 125 GeV/c² [40, 41].

The A Large Ion Collider Experiment (ALICE) (cf. [137]) has the primary goal to improve the understanding of heavy-ion interactions at high energies. Heavy-ion interactions are expected to result in an extremely hot and dense state of matter, the so-called “quark-gluon plasma”. Quarks and gluons are deconfined within this experimentally challenging environment.

The Large Hadron Collider beauty (LHCb) experiment (cf. [138]) is designed to address heavy-flavor physics. LHCb analyses rare decays of charm and bottom (beauty) hadrons. The precise measurement of the amount of CP violation in heavy flavor processes is a crucial test of the SM and a window to searches for new physics phenomena. LHCb is designed to operate at an instantaneous luminosity of $2 \times 10^{32}$ cm$^{-2}$ s$^{-1}$. At this luminosity, radiation damage to the detector is minimized and the number of simultaneous interactions is lower, such that the lifetime of the detector is extended. For this reason, only a fraction of the total luminosity was delivered to the LHCb experiment in 2011 and 2012. LHCb recorded data corresponding to an integrated luminosity of 1.1 fb$^{-1}$ at a center-of-mass energy of $\sqrt{s} = 7$ TeV and 2.1 fb$^{-1}$ at $\sqrt{s} = 8$ TeV.

The Total Elastic and Diffractive Cross Section Measurement (TOTEM) experiment (cf. [139]) studies very forward charged particles within a pseudo-rapidity range of $3.1 \leq |\eta| \leq 6.5$. The main objectives of the TOTEM experiment are a luminosity-independent measurement of the total proton-proton cross section, a luminosity measurement and luminosity monitoring that is independent and complementary w.r.t. the other experiments, and measurements of the elastic and diffractive proton-proton scattering. The detectors are installed in the cavern of the CMS experiment at distances of $\pm 145$ m and $\pm 220$ m from the interaction point.

The Large Hadron Collider forward (LHCf) detectors (cf. [140]) are installed beyond the interaction point at which the ATLAS detector is located. LHCf is the smallest of the six experiments and studies very forward neutral particles. These particles have “laboratory equivalent energies” up to $10^{17}$ eV. The LHCf experiment facilitates the understanding of hadron-interaction models at high energies, in particular at energies between the “knee” and the Greisen-Sazepin-Kusmin (GZK) cut-off [141, 142] of the cosmic-ray-energy spectrum.

Besides these large experiments that study highly energetic hadrons from the LHC, a couple of smaller experiments work with low-energetic beams at intermediate steps of the accelerator chain, e.g. in fixed target experiments.
3 Experimental Setup

3.2 The Compact Muon Solenoid Detector

The CMS detector (cf. [136]) is one of the two general-purpose detectors at the LHC. It is installed in an experimental cavern which is located underground near Cessy in France. The CMS detector is designed for proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 14$ TeV with a peak luminosity of $10^{34}$ cm$^{-2}$ s$^{-1}$ and for heavy-ion collisions at $\sqrt{s} = 1.15$ PeV with a peak luminosity of $10^{27}$ cm$^{-2}$ s$^{-1}$. A collaboration of over 4500 active people from over 40 countries contributes to the success of the CMS experiment.

The scientific goals of the LHC experiments are outlined in the previous section (3.1). The implementation of the general physics program of the CMS experiment is further described in detail in ref. [143]. Especially the search for the Higgs boson with its diversified decay pattern motivated the following four design goals (cf. [136]).

First, the CMS detectors ability to identify muons is excellent. Muons are reconstructed with high momentum resolution and low charge-misidentification probability. Dimuon masses are measured with a resolution of 1% at 100 GeV/c$^2$. A muon system consisting of three different technologies covers a wide range of pseudo-rapidity. The muon system is complemented by the inner-tracking system in combination with a strong magnetic field.

Second, the inner-tracking system reconstructs the trajectories of charged particles precisely and highly efficient. Jets from b-quark fragmentation and $\tau$-leptons are efficiently identified. Furthermore, the inner-tracking system is able to disentangle the hard-scattering process from many additional soft interactions, and is able to cope with thousands of tracks in heavy-ion collisions. For that reason, CMS has an inner-tracking system that is made of silicon detectors.

Third, dielectrons and diphotons are reconstructed with a mass resolution which is comparable to the dimuon mass resolution. Neutral pions are efficiently rejected, since this process is a background process to a Higgs boson that decays into two photons. An electromagnetic calorimeter with a high granularity supports the identification of photons and charged leptons. Therefore, CMS has a homogeneous electromagnetic calorimeter, which fully consists of scintillating crystals, and dedicated preshower detectors with silicon-strip sensors, which have an even higher granularity than the crystals.

Fourth, the hadron calorimeter provide a good missing transverse energy and dijet-mass resolution. CMS has a hadron calorimeter that is “hermetic” in terms of a large geometrical coverage, accurately-fitting layers of absorbers and scintillators, and dense absorbers with large interaction lengths. The hadron calorimeter covers a pseudo-rapidity up to $|\eta| < 5.2$ around the whole beamline. It is made of over 70000 thin scintillator tiles that fit into the small gaps between the brass absorbers. The hadron calorimeter has an interaction length corresponding to about 10 $\lambda_l$.

The CMS detector is kind of a cylinder that fully surrounds the beamline, and which is centered at the nominal collision point. The coordinate system (cf. [136]) of the CMS detector is a right-handed coordinate system. The $z$ axis points in counterclockwise beam direction. The $x$ axis points radially towards the center of the LHC storage ring. The $y$ axis points vertically upward, i.e. perpendicular to the plane which is spanned by the LHC storage ring. The azimuthal angle $\phi$ is measured in the $x$-$y$ plane, in which $\phi = 0$ corresponds to a trajectory that is parallel to the $x$ axis, and $\phi = \pi/2$ corresponds to a trajectory that is parallel to the $y$ axis. Transverse components, e.g. transverse momenta, are always calculated transverse to the beam direction, i.e. in the $x$-$y$ plane. The radial distance to the beam axis (i.e. $z$ axis) in the $x$-$y$ plane is referred

---

1 June 2013
3.2 The Compact Muon Solenoid Detector

to as $r$. The polar angle $\theta$ is measured w.r.t. the positive $z$ axis. More usual in the field of high-energy physics is to use the pseudo-rapidity $\eta$ instead of the angle $\theta$. The pseudo-rapidity is defined [11] as

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

(3.6)

Trajectories along the beamline have angles $\theta = 0^\circ$ and pseudo-rapidity $\eta \to \pm \infty$. Trajectories transverse to the $z$ axis and parallel to the $y$ axis have angles $\theta = \pm 90^\circ$ and pseudo-rapidity $\eta \to 0$.

The CMS detector is of cylindrical shape with a length of 21.6 m and diameter of 14.6 m. CMS has a total weight of about 12500 t. The heavy return yoke of the solenoid, which is made of iron, contributes most with a weight of about 10000 t.

Figure 3.2 shows a perspective of the CMS detector. The structure of the CMS detector follows the “classical principles” of modern collider-detector design, in which the beamline is consecutively enclosed by several layers of subdetectors. The most inner layer is the tracker. The next layers are the Electromagnetic Calorimeter (ECAL) and the Hadron Calorimeter (HCAL). They are surrounded by a solenoid that provides a huge magnetic field of up to 4 T. The solenoid is enclosed in a massive iron return yoke, in which the muon chambers are installed. The forward regions in both directions are further covered by the Hadron-Forward Calorimeter (HF). The HF detectors surround the beamline at a distance of about 11 m from the center of the detector. The following subsections discuss each subdetector system, beginning with the inner tracker, which is closest to the beamline, up to the muon chambers, which are located inside the iron yoke.
3.2.1 The Tracker

The tracker (cf. [136]) is the subdetector which is closest to the beamline. The tracker covers a pseudo-rapidity range of $|\eta| < 2.5$ and is composed of silicon-pixel and silicon-strip modules. The whole tracker is cooled down to less than $-10^{\circ}C$ in order to avoid damage due to thermal runaway and radiation. The overall material budget of the tracker corresponds to radiation lengths of up to $2X_0$ or hadronic interaction lengths of up to $0.55\lambda_l$.

Figure 3.3 shows the profile of the inner-tracking system. The beamline is at $r = 0$ along the $z$ axis. The tracker has a length of 5.8 m and a diameter of 2.5 m. It consists of single-sided and double-sided modules, which are represented by single and double lines. The double-sided modules provide the measurement of a second coordinate.

The pixel tracker is the most inner subdetector component and directly surrounds the beamline. In the barrel, three layers provide three points for the reconstruction of a particle trajectory. The layers are separated by a few centimeters in $r$ direction. The pixel detector encloses the beamline with two disks in the endcaps. A total of 1440 silicon-pixel modules cover an area of $1m^2$ with about 66 million pixels. Each pixel has a size of $100 \times 150\mu m^2$ (in $r-\phi \times z$). The pixel detector allows for a precise three-dimensional reconstruction of primary vertices and secondary vertices from heavy-flavored hadrons and $\tau$ leptons.

The pixel modules are surrounded by the 15148 silicon-strip modules. In total 9.3 million strips yield an active silicon area of about 198 m$^2$. The silicon-strip modules are organized in three subdetector components, namely the Tracker Inner Barrel (TIB) and the Tracker Inner Disk (TID), the Tracker Outer Barrel (TOB) and the Tracker EndCaps (TEC). Each subdetector component has a certain number of layers, and provides several point measurements of a trajectory. The TIB and TID modules are composed of up to four layers with $320\mu m$ thick silicon micro-strip sensors, and they have single-point resolutions of $23\mu m$ and $35\mu m$. The TOB consists of $500\mu m$ thick strips with six layers and has single-point resolutions of $35 - 53\mu m$. Each TEC has nine layers with $320\mu m$ and $500\mu m$ thick strips. The inner layers of each subdetector component consist of double-sided modules. These provide a measurement of a second coordinate, resulting in single-point resolutions of $230\mu m$ and $530\mu m$ in TIB and TOB and varied resolutions in TID and TEC which depend on the pitches of the strips. A dedicated laser-alignment system monitors the positions of the subdetector components.

The resolution of the CMS-tracking system for muons is shown in figure 3.4. The CMS-tracking system is designed to provide a transverse impact-parameter resolution of about $10\mu m$ for muons with a $p_T$ of $100\text{GeV}/c$ in the central region, and slightly larger resolution at higher $|\eta|$. The longitudinal impact-parameter resolution for muons with a $p_T$ of $100\text{GeV}/c$ is $20 - 40\mu m$ for $|\eta| < 1$, but linearly increasing to $70\mu m$ at $|\eta| = 2.4$. Impact parameters for muons with a transverse momentum of $1\text{GeV}/c$ are measured with a resolution of $100 - 200\mu m$ in the transverse plane, and $100 - 1000\mu m$ in $z$ direction, depending on the pseudo-rapidity. The transverse-momentum resolution for $p_T = 100\text{GeV}/c$ muons is about $1 - 2\%$ within $|\eta| < 1.6$, and linearly increasing to $7\%$ at $|\eta| = 2.4$, if measured in the tracker only. Muon tracks with a $p_T$ up to $10\text{GeV}/c$ are measured with a transverse-momentum resolution of $\approx 0.5 - 2\%$. Their tracks are more bent, which allows a more precise measurement.

Furthermore, the CMS-tracking system precisely reconstructs vertices in three dimensions. Also secondary vertices, e.g. from heavy-flavored hadrons, are reconstructed with a resolution of $\approx 15\mu m$.
3.2 The Compact Muon Solenoid Detector

Figure 3.3: Profile of the CMS inner tracker composed of modules with silicon pixels and silicon strips. The beamline is at $r = 0$ along the $z$ axis. The tracker consists of single-sided and double-sided modules, which are represented by single and double lines.

Figure 3.4: Resolution of the CMS tracking system for muons expressed as a function of the pseudo-rapidity $|\eta|$. Shown are the transverse momentum (left), the transverse impact parameter (center), and the longitudinal impact parameter (right plot). This figure is taken from ref. [136].
3.2.2 The Calorimeter

The CMS detector has two main calorimeter. The ECAL measures particles that interact electromagnetically, i.e. electrons and photons. The HCAL measures the energy of particles that interact via the strong force, i.e. the deposits of hadrons, which stem from fragmentation of quarks and gluons. The measurement of an imbalance of the transverse momenta requires the HCAL to be hermetic. The CMS detector has a (mostly) homogeneous ECAL and a sampling HCAL. The calorimeter are explained in the following (cf. [136]).

The Electromagnetic Calorimeter

The ECAL (cf. [136]) is organized in three parts, namely a barrel, two endcaps, and two preshower detectors. All ECAL subdetectors are finely segmented. The fine segmentation allows for identification of isolated electrons and electrons from photon conversions. Figure 3.5 shows the overall layout of the ECAL. The barrel and endcaps are made of lead tungstate (PbWO$_4$) crystals. PbWO$_4$ is a material that is very dense (8.28g/cm$^3$) with a short radiation length of 0.89 cm and a Molière radius of 2.2 cm, which allows the ECAL to be very compact. Furthermore, PbWO$_4$ is optically transparent to the emitted scintillation light. The crystals emit scintillation light at the order of the time intervals of LHC bunch crossing, which makes a fast readout possible. The readout is done with photodetectors, either Avalanche photodiodes in the barrel or vacuum phototriodes in the endcaps. The ECAL is operated at a constant temperature of $(18 \pm 0.05)$ °C, since the performance of both crystals and photodetectors is sensitive to temperature. The crystals are subject to aging effects. A dedicated laser-monitoring system is used to measure the transparency loss of the ECAL crystals, which can be corrected for.

The ECAL barrel covers the pseudo-rapidity range of $|\eta| < 1.479$ and has 61200 crystals. Each crystal has a front face of $22 \times 22$ mm$^2$, rear face of $26 \times 26$ mm$^2$, is 230 mm long, and has a radiation length of $25.8 X_0$. The crystals are organized in modules of 400 or 500 crystals, and

Figure 3.5: Profile of the CMS ECAL [136].
four modules are grouped in super-modules due to mechanical reasons. The endcaps have each 7324 crystals and cover a pseudo-rapidity range of $1.479 < |\eta| < 3.0$. The crystals are 10 mm shorter in the endcaps than in the barrel and have radiation lengths of $24.7 \times X_0$. They also have a larger front face of $28.62 \times 28.62 \text{mm}^2$, as well as rear face of $30 \times 30 \text{mm}^2$ in the endcaps. Crystals are organized in super-crystals of dimension $5 \times 5$, which are arranged in two dees.

Preshower detectors are installed in the front of the ECAL crystals in the endcaps. They extend the ECAL within the pseudo-rapidity range $1.653 < |\eta| < 2.6$. The preshower detectors are designed as a “classical” two-layer sampling calorimeter, and they consist of lead absorbers and silicon-strip sensors. In total, the preshower detectors are 20 cm thick. Their silicon-strip sensors provide an even finer granularity than the PbWO$_4$ crystals. Each sensor contains 32 silicon strips which are 320 $\mu$m thick and have distances of 1.9 mm among each other. The main reason for the preshower detector is the identification of neutral pions which decay into two photons, since neutral-pion background can mimic a high-energetic, prompt photon. Two close photons from neutral pion background would simply hit the same crystal, and neutral pion background is enhanced in the endcap-$\eta$ region. That’s why the fine granularity of the preshower detectors is important. Such a high accuracy in discriminating among photon sources is important, e.g. to identify Higgs bosons in the two-photon-decay channel.

The overall energy resolution of the ECAL is measured in beam tests to be

$$
\left( \frac{\sigma_E}{E} \right)^2 = \left( \frac{2.8\% \sqrt{\text{GeV}}}{\sqrt{E \ [\text{GeV}]}} \right)^2 \oplus \left( \frac{0.12 \text{GeV}}{E \ [\text{GeV}]} \right)^2 \oplus (0.30\%)^2,
$$

in which $E$ is the electron energy. The first term refers to a stochastic component (e.g. the evolution of the shower), the second term describes noise (e.g. electronics or additional interactions), and the third term is an offset.

The Hadron Calorimeter  The HCAL (cf. [136]) consists of four subdetector components. A longitudinal view is shown in figure 3.6. The barrel ($|\eta| < 1.3$) and endcaps ($1.3 < |\eta| < 3.0$) are inside the solenoid. The outer calorimeter ($|\eta| < 1.3$) is located outside the solenoid coil. The HF covers particles with large pseudo-rapidity ($3.0 < |\eta| < 5.2$).

The HCAL is a sampling calorimeter consisting of several layers of brass absorbers and plastic scintillators. These are embedded in plates made of stainless steel. The steel plates are 40 mm and 75 mm thick, the brass plates are about 50 mm to 80 mm thick, and the scintillators are 4 mm and 9 mm thick. The thinner components are used in the barrel. The brass is composed of 70% copper and 30% Zinc, and is non-magnetic. It has a density of 8.53g/cm$^3$, an interaction length of 16.42 cm, and a radiation length of 1.49 cm. The total interaction length of the detector up to the HCAL corresponds to about $10 \lambda_l$. Hence, a high probability to absorb hadronic showers is given. The HCAL is further extended outside the solenoid, which is the next layer, with an outer calorimeter. The solenoid coil is used as an additional absorber, extending the interaction length to about $11.8 \lambda_l$. Wavelength-shifting fibers transport the collected scintillation light to photo-detectors. Those fibers have a diameter of about 1 mm. The photo-detectors are located far away behind the return yoke.

Particles at high pseudo-rapidity are much more energetic. Thus, the HF has to be much more resistant against radiation damage. Here, steel absorbers combined with quartz fibers as active material are used. The shower particles are detected by emitted Cherenkov light in the quartz fibers. Furthermore, information about the particle trajectory is provided [144].
Figure 3.6: Profile of one quarter of the CMS HCAL [136]. The HCAL is composed of four subdetectors. The barrel, endcaps, and Hadron-Forward Calorimeter (HF) as well as the outer calorimeter beyond the solenoid coil. The pseudo-rapidity \( \eta \) is covered up to \( |\eta| < 5.2 \).

The quartz-fibers have a core diameter of about 600\( \mu \)m. The HF is instrumented with about 1000 km of quartz-fibers altogether.

The overall energy resolution of the HCAL is measured with a pion test beam (cf. [145]) to

\[
\frac{\sigma_E}{E} = \left( \frac{110\% \sqrt{\text{GeV}}}{\sqrt{E \text{[GeV]}}} \right) \oplus 4.3\%. \tag{3.8}
\]

The first term refers to a stochastic component, and the second term is an offset. In fact, the total energy resolution of hadrons is always a combination of ECAL and HCAL resolutions. Hadrons that minimally interact with the ECAL material have a much better energy resolution, but the resolution significantly degrades if particles also interact with the ECAL material. The single-charged-particle response is tuned to test-beam data, but is reasonably described also in minimum-bias data of proton-proton collisions at \( \sqrt{s} = 7 \) TeV [146]. The HCAL is calibrated to an absolute scale measured with pions at an energy of 50 GeV. However, the overall response of the calorimeter is non-linear, and the energy scale of jets is subject to a dedicated jet-energy calibration. The latter is discussed in detail in section 3.3.5.

### 3.2.3 Superconducting Solenoid

A superconducting solenoid (cf. [136]) with a free bore encloses the HCAL. The magnet is designed to provide a magnetic-flux density of 4 T, but operated at 3.8 T in the first few years until the aging of the coil can be quantified [147]. The solenoid is 12.5 m long and has a diameter of about 6 m. The superconducting magnet is made of NbTi with a cold mass of 220 t. The cold mass is surrounded by a liquid-helium cryostat. The magnet is operated at a temperature of

\(^2\)Details are given in figure 1.25 of this reference.
about 4.7 K with a margin of 1.8 K, and has a nominal current of 19.14 kA when operated at 4 T. Both coil and cryostat are enclosed in a heavy flux-return yoke, which is made of iron and weighs about 10000 t.

3.2.4 The Muon System

The iron yoke also serves as a massive hadron absorber for the muon system (cf. [136]). The muon system is embedded in the return yoke and covers an area of about 25000 m². It consists of three gaseous subdetectors with different technologies (cf. figure 3.7), and is used for muon triggering, muon reconstruction, and muon identification. The three subdetectors are precisely aligned by an optical system.

DT chambers cover the barrel region with a pseudo-rapidity range of $|\eta| < 1.2$. In this region, the muon rates due to neutron-induced backgrounds are low, but also the magnetic field is relatively low and uniformly distributed. DT chambers are organized in four stations with over 172000 sensitive wires to measure the particle trajectory. Each of the first three stations has four chambers that measure $r$-$\phi$ coordinates, and four chambers that measure the $z$ coordinate. The last station consists of four chambers that measure only the $r$-$\phi$ coordinates. Each chamber consists of several layers with DTs. DTs are operated with an Ar/CO₂ (85%/15%) gas mixture. They have a maximum drift length of 2.1 cm, which corresponds to a drift time of 380 ns.

CSCs cover the pseudo-rapidity range of $0.9 < |\eta| < 2.4$. This region has a high and non-uniformly distributed magnetic field as well as high muon-background rates. CSCs have a
fast response time and allow for a fine segmented detection area, which is ideal also for muon triggering. A total of 468 CSCs cover an area of about 5000 m² with about 2 million wires.

The RPCs cover the pseudo-rapidity range of $|\eta| < 1.6$. RPCs are gaseous detectors in which the gas is separated by two parallel, resistive plates. The outside of the plates is coated with an electrically conducting material. RPCs provide an independent and “sharp” trigger of muons at even low transverse momenta. In particular, RPCs are well suited also for high particle rates, since they can be operated in dedicated configurations of the electrical field and gas mixture. In CMS, RPCs are operated in the so-called “avalanche mode”.

Muons are reconstructed using the combined information from the inner-tracking system and the muon system. The transverse-momentum resolution of the muon reconstruction is shown in figure 3.8. The momentum resolution of the muon system alone is about 8–15% for muons for with a transverse momentum of 10 GeV/c, and about 15–30% for $p_T = 1$ TeV. The combined $p_T$ resolution of both muon system and inner-tracking system is 1–2% for muons with a $p_T$ of 10 GeV/c, and about 4–10% for muons with $p_T = 1$ TeV. All momenta resolutions are functions of the pseudo-rapidity $\eta$ (besides the $p_T$ dependency), and they are more precise in the central region. The muon system significantly improves the combined momentum resolution for muons with $p_T > 100$ GeV/c. Muon identification, reconstruction, and momentum measurement as well as their performances are described further in section 3.3.3.

![Transverse-momentum resolution for different pseudo-rapidity ranges](image1.png)

Figure 3.8: Transverse-momentum resolution of the muon reconstruction for a pseudo-rapidity of $|\eta| < 0.8$ (left) and $1.2 < |\eta| < 2.4$ (right). This figure is taken from ref. [136].

### 3.2.5 The Trigger System

The inelastic proton-proton cross section is $\approx 100$ mb at a center-of-mass energy of 14 TeV, resulting in an event rate of about 1 GHz at a peak luminosity of $10^{34}$ cm$^{-2}$ s$^{-1}$ [136]. With a typical event size of at least 150 kb, as used for the offline high-level analysis, a huge amount of data would have to be processed by the CMS experiment. However, only an event rate of
3.2 The Compact Muon Solenoid Detector

$O(400\text{ Hz})$ can be processed by the computing facilities that are used for “online” processing and reconstruction of events. Hence, only a couple of “interesting” events can be recorded. This rate reduction is achieved by triggering events using typical signatures, i.e. a first event selection is done. Triggered data are reconstructed at central CERN computing facilities, and they are usually available for high-level analyses within 48 hours \[148\]. The CMS trigger system (cf. \[136\]) consists of two steps, the Level-1 Trigger (L1) with a design output rate of 100 kHz and access to limited detector data, and the High Level Trigger (HLT) with access to the full detector readout.

The L1 trigger inspects every bunch crossing, and has to initiate a readout of the detector-electronics within $3.2\mu s$. Therefore, L1 triggers are implemented as customized hardware triggers. Fundamental ingredients to the L1 trigger are either muon-trigger candidates from CSC, DT or RPC subsystems, or calorimeter deposits in ECAL, HCAL or HF.

The muon-trigger system starts from track segments and hit pattern in CSC and DT subsystems. It further provides an assignment of the muon-trigger candidate to a certain bunch crossing. Algorithms combine segments to tracks for both CSC and DT subsystems, while an individual trigger candidate is reconstructed with the RPC subsystem. A global-reconstruction algorithm combines the information of the three subdetector systems, and up to four muon candidates are passed to a global-trigger algorithm.

Calorimeter deposits are segmented into “trigger towers” with a size of at least $(\eta \times \phi) = (0.087 \times 0.087)$. Light-weight algorithms check the compatibility and isolation of calorimeter deposits with electron/photon, or muon signatures. For this reason, trigger towers are combined to dimension $4 \times 4$ in ECAL or HCAL, and $1 \times 1$ in the HF. The next layer of algorithms reconstructs jet-trigger candidates, jets from $\tau$ leptons, missing transverse energy, as well as simple global quantities such as jet counts, sum of transverse energies etc. The global quantities with up to eight electron/photon candidates, eight jet candidates, and four $\tau$-lepton candidates are passed to the global-trigger algorithm as well.

The global-trigger algorithm first ranks all reconstructed trigger objects. It takes the decision if an event passes the L1 trigger criteria according to a set of kinematics properties and quality criteria of the candidate physics objects, as well as global event quantities.

The HLT \[149\] is a software trigger that runs on common central processing units. It is fed with events that are accepted by the L1 trigger, and its purpose is to reduce the event rate from 100 kHz down to about 400 Hz. HLTs have access to the full event data, and they can be customized to a certain event signature. Several trigger paths are organized in trigger menus, and each trigger menu is adapted to a certain luminosity scenario. HLT menus can be used with high flexibility and can be switched even during a single run. This is useful since the instantaneous luminosity significantly decreases over the period of a run. “Pre-scaled” triggers with looser criteria only pick every $N$-th event, they are often used for monitoring purposes. Pre-scaled triggers are also useful for trigger-efficiency measurements. HLTs first access information that have a fast readout, i.e. the calorimeter and the muon system. Track reconstruction usually requires significant computing time due to the high track multiplicities. Therefore, tracking information is handled last.

Complementary trigger strategies are “data parking” \[148\] and “data scouting” \[148\] \[150\]. “Parked” datasets have dedicated trigger and data-streaming paths. Here, events which are streamed to “parked” datasets are not instantly reconstructed within a certain data-taking period, but stored on disk for later processing. The used triggers may have looser criteria than the triggers used for “online” reconstruction, or they can be completely orthogonal. The processing of events can be delayed until enough computing resources are in idle state, e.g. during
a shutdown period of the collider. In 2012, CMS attributed the same amount of bandwidth to parked datasets as for the prompt reconstruction.

The “data-scouting” strategy has the objective to monitor the non-triggered phase space, e.g. to search for signs of new physics, and to monitor the parked data. Dedicated triggers for such a dataset require criteria that are more loose than in the default trigger menu. The recorded data are processed with a condensed event content, which is restricted to a few (simple) quantities. Thus, the computing consumption at HLT level is minimized, but a high trigger rate can be facilitated. Event rates can be larger than one kHz. However, the reduced event content makes a more precise offline reconstruction mostly impossible, and “data scouting” is usually constrained to simple analyses. In 2011, the CMS collaboration utilized data-scouting for one analysis that searched for narrow resonances within the dijet-mass spectrum \[ L = 0.13 \text{ fb}^{-1} \] complemented the analysis strategy at low invariant masses.

### 3.2.6 Detector Simulation

An important part of the scientific work is an accurate prediction of the outcome of an experiment for a particular hypothesis. In high-energy physics, any physics analysis is designed by using simulated events at some step. An accurate prediction does not only involve the simulation of \( pp \) collisions, i.e. event generation, but also the precise modeling of particle interactions with the detector material and the detector response to a certain hit pattern. Simulations are in particular important for the calibration of physics objects, optimization of physics object reconstruction, and estimation of systematic uncertainties.

Events of the hard-scattering process in \( pp \) collisions are typically simulated with dedicated matrix-element generators at various orders of perturbative QCD calculation, e.g. MADGRAPH \[ 59 \] or POWHEG BOX \[ 60, 61 \]. Furthermore, general-purpose generators like PYTHIA \[ 63 \] or HERWIG \[ 64 \] exist. A general overview of the simulation of physics events is given in section 2.2.1, while the generated event samples as used in this analysis are summarized in 3.5.

For the CMS experiment, a full detector simulation based on GEANT \[ 151, 152 \] is used. In a first step, the beam profile and the luminous region are modeled according to realistic conditions. Then, the detector simulation describes the passage of particles through the detector material \[ 11 \] and their interactions with it. Physics effects such as energy loss and deflections to the particle trajectory due to multiple scattering effects are implemented. The detector simulation also describes the development of electromagnetic and hadronic showers. One of the two important ingredients to the detector simulation is a precise and complete model of the detector geometry. Such a model is needed in order to properly account for the material budget of the detector. The detector geometry includes information such as the dimensions, structure, and position of all elements, as well as their material type. The second important ingredient to the simulation is a precise and detailed map of the magnetic field of the CMS detector \[ 147 \]. The simulated event may be also superimposed with additional soft interactions which are generated e.g. with a dedicated PYTHIA simulation. The detector response with its limited resolution of subdetector components, noise modeling, calibration constants, electronics readout, reconstruction efficiencies etc. is modeled within a separate step of the simulation. This step is referred to as “digitization” (cf. \[ 153 \] for details).
3.3 Reconstruction of Physics Objects

Particle that are produced within the CMS detector acceptance are identified by their characteristic signatures. These signatures may be composed of tracks in the silicon-tracker that are bend by the magnetic field, energy deposits in a few or more crystals of the ECAL, energy deposits in the HCAL, a shower depth, a characteristic decay, tracks from secondary vertices for long-lived particles, or many other features. The identification of a particle type, the measurement of its energy and momentum is done by exploiting this characteristic signature. The reconstruction of physics objects from the CMS detector readouts is briefly described in the following section.

This section starts with a description of the vertex-reconstruction in section 3.3.1. Vertices are reconstructed from silicon-tracker tracks. They are important e.g. whether individual particles belong to the primary-scattering process. In particular, vertex reconstruction becomes challenging in an environment with many additional interactions due to the large instantaneous luminosity.

The PF algorithm is used as a tool for a global and non-ambiguous event reconstruction. It is described in section 3.3.2. The output of the PF algorithm is a list of all reconstructed stable particles in an event, similar to the list one obtains from simulated events.

Afterwards, the reconstructions of muons (sec. 3.3.3) and electrons (sec. 3.3.4) are discussed in detail. High-\(p_T\) leptons are expected from the W-boson decay. The reconstruction and identification of electrons and muons is an important instrument to identify electroweak top-quark production among the overwhelming number of background events without prompt charged leptons.

Another indicator for \(t\)-channel-top-quark production is the production of particle jets from (b-) quark fragmentation. Jets (section 3.3.5) are highly complex objects, and they are reconstructed using the individual, stable constituents of the PF-algorithm output. Jet-flavor tagging is an important tool to identify jets that originate from fragmenting b-quarks, as it is the case for the top-quark decay. Jet-flavor tagging is described in section 3.3.6.

Neutrinos leave undetected, but still carry a significant amount of momentum in case of \(t\)-channel events. Their transverse-momentum components are inferred from a momentum imbalance of all visible particles. The reconstruction of missing transverse energy (\(E_{\text{T}}^{\text{miss}}\)) is discussed in section 3.3.7.

3.3.1 Vertex Reconstruction

Vertex reconstruction includes the reconstruction of both primary vertices of proton-proton interactions and secondary vertices from long-lived particles such as b-flavored hadrons. The reconstruction of the primary vertex in an event is in particular helpful to check the compatibility of reconstructed physics objects, e.g. muons or electrons, with the hard-scattering process. The compatibility with the primary vertex provides also useful information to suppress contributions from additional interactions. Secondary-vertex reconstruction is used to improve jet-flavor tagging.

In the CMS experiment, vertex reconstruction is performed in two subsequent steps, namely vertex finding and vertex fitting. “Vertex finding” is a procedure to group tracks into a vertex candidate. Tracks are clustered according to their \(z\)-coordinates with the Deterministic-Annealing-clustering algorithm [154]. A linear dependence of the vertex-reconstruction effi-
ciency and the number of additional interactions is achieved. “Vertex fitting” refers to the
determination of the best estimate of vertex parameters of a particular vertex candidate, i.e.
a group of tracks, and the fit quality. Parameters include position coordinates and covariance matrix. A three-dimensional fit which is based on an “iterative re-weighted Kalman filter”, the Adaptive-Vertex Fitter, is used [155, 156]. Individual tracks are weighted according to their compatibility with the vertex candidate. The standard Kalman-filter technique assumes Gaussian-probability-density functions [157]. The adopted version of a Kalman filter is more robust against outliers than the standard Kalman filter, but takes all tracks into account.

Primary vertices are accepted if they are compatible with the beam spot, i.e. the luminous
region. The beam spot is reconstructed with a fit to all good silicon-tracker tracks; details of the
beam-spot reconstruction are given in ref. [158]. Typically, a number of additional interactions
occurs parallel to the hard-scattering process due to Pile-Up. Then, several primary vertices
are reconstructed in each event. The selection of a primary vertex that originates from the
primary hard-scattering process is part of the event selection as described in section 4.3. The
vertex-reconstruction efficiency for vertices of additional interactions approximately is 70%.
The reconstruction efficiency for primary vertices is close to 100% (cf. ref. [159]).

3.3.2 Particle-Flow Reconstruction

The Particle Flow (PF) event reconstruction [160] is the default event-reconstruction algorithm
for physics analyses carried out at the CMS experiment. It is described in detail in ref. [160]
and summarized in this section. This section starts with a brief motivation which explains the
importance of the PF event reconstruction for physics analyses. Afterwards, the PF algorithm
itself is described. The reconstruction of high-level-physics objects and corresponding system-
atic uncertainties are not part of this section, but are discussed in the following sections.

The PF algorithm performs a global event reconstruction in the following sense. First, it pro-
vides a complete list of reconstructed, identified, stable, and elementary particles in an event.
Second, it combines all available information from subdetectors for an optimal reconstruction
of particle momenta and directions. Third, it allows for a non-ambiguous event reconstruc-
tion. Elementary particles, which are reconstructed by the PF algorithm, are muons, electrons,
photons, charged hadrons, and neutral hadrons.

The output of the PF algorithm is a complete list of identified visible particles, similar to a list
one would obtain from simulated events. This list of stable particles serves as an input for algo-
rithms that reconstruct high-level objects and their properties. These include jet reconstruction
and identification of jet flavors, identification of τ leptons, calculation of momentum imbalance
and missing transverse energy, and Pile-Up compatibility criteria among others. Particle-based
isolation and a dedicated reconstruction of Bremsstrahlung photons improve identification of
prompt electrons and muons.

Reconstruction algorithms for high-level objects consecutively run one after another. Usually,
these algorithms are independent of each other. However, ambiguities in the global event

---

3 In this publication, the efficiency measurement is obtained with a different clustering method, namely the “gap-z clustering”. This vertex-finding algorithm groups together tracks which are separated in less than 1 cm distance in $z$ direction. However, the cited study was done with data recorded in the 2010 data-taking period with low Pile-Up conditions. Both clustering algorithms, gap z and Deterministic Annealing, perform similar for a low number of additional interactions.
reconstruction can arise during the reconstruction of individual high-level objects, i.e. a double-counting of measured particles in high-level objects. As an example, an electron can be reconstructed both as a jet or reconstructed as a final-state electron. Within the PF-event reconstruction, ambiguities can be resolved by removing dedicated reconstructed particles, e.g. well-isolated charged leptons, from the list of available particles for subsequent high-level-reconstruction algorithms.

Furthermore, the PF algorithm significantly improves the response and resolution of $E_T^{\text{miss}}$ \cite{161}. The $E_T^{\text{miss}}$ reconstruction benefits from the improved reconstruction of all individual particles it is composed of. Moreover, the PF algorithm facilitates the cleaning of calorimeter deposits and tracks due to Pile-Up contributions. The compatibility of charged tracks with the primary vertex is checked on an event-by-event basis, and corresponding calorimeter deposits due to additional interactions are subtracted. Sophisticated algorithms utilize the PF algorithm to reconstruct $\tau$ leptons. These algorithms use the list of reconstructed jet constituents to analyze all hadronic decay modes of the $\tau$ lepton individually \cite{162}.

In general, the PF algorithm profits from the good resolution of the large silicon tracker that provides, combined with the large magnetic field, a high tracking efficiency and low fake rates. It further profits from the granularity of the ECAL and pure muon reconstruction \cite{160}.

**Particle-Flow Algorithm** The basic ingredients of the PF algorithm are an iterative-tracking procedure for a track reconstruction with a small fake rate and a clustering of calorimeter deposits. Reconstructed silicon-tracker tracks, calorimeter clusters, and muon-subdetector tracks are combined into “blocks” by a linking procedure. As an example, a block for a charged hadron may consist of a silicon-tracker track, an ECAL cluster, and an HCAL cluster. Blocks are interpreted in terms of final-state particles, namely muons, electrons, photons, charged hadrons, and neutral hadrons. Furthermore, calorimeter clusters need to be calibrated for a precise response to photons and hadrons. This list of reconstructed stable particles is passed to algorithms that reconstruct high-level objects like jets.

**Iterative Tracking** Since an efficient, fast, and precise determination of tracks with even low momenta is important for the PF algorithm, an iterative-track reconstruction is developed \cite{160}. In a first step, tight seeding and identification criteria are applied, and a pure track reconstruction is achieved. In the next steps, hits assigned to the previous iterations are removed, and quality criteria are relaxed. This leads to a high reconstruction efficiency. A small fake rate is obtained since combinatorics are reduced. Further iterations relax constraints to take into account so-called “secondary tracks” from nuclear interactions, photon conversions, Bremsstrahlung, or long-lived particles. The former three processes are induced by the material budget of the tracker, which corresponds to radiation lengths of up to $2X_0$ and hadronic interaction lengths of up to $0.55\lambda_l$. The iterative-tracking procedure facilitates a track reconstruction down to a $p_T$ of 150 MeV/c with a fake rate at $O(1\%)$ \cite{160}.

**Calorimeter Clustering** Nearby calorimeter deposits are grouped into “calorimeter clusters”. The clustering algorithm runs in each subdetector of ECAL (barrel and endcap), HCAL (barrel and endcap), as well as first and second layer of the ECAL Preshower (PS). In the HF, each cell corresponds to a cluster. More details are given in ref. \cite{160}.
Linking of Tracks and Calorimeter Clusters  The “linking” of ECAL clusters, HCAL clusters, and tracks is described in the following (cf. [160]). Tracks are consecutively extrapolated from the last hit in the inner tracker to the PS to the ECAL within the typical electron-shower depth, and to the HCAL subdetectors within the typical interaction length of a hadron shower. Any cluster that matches the extrapolated track within its spatial boundaries is linked to a track. The linking algorithm further collects energy of Bremsstrahlung photons by considering tangents to extrapolated tracks. Furthermore, a PS cluster is linked to an ECAL cluster if the PS cluster is within \((\eta, \phi)\) boundaries of the ECAL cluster. ECAL clusters are linked to HCAL clusters accordingly. Muon tracks reconstructed in the muon-subdetector systems are linked to tracker tracks within a global-fit procedure (cf. section 3.3.3). Small blocks, which typically consist of a few elements, are reached due to the granularity of the subdetectors, such that the PF algorithm performs robust in even complex events.

Particle Reconstruction and Identification  Particle reconstruction and identification of each particle type is performed consecutively. First, well identified and isolated muons are reconstructed and removed from the list of input particles for subsequent reconstruction algorithms. An efficient selection with a low fake rate is used to avoid any bias on jet or \(E_{\text{miss}}\) reconstruction. Second, electrons are reconstructed. PF-based-selection criteria include a dedicated conversion rejection, electron-identification criteria, as well as particle-based isolation. Especially for electrons, an identification of tracks from photon conversions and Bremsstrahlung photons is taken into account. The PF-based electron and muon reconstruction is commissioned with data in J/\(\Psi\) events for low-\(p_T\) leptons and in W-boson events for high-\(p_T\) leptons [163]. The reconstruction of muons is explained in section 3.3.3, and the reconstruction of electrons is explained in section 3.3.4. More details on the PF-specific-lepton reconstruction can also be found in ref. [163].

Energies and directions of charged hadrons are reconstructed from silicon tracks which can be linked to ECAL clusters and HCAL clusters. Single tracks which cannot be linked to any cluster are interpreted as charged hadrons as well. In this case, the track is assumed to stem from a charged pion [160]. Photons and neutral hadrons are reconstructed from calorimeter clusters which cannot be linked to tracks. Photons only deposit energy in the ECAL. Remaining ECAL clusters without linked tracks are assigned to photons, and remaining pure HCAL clusters (or HCAL clusters which are linked to ECAL clusters) are assigned to neutral hadrons. Photons and neutral hadrons can even be reconstructed if their energy deposits superimpose or merge with that from charged hadrons. An energy excess in one or both calorimeters can be determined by checking the compatibility of the calorimetric energy with that from their linked tracks. If either HCAL or ECAL cluster deposits are incompatible (within the calorimeter resolution) with the momentum of the linked track, all tracks associated to charged hadrons are subtracted. Depending on the typical footprint of HCAL and ECAL deposits of photons and neutral hadrons, the remaining energy is assigned to either photons or charged hadrons.

Calorimeter-Cluster Calibration  The reconstruction of photons and hadrons relies on a well-calibrated response of the calorimeter system. Energy thresholds are applied to ECAL crystals in order to reduce noise. Therefore, a residual calibration, which is determined from simulation, is applied to reconstructed photons (cf. [161]). The photon-energy scale is checked with \(\pi^0\) events in data and found to be well understood [161]. Hadrons deposit energies in both ECAL and HCAL. The ECAL response to neutral hadrons, however, is different from its response to photons, and the HCAL response to hadrons is nonlinear and calibrated to 50 GeV pions that solely interact in the HCAL [160]. Therefore, the hadron energy is inferred with an
3.3 Reconstruction of Physics Objects

η-dependent linear calibration function of both ECAL and HCAL energy deposits \[160\]. The same calibration coefficients are assumed for charged and neutral hadrons. The calibration of the hadron response is confirmed in minimum-bias collisions at \( \sqrt{s} = 7 \) TeV up to a few percent \[161\]. Furthermore, residual calibrations are applied to reconstructed jets as discussed in section 3.3.5. Cleaning procedures for calorimeter noise are described in ref. \[161\].

**Charged Hadron Subtraction (CHS)** The PF algorithm can be used to clean Pile-Up contributions on particle level. The algorithm removes charged tracks from an event if they cannot be matched to the primary vertex. CHS is applied on an event-by-event basis.

### 3.3.3 Muons

An efficient and precise reconstruction and identification of muons is an essential element of the CMS physics program. Muons provide a clear and powerful signature for many SM processes and new physics scenarios. In the multi-GeV regime, they are minimally ionizing the inner-tracker material and calorimeters, and they are the only particles that penetrate through the massive steel yokes of the solenoid into the muon-subdetector system; except for weakly-interacting particles such as neutrinos, which leave undetected at all. This section is organized as follows. First, muon reconstruction and muon identification are described. An important criterion to identify muons is their isolation from hadronic activity around their trajectory. Afterwards, the performance of the muon reconstruction in data is compared to that from simulation.

**Muon Reconstruction and Identification** Tracks are reconstructed in both the silicon tracker and muon-subdetector system with at least one valid hit. In the following, the former are referred to as silicon-tracker tracks, and the latter as stand-alone-muon tracks. Two complementary reconstruction algorithms exist to combine the information from tracker and muon-systems to suppress cosmic muons, muons from hadron decays, or hadron punch-through.

In the global-muon reconstruction, a trajectory for a muon candidate is obtained by fitting each stand-alone-muon track to all available silicon-tracker tracks. The fit is done using the Kalman-filter technique \[164\]. If a good quality is obtained, the muon candidate is referred to as a “global muon”. The global-muon reconstruction typically requires segments in at least two muon stations. In the tracker-muon reconstruction, silicon tracks are extrapolated to the muon-subdetector system. If the extrapolated silicon track can be matched to segments in the muon system within distances of a few cm, it is considered as a tracker-muon candidate. The tracker-muon reconstruction is in particular useful for low-\( p_T \) muons, since only one muon station is required here. In this analysis, only muons that are reconstructed as both global and tracker muons are used. Furthermore, dedicated algorithms exist that optimize the reconstruction of muons with \( p_T > 200 \) GeV/\( c \), since radiation losses in the return yoke can be crucial for high-\( p_T \) muons. A detailed description is given in ref. \[163\].

Muons candidates during the PF reconstruction are required to have a transverse momentum of at least 5 GeV/\( c \), to be well isolated, and reconstructed as global muons (cf. \[163\]). Only loose identification criteria are applied in this first step in order to maximize the reconstruction efficiency for the high-level analysis. Additional reconstruction efficiency is harvested by applying looser isolation criteria to the remaining muon candidates, but instead tighter identification criteria. Furthermore, it is checked if the energy footprint of the muon candidate is compatible with a typical signature of a muon from a hadron decay. Compatibility with the
primary-hard-scattering process is achieved by requiring a distance in the transverse plane of less than 0.02 cm w.r.t. the (average) beam spot, as well as less than 0.5 cm in z-direction.

Further tight identification criteria on the fit quality and the muon-hit pattern are applied during the offline event selection. A momentum threshold of at least $p_T > 20 \text{ GeV/c}$ is used to match offline-reconstructed muons to trigger criteria. The full event selection is discussed in section 4.3.

Muons Isolation In order to discriminate between muons from the primary-hard-scattering process and muons from hadron decays, muons are required to be isolated both in the high-level analysis and in the PF-event reconstruction. The relative particle-based isolation $\text{relIso}$ is defined as

$$\text{relIso} = \frac{p_T^\mu}{\sum_{\Delta R<0.4}^\text{charged hadrons} p_T} + \sum_{\Delta R<0.4}^\text{neutral hadrons} E_T + \sum_{\Delta R<0.4}^\text{photons} E_T,$$

in which the sums run over all PF-reconstructed particles within a cone $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} < 0.4$ around the muon trajectory. In order to make the isolation robust against Pile-Up contributions, charged hadrons are excluded from the isolation calculation if they cannot be associated to the leading primary vertex of the event. The obtained isolation is corrected for neutral hadrons from Pile-Up, assuming a ratio of 0.5 between charged and neutral hadrons. Furthermore, a veto cone within $\Delta R < 0.01$ around the muon trajectory, and a transverse-energy threshold of $E_T > 0.5 \text{ GeV/c}$ are applied for neutral hadrons and photons. Thus, deposits due to ionization and radiation are removed from the isolation calculation. Typically, a muon is referred to as “isolated” if the transverse energy deposits and track-momenta around its trajectory are less than 15%.

Performance of the Muon Reconstruction Muon-momentum scale and resolutions are found to well agree between data and simulation [165]. The total muon-reconstruction efficiency $\epsilon_\mu$ is factorized into four components (cf. [164])

$$\epsilon_\mu = \epsilon_{\text{track reconstruction}} \times \epsilon_{\text{reconstruction and identification}} \times \epsilon_{\text{isolation}} \times \epsilon_{\text{trigger}}.$$ (3.10)

The total muon-reconstruction efficiency in data is about 84% within $|\eta| < 2.1$ [164]. The reconstruction efficiency depends on $\eta$, but mostly is independent of $p_T$. Each efficiency component is individually measured in data and compared to simulation. Residual corrections to the simulated efficiencies are applied in terms of data-simulation-scale factors. Uncertainties on the performance and the measurement technique are taken into account within the statistical inference of this analysis (cf. sections 4.6 and 6).

Efficiencies are typically measured in dedicated analyses using the tag-and-probe method [165] in events that are enriched in dimuon resonances, e.g. $J/\Psi \rightarrow \mu^+\mu^-$ or $Z \rightarrow \mu^+\mu^-$ events. Here, an identified muon with tight identification criteria is referred to as a “tag”. Tight criteria are applied to obtain a pure event sample. A muon candidate without (or with loose) identification criteria is referred to as a “probe”. In particular, the probe is selected without using any criteria that are being tested for. Usually, events are selected, in which both tag and probe yield a reconstructed invariant mass within a certain window around the resonance mass. Thus, the probe is expected to also stem from the resonance. It is tested if the probe passes or fails the tight reconstruction or identification criteria, i.e. the part of the efficiency that is being studied. Thus, an efficiency for a certain criterion can be calculated by looking at many events. The tag-and-probe method can be further applied twice to the same event sample to avoid any bias.
3.3 Reconstruction of Physics Objects

on the procedure, e.g. by using once positively charged muons as tags, and once negatively charged muons, while requiring tags and probes to have opposite charges.

The track-reconstruction efficiency is measured in $J/\Psi \rightarrow \mu^+\mu^-$ and $Z \rightarrow \mu^+\mu^-$ events. It is about 99% in both data and simulation [166]. The agreement between simulation and data is better than 1% [165][166].

Reconstruction and identification efficiencies are about 96% in both data and simulation [164]. The isolation efficiencies are about 97% in data and simulated events [164]. The quoted efficiencies are averaged over the full pseudo-rapidity range ($|\eta| < 2.4$). Reconstruction and identification efficiencies, as well as the isolation efficiency, are measured also in Z-boson-plus-jets events for exactly the same selection criteria as used in this analysis [167]. The reconstruction, identification, and isolation of muons are found to be well simulated in Z-boson-plus-jets events, and no additional scaling of simulated events needs to be applied. While the statistical uncertainties for the scale factor are negligible, the systematic uncertainties are 3%. The systematic uncertainties mainly cover the extrapolation of scale factors from the measurement with Z-boson-plus-jets events with a low jet-multiplicity to the analysis phase space with a higher jet multiplicities.

The trigger efficiency is about 91% in data events [164]. The trigger efficiency used in this analysis is found to be flat in muon $p_T$. However, it is expressed as a function of the pseudo-rapidity $\eta$ [167], since the three different technologies of the muon system cover different ranges in $\eta$. The obtained data-simulation-scale factors are up to $\mathcal{O}(5\%)$ at $1.6 < |\eta| < 2.1$ and of $\mathcal{O}(-2\%)$ within $|\eta| < 1.6$. Uncertainties on the scale factors are well below 1%.

Figure 3.9 shows the resulting invariant-dimuon-mass spectrum in data corresponding to an integrated luminosity of 40 pb$^{-1}$ [164]. Two muon candidates are required to trigger an event. Staggered muon-$p_T$ criteria, which are adopted to the instantaneous luminosity conditions, are used in the trigger; the various trigger thresholds are also visible in fig. 3.9 (left plot). Dimuon resonances are resolved on top of the overwhelming Drell-Yan spectrum over about three orders of magnitude down to a few hundred MeV/$c^2$. The muon system is able to discriminate among the $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ mesons (right plot). A mass resolution of about 100 MeV/$c^2$ is achieved within the full pseudo-rapidity range of $|\eta(\mu)| < 2.4$ (right plot), and an even better mass resolution of 67 MeV/$c^2$ is achieved in the central region within $|\eta(\mu)| < 1.0$ (cf. [164]).

Figure 3.9: The left plot shows the invariant-dimuon-mass spectrum as obtained with loose dimuon triggers in data corresponding to an integrated luminosity of 40 pb$^{-1}$. The right plot highlights the reconstruction of the $\Upsilon(1S)$, $\Upsilon(2S)$, and $\Upsilon(3S)$ mesons with a mass resolution of about 100 MeV/$c^2$ within the full pseudo-rapidity range of $|\eta(\mu)| < 2.4$. Both figures are taken from ref. [164].
Therefore, precise \( t\bar{t} \) analyses require at least one W boson which decays into either a muon or an electron. Furthermore, for single-top-quark events with hadronically decaying W bosons, the QCD-multijet and W-boson-plus-jets processes are by far dominant. Thus, the reconstruction and identification of electrons from leptonically decaying W bosons is an important and complementary addendum besides the reconstruction of muons.

### Electrons

Electrons are an important part of the event signature for many processes. The electron reconstruction and identification is in general more challenging, i.e. less efficient and less pure, than it is for muons. Higher \( p_T \) thresholds are required already at the trigger level, and the high-level analyses usually have a smaller acceptance in the electron decay channels. The reason for that is twofold.

**First**, electrons radiate much more Bremsstrahlung \( \mathcal{O} \left( \frac{m_e^2}{m^2} \right) \) \(^4\) compared to muons with the same momentum. Bremsstrahlung is already induced by the silicon-tracker material, which corresponds to radiation lengths of up to \( 2X_0 \) at \( |\eta| \approx 1.5 \). Furthermore, the tracker-material budget is \( \eta \)-dependent. ECAL deposits of radiated photons are spatially separated from the electron trajectory due to the large magnetic field of the CMS experiment. Moreover, the electron trajectory itself is distorted due to radiated photons and the large magnetic field. Instead, the ionization loss is comparable for electrons and muons at high momenta (\( > 10 \text{ GeV}/c \)).

**Second**, electrons deposit energy in the ECAL, which is located in front of the magnet yoke and in front of the HCAL. Fake electrons can be induced if energy deposits of charged hadrons, neutral hadrons, or photons are accidentally matched with any reconstructed track. In addition, “real” electrons can also stem from hadron decays, and are a background to prompt electrons from the hard-scattering. Furthermore, electron-positron pairs are induced by photon conversions.

In this section, electron reconstruction and identification are described first. Then, electron isolation and the performance of the electron reconstruction are discussed.

#### Electron Reconstruction and Identification

Electron trajectories are quite “distorted” due to photon radiation in combination with the large magnetic field of the CMS detector. A Gaussian-Sum Filter (GSF) \([157]\) is used to reconstruct the tracks of electrons. The GSF is a generalized Kalman Filter that is able to cope with non-Gaussian probability-density functions. Here, the electron-energy loss due to Bremsstrahlung is described as a sum of Gaussian distributions. Hence, the electron trajectory, which is defined by a vector containing momentum, direction, and position at a reference point, can be described as a mixture of many single Gaussian distributions, i.e. an ensemble of Kalman filters.

The GSF-track-reconstruction algorithm is seeded in two different ways. An ECAL-driven seed uses energy deposits in the ECAL, the so-called “superclusters”. The ECAL-driven seed is in particular useful for high-\( p_T \) electrons. A tracker-driven seed is build from silicon-tracker tracks with certain properties. The ECAL-driven seed is in particular useful for low-\( p_T \) electrons. In the following two paragraphs, the building of seeds is briefly explained.

**ECAL-driven-seed reconstruction \([163]\)** proceeds as follows. Electrons or photons typically deposit their energy in a few ECAL crystals. A “basic” squared block of crystals, which is

---

\(^4\)This estimate does not take into account that also the ionization loss is different between muons and electrons. These differences become important at high momentum scales.
3.3 Reconstruction of Physics Objects

defined in the $\eta$-$\phi$ plane, is a good estimator for the electron or photon energy. Basic-cluster reconstruction is seeded by local maxima of energy deposits in the ECAL. However, the electron trajectory is distorted due to photon Bremsstrahlung and the large magnetic field of the CMS detector. In order to take radiated energy into account, so-called “superclusters” are built from blocks of basic clusters in $\phi$-direction.

The pre-defined width of the superclusters is, however, not optimal for low-$p_T$ electrons. Radiated Bremsstrahlung photons of low-$p_T$ electrons are spread over a wide $\phi$-range as the electron trajectory is highly bent due to the magnet field. Thus, the tracker-driven-seed reconstruction [160, 163] starts with a subset of all tracks that are likely to be compatible with electrons. Electron tracks are efficiently selected, but the subset also contains many tracks from charged hadrons. Tracks are extrapolated to the ECAL and matched to ECAL deposits. Energy from Bremsstrahlung photons is collected by considering also tangents to extrapolated tracks, which are expected to match the trajectories of radiated photons, and matching those “tangents” to the ECAL clusters as well.

Both seed collections are combined and passed to the GSF-track reconstruction. Duplicates are cleaned. Rejection criteria for photon conversions are applied. The electron candidate is required to be isolated with deposits less than 15% relative to its $p_T$. Only electron candidates with $p_T > 5$ GeV/$c$ are considered as “real” electrons by the PF algorithm. Finally, electrons that passed all criteria are cleaned from high-level algorithms like jet reconstruction, and are considered as electron candidates for the (tighter) high-level-electron selection as described in the section “event selection” (4.3). Electrons in the transition region between ECAL barrel and ECAL endcap, which corresponds to $1.44 < |\eta| < 1.57$, are excluded from this analysis.

**Electron Isolation** The isolation calculation for electrons is similar to the calculation for muons, which is described in the previous section 3.3.3. The calculation differs in two criteria. First, there aren’t any Pile-Up corrections for electrons. Second, deposits of neutral hadrons and photons are vetoed with transverse energies $E_T \leq 0.5$ GeV/$c^2$, but they are applied independently of their spatial distance to the electron trajectory.

**Performance of the Electron Reconstruction** Electron-momentum scale and response are measured with resonances of Z bosons, J/$\Psi$ mesons, and $\Upsilon$(1S) mesons and found to be well reproduced in data [168, 169]. The electron-reconstruction efficiency is factorized into three components,

$$\epsilon_e = \epsilon_{\text{GSF-electron reconstruction}} \times \epsilon_{\text{identification and isolation}} \times \epsilon_{\text{trigger}}.$$  

(3.11)

All efficiency terms are measured in simulation and data with the tag-and-probe method as described in section 3.3.3. The GSF-reconstruction efficiency is defined as the efficiency to reconstruct a GSF-electron candidate from an ECAL cluster within the detector acceptance. For electrons in data, the efficiency is about 97% in the barrel, and 94% in the endcap [165]. The simulated efficiency in both ECAL barrel and ECAL endcap agrees very well with the efficiencies as observed in data [165]. A working point with an electron-identification efficiency of 70% is used. Identification and isolation efficiencies are measured in ref. [167] for the same selection criteria as used in this analysis. The obtained scale factors between simulation and data are very close to unity and neither a $p_T$-dependence nor an $\eta$-dependence are observed. However, the systematic uncertainties are 3% and cover the extrapolation of scale factors from the measurement with Z-boson-plus-jets events with a low jet-multiplicity to the analysis phase space with higher jet multiplicities. Electron-trigger efficiencies are measured in a sample that is enriched in Z bosons and are about 98% in the barrel and 97% in the endcaps [165]. The obtained
data-simulation-scale factors are flat in \( p_T \) and \( \eta \) \cite{165}. The measured trigger efficiencies well agree with the simulated efficiencies, and uncertainties of the data-simulation-scale factors for the trigger efficiency are below 1%. The total electron-selection efficiency is about 67%.

3.3.5 Jets

Partons from the hard-scattering process or radiated quarks hadronize according to the QCD theory, since free partons cannot exist due to the color confinement. However, the top quark has a lifetime \( \tau_t \approx (0.5 \times 10^{-24}) \, \text{s} \) \cite{11} smaller than the typical QCD hadronization timescale. Due to their relatively short lifetime, top quarks decay before they hadronize into bound states of subatomic particles. In this analysis, the dynamics of parton hadronization are modeled with the PYTHIA \cite{63} generator based on the Lund string model. As a result of the hadronization process, many stable particles like charged and neutral hadrons are produced. The mapping of multiple stable particles, which result from the hadronization process, to the original parton is subject to jet-clustering algorithms. Jet-clustering can be understood as a matching of experimental observations to theory predictions that are formulated on parton level. Typically, the direction and energy of reconstructed jets is related to the original (fragmenting) parton. However, non-linear responses of the calorimeters, additional interactions, and noise effects require a jet-calibration of reconstructed jets.

Approximately 65% of the jet energy is carried by charged hadrons, 25% by photons, and 10% by neutral hadrons \cite{160,161}. Since the PF-event reconstruction is able to distinguish between these types of particles, an improved jet reconstruction is possible. The charged-hadron momentum resolution is greatly improved when combining HCAL calorimeter deposits with the tracker information. Photons can be separated from charged-hadron-energy deposits such that also jet-energy reconstruction profits from the high granularity and energy resolution of the ECAL. Thus, about 90% of the typical jet-energy deposits are reconstructed with improved resolution when using the PF algorithm, leading to smaller particle-level-correction coefficients and jet-energy-scale uncertainties than conventional jet-reconstruction algorithms.

Jet-clustering algorithm

The anti-\( k_t \) algorithm \cite{170} with a distance parameter of \( R = 0.5 \) is used as the jet-clustering algorithm. The algorithm runs on the list of stable, elementary particles which is obtained by the PF algorithm.

The following definition of the anti-\( k_t \) algorithm is given in ref. \cite{170}. Constituents which have the smallest distance \( d_{i,j} \) among each other are consecutively recombined. Those constituents \( i \) and \( j \) include particles as well as pseudo-jets, and jet-clustering of a particular jet continues until \( d_{i,B} \) is the smallest distance, whereas \( d_{i,B} \) is between reconstructed jet \( i \) and the beam \( B \). If a jet is found, particles related to that jet are removed from the list and the clustering algorithm continues. The distance measure \( d_{i,j} \) scales the transverse momentum relative to the geometrical distance and is defined as

\[
d_{i,j} = \min(k_{t,i}^{2p}, k_{t,j}^{2p}) \frac{\Delta_{i,j}^2}{R^2} \tag{3.12}
\]

and

\[
d_{i,B} = k_{t,i}^{2p} \tag{3.13}
\]

in which \( \Delta_{i,j}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \) with rapidity \( y \), azimuth \( \phi \), and transverse momentum \( k_{t,i}^{2p} \) of particle \( i \). For the anti-\( k_t \) algorithm, \( p = -1 \) is chosen as the distance parameter.

The anti-\( k_t \)-clustering algorithm provides infrared- and collinear-safe jets \cite{170}, which means that the number of hard jets is insensitive w.r.t. the addition of new soft particles or collinear
3.3 Reconstruction of Physics Objects

splittings \[171\]. Moreover, the resulting jet boundaries are robust against soft-radiation effects, i.e. jets reconstructed with the anti-\(k_t\) algorithm usually have a regular shape with a circular cone with radius \(R\) if they are not too soft \[170\].

A potential ambiguity in the event reconstruction arises if electrons or muons from subsequent hadron decays are used once within the jet clustering and once again reconstructed as charged leptons. This ambiguity is resolved by using a relIso criterion during the PF-event reconstruction. Well-isolated charged leptons are excluded from the input-particle list of the jet-clustering algorithm and reconstructed as prompt, charged leptons, and vice versa. Reconstructed \(\tau\)-lepton candidates and photons are clustered always into jets in this analysis.

Furthermore, charged hadrons are subtracted from the input-particle list of the jet-clustering algorithm if their tracks are not compatible with the primary vertex of the event. The vertex-compatibility criterion aims at suppressing deposits from additional interactions (Pile-Up) that cannot be related to partons from the hard-scattering process. Remaining energy deposits from neutral hadrons are subtracted during jet-energy calibration.

![Reconstructed Jets - Detector Level -](image1)

![Calibrated Jets - Particle Level -](image2)

Figure 3.10: Sequence of jet-energy corrections that are applied to relate the energy of reconstructed detector jets to corresponding particle jets. First, jet-energy corrections are derived from simulated events. Second, residual corrections are derived from data and are applied on top of the previous corrections.

**Jet-Energy-Calibration Strategy** On detector level, jet reconstruction uses combined information of calorimeter and tracker with the PF algorithm (sec. 3.3.2). The jet-clustering algorithm is applied to the list of stable, elementary particles, and a detector-jet-four-momentum vector \(P_{\text{jet}}^{\text{raw}}\) is obtained. On particle level (or generator level), jets are clusters of stable particles which stem from the fragmentation process of a parton. Their four-momentum is referred to as \(P_{\text{jet}}^{\text{true}}\) in the following. Technically, the same jet algorithm is used for jet-clustering on both detector and particle level. However, the energy of a detector jet cannot be directly related to the jet energy on particle level due to non-linear responses of the calorimeters, additional interactions, and noise effects. Jet-energy calibration is needed to translate, on average, the jet reconstructed on detector level to a particle-level jet. The jet-energy calibration is often also referred to as “jet-energy scale”.

The CMS experiment uses a multiplicative approach to account for jet-energy calibration,

\[
P_{\text{jet}}^{\text{true}} = C(p_{\text{T}}^{\text{raw}}, \eta) \cdot P_{\text{jet}}^{\text{raw}},
\]

in which \(C(p_{\text{T}}^{\text{raw}}, \eta)\) refers to the \(p_T\)- and \(\eta\)-dependent corrections, which are applied to every component of the four-momentum \[71\]. The jet-energy corrections \(C\) itself are factorized further into four sub-components,

\[
C(p_{\text{T}}^{\text{raw}}, \eta) = C_{\text{Offset}}(p_{\text{T}}^{\text{raw}}) \cdot C_{\text{Simulation truth}}(p_{\text{T}}, \eta) \cdot C_{\text{Relative}}(\eta) \cdot C_{\text{Absolute}}(p_{\text{T}}'),
\]

namely the Pile-Up-and-noise-offset correction, the simulation-truth calibration, the \(\eta\)-dependent residual relative correction, and the residual \(p_T\)-dependent absolute correction \[71\].

59
3 Experimental Setup

While all four corrections are sequentially applied to reconstructed jet four-momenta in data, only the former two corrections are applied to reconstructed jets in simulation.

The offset correction removes energy from Pile-Up contributions, electronics noise, and underlying event. The offset due to deposits from additional interactions is corrected for with the jet-area method [71]. The offset linearly depends on the number of reconstructed vertices and typically is of the order of a few GeV/c [172]. The offset from electronics noise is approximately 250 MeV/c [71].

The simulation-truth calibration corrects the reconstructed detector-jets back to particle-level jets using information from simulated events. An average correction is derived in QCD-multijet events, which are simulated with the PYTHIA generator [71]. Here, detector-level jets are matched to particle-jets within a $\Delta R < 0.25$ cone, and the average response is expressed as a function of $p_T$ and $\eta$. The calibration factor can be quite large with up to $O(20\%)$ [71] in the transition region between barrel and end-cap, while being much less elsewhere. The parton-flavor composition of low-$p_T$ jets from the QCD-multijet sample is dominated by jets from gluon fragmentation. Jets from fragmenting quarks usually have higher transverse momenta. The energy response is expected to be dependent on the flavor of the fragmenting parton due to diversified particle-multiplicity patterns, as well as different energy spectra. A small flavor dependence is indeed observed at a level of $O(3\%)$ for jets with $p_T > 10$ GeV/c within the barrel region ($|\eta| < 1.3$). The response significantly profits from the precise charged-particle-momentum resolution of the PF algorithm. Simulations with PYTHIA show that the response among different jet-flavors is enveloped by the response of light jets and gluon jets [71, 172]. Jets from c- or b partons lie in between those two extremes. Furthermore, a comparison between PYTHIA and HERWIG++ simulations shows that the light-jet (u, d, s partons) and gluon-jet responses marginally depend on the fragmentation model. The difference in responses is $O(1\%)$ for a jet with $p_T > 30$ GeV/c within the central region ($|\eta| < 1.3$) of the detector (cf. [71]).

However, small differences are found when comparing the simulated jet response with the jet response measured in data. Thus, empirical corrections, the so-called “residual jet-energy corrections”, are applied in addition to the jet-energy corrections obtained from simulation. These corrections are applied to jets in data only. Residual jet-energy corrections (cf. [71]) are derived in measurements with data corresponding to an integrated luminosity of 4.7 fb$^{-1}$ at $\sqrt{s} = 7$ TeV, and applied to reconstructed jets which are used in this analysis. These corrections include an $\eta$-dependent relative correction and a $p_T$-dependent absolute correction.

The $\eta$-dependent correction is derived in dijet events. Here, the conservation of transverse momentum is used. The relative response of a (probe) jet at an arbitrary pseudo-rapidity $\eta$ w.r.t. a jet within the central region of the detector ($|\eta| < 1.3$) is measured in events in which both jets are back-to-back in azimuth $\phi$ [71]. The $\eta$-dependent correction mostly affects the transition region between barrel and end-cap with scale factors up to $O(15\%)$ [71].

The $p_T$-dependent absolute correction is determined in $\gamma$-plus-jet and Z-plus-jet events with leptonically decaying Z bosons, since their $p_T$ response is precisely known from the ECAL tracker, or muon subdetectors [71]. Both processes provide complementary information since they cover varied $p_T$ ranges, use different subdetectors with diversified resolution patterns, and have different production cross sections, which means varied trigger requirements and data-taking periods. A $p_T$ balancing between the jet and the $\gamma$ or Z boson is used, in which a central ($|\eta| < 1.3$) jet is required. Isolation criteria are used to reduce effects due to initial- and final-state radiation. A good agreement between simulation and data is observed, resulting in a small $p_T$-dependent absolute correction of approximately 1%.
Jet-Energy-Scale Uncertainties   The total uncertainties of the jet-energy calibration (fig. 3.11) typically are of $O(3\%)$ for $\text{PF}$ jets with $p_T = 30\text{GeV}/c$, and $O(1\%)$ for $\text{PF}$ jets with $p_T = 100\text{GeV}/c$ [172]. The total uncertainties are much larger (up to 5\%) in the barrel-endcap-transition region due to an observed instability of the derived jet-energy corrections with increasing run numbers (“time stability”). This instability is expected to be caused by radiation damage to the HF and transparency loss of the ECAL crystals that is not yet corrected for. The most important uncertainty at low $p_T$ is the uncertainty of the Pile-Up correction. At medium-jet-$p_T$ range, the most important uncertainty is due to an altered response of quark and gluon jets in different fragmentation models, which estimated by comparing PYTHIA and HERWIG++ simulations. At high jet-$p_T$, data statistics are limited. An extrapolation of the single-particle response and fragmentation modeling is done by combining information from simulation and data [71]. The uncertainty due to this extrapolation is the most important contribution for jets with large transverse momenta. In conclusion, the uncertainties of the jet-energy calibration are rather small when $\text{PF}$ particles are used as input to the jet-reconstruction algorithm.

![Figure 3.11: Uncertainties on jet-energy calibration expressed as a function of the jet $p_T$ for jets at $\eta = 0$ (left) and as a function of jet $\eta$ for jets at $p_T = 100\text{GeV}/c$ (right). Both figures are from ref. [172].](image)

Correction of Jet-Transverse-Momentum Resolution   The resolution of transverse momenta of jets is found to be lower in data than in simulated events. The conservation of transverse momenta is utilized in dijet and $\gamma$-plus-jet events [71] to derive scale factors that correct the jet-$p_T$ resolution in simulated events. The transverse-momentum resolution of each jet is corrected for by scaling the reconstructed jet $p_T$ with the difference between reconstructed jet $p_T$ and matched generator-jet $p_T$, where the difference is multiplied with a certain correction factor. The resolution correction depends on the pseudo-rapidity $\eta$ of the jet, and is determined by the CMS collaboration (based on the methodology in ref. [71]) to

- $1.05 \pm 0.06$ for jets within $|\eta| \leq 0.5$,
- $1.06 \pm 0.06$ for jets within $0.5 < |\eta| \leq 1.1$,
- $1.10 \pm 0.06$ for jets within $1.1 < |\eta| \leq 1.7$,
- $1.13 \pm 0.10$ for jets within $1.7 < |\eta| \leq 2.3$,
- and $1.29 \pm 0.20$ for jets within $|\eta| > 2.3$. 
Identification of Jets from b-Quark Fragmentation

Many SM processes include the production of a final-state-b quark. In particular, the top quark almost exclusively decays into a W boson and a b quark. A final-state-b quark is also an important signature for searches of various beyond-the-SM processes. But even more SM processes with much larger cross sections do not include b-quark production. Jet-flavor-tagging is an important tool to identify a particular event signature, but also to efficiently suppress background processes.

In the CMS experiment, jets originating from b-quark fragmentation are tagged using dedicated algorithms that exploit the semi-leptonic-decay mode of b-flavored hadrons, utilize the long lifetime of B hadrons, or combine information from both classes. The former algorithms use kinematics and track properties of muons or electrons to distinguish among jet flavors. The second class of algorithms benefits from the long lifetime of B hadrons which typically is about $1.6 \times 10^{-12}$ s [11]. The B hadrons have flight distances of a few millimeters which can be measured with the CMS tracking system. Algorithms may calculate impact parameters of charged-particle tracks w.r.t. the primary vertex, check the compatibility of all jet tracks with the primary vertex, or reconstruct displaced secondary vertices. The B-hadron-flight distance follows an exponential distribution. As an example, the average of the B-hadron-flight distance from a top quark at rest is

$$l = c(\beta\gamma)_{\tau} = \left(\frac{p}{m_b}\right)\tau = \frac{70}{5}c\tau = 6.7 \text{ mm.}$$

Here, the momentum is given by $p = m_0\gamma v = m_0 c\gamma \beta \Leftrightarrow \gamma \beta = \frac{p}{m_0 c}$, a b-quark mass of $m_b = 5 \text{ GeV}/c^2$ is used, and an average b-quark $p_T$ of 70 GeV/c is used, since the b quark is boosted due to the large top-quark mass. The average $p_T$ of a b quark from the top-quark decay is even larger at the LHC since the top quarks themselves are boosted.

In this analysis, the Track Counting High Purity (TCHP) algorithm [173] is used for jet-flavor tagging. This algorithm belongs to the class that exploits the long lifetime of B hadrons. It uses the three-dimensional impact-parameter significance $d_{xyz}/\sigma(d_{xyz})$ of the third track of a jet as a discriminator. Here, $d_{xyz}$ is calculated w.r.t. the primary vertex in each event, and the $d_{xyz}/\sigma(d_{xyz})$ values of all tracks in a jet are sorted in decreasing order. The determination of the three-dimensional impact parameter profits from the high resolution of the CMS pixel detector even in $z$ direction. The TCHP is preferred due to two reasons. First, it is expected a priori that the TCHP algorithm profits from a low dependence on the number of additional interactions since additional tracks are rejected if they are too far away from the jet axis. This is confirmed in measurements in data [173]. Second, the Track Counting High Efficiency (TCHE) algorithm, which is a variant of the TCHP algorithm, is used already at trigger level for events with an electron in the final state. The TCHE utilizes the impact parameter of the second track instead of the third track. The TCHP algorithm is used also offline as it is highly correlated to the trigger decision.

Figure 3.12 shows the discriminator values for the TCHP algorithm for the inclusive jet-$p_T$ and jet-$\eta$ range of QCD-multijet events. A reasonable agreement between simulated and observed distributions is obtained. Scale factors are derived in order take the remaining differences between data and simulation into account. These scale factors are discussed later in this section.

Even if the jet-flavor-tagging algorithms have continuous discriminator values, their efficiency in data is only estimated for dedicated working points. Working points are defined according to the probability of tagging jets from u, d, or s parton fragmentation, also referred
3.3 Reconstruction of Physics Objects

Figure 3.12: Comparison of jet-flavor-tagging-discriminator values between observed and simulated QCD-multijet events for the TCHP algorithm. The tight working point corresponds to a TCHP discriminator > 3.41. The figure is taken from ref. [174].

to as “mistag” probability. The “loose” working point corresponds to a mistag probability (in simulated events) close to 10%, the “medium” working point to 1%, and the “tight” working point to 0.1%. In this analysis, the tight working point is used. The tight working point for the TCHP algorithm corresponds to a discriminator value larger than 3.41.

The performance of various b-tagging algorithms is measured in QCD-multijet events in data [173]. Figure 3.13 shows the tagging efficiency of b-flavored jets vs. the mistag probability. At the tight working point, a b-tagging efficiency of about 45% is reached. The b-tagging efficiency is degraded by a few percent in events with many additional interactions.

Figure 3.13: Performance of the TCHP algorithm for two Pile-Up scenarios. The loose, medium, and tight working points correspond to mistag probabilities of 10%, 1%, and 0.1%. In this analysis, the tight working point is used. The figure is taken from ref. [173].
Furthermore, simulated efficiencies are compared to efficiency measurements in data. Jet-flavor-tagging-efficiency measurements for jets originating from b-parton fragmentation are carried out with four different methods in QCD-multijet events. In the low-jet-$p_T$ regime, kinematic properties of muons from subsequent decays of B hadrons are used to discriminate among different jet flavors. Events with muons that are nearby to the jet axis are selected. The muon transverse momentum relative to the jet axis ($p^\text{rel}_T$) has a characteristic distribution due to the relatively large mass of the b quark. Another measurement exploits the muon $p^\text{rel}_T$ to extract tagging efficiencies with a tag-and-probe of jets. In the high-jet-$p_T$ regime, $p^\text{rel}_T$ provides less discrimination power among different jet flavors. Instead, one method fits the three-dimensional impact parameter of the muon track in events with muons nearby the jet axis. The impact parameter has substantial discrimination power due to the long lifetime of B hadrons. Another method uses a lifetime tagger which can be directly calibrated with data. The lifetime tagger provides a reference to calibrate other taggers. Misidentification probabilities, i.e. tagging efficiencies for jets from u, d, or s parton fragmentation, are derived by inverting tagging criteria such that non-b-jets are selected.

Scale factors are derived in order to match the jet-flavor-tagging efficiencies observed in data events. They are expressed as a function of jet-$\eta$ and jet-$p_T$ and applied to simulated events in order to account for differences between simulated and observed jet-flavor-tagging efficiencies. Figure 3.14 refers to the $p_T$-dependent scale factors and their uncertainties, in which the values are taken from ref. [173]. For the TCHP algorithm at the tight working point, the scale factors are constant in pseudo-rapidity $\eta$. The observed tagging efficiency is in agreement with the simulated efficiencies at a level of $O(10\%)$ for jets from b- or c-parton fragmentation, while the agreement is of $O(30\%)$ for jets from fragmentation of u, d, or s partons.

Figure 3.14: Scale factors and uncertainties to the TCHP-tagging efficiencies as applied to the simulated events. The scale factors depend on the jet flavor and reconstructed jet-$p_T$, while they remain constant in $\eta$ for the TCHP algorithm at the tight working point. Values are from ref. [173].

Uncertainties on the obtained scale factors are due to additional interactions, the fraction of gluon splitting into $b\bar{b}$ pairs, uncertainties related to the muon selection, and individual uncertainties to the particular measurement procedure. Additional uncertainties for the determination of the mistag probability account for the amount of long-lived $K^0_S$ and $\Lambda$ particles, as well as mismeasured tracks among others. Some methods exploit only events with muons from semi-leptonic hadron decays, which is a small fraction of the total inclusive phase space of jets from b-parton fragmentation. This extrapolation required additional studies with simulation that confirmed the extrapolation to be reasonable. The scale factors for jets from c-quark
3.3 Reconstruction of Physics Objects

Efficiency measurements are also carried out in a phase space which is enriched in \( t\bar{t} \) events [173]. These measurements are a cross-check with varied jet multiplicities and provide an inclusive (average) data-simulation-scale factor which is adapted to the typical \( t\bar{t} \) jet-\( p_T \) and jet-\( \eta \) spectra. Both semi-leptonic-decay channels and dilepton-decay channels are exploited to obtain two statistically independent samples. The used methods exploit the two-dimensional distribution of jet-multiplicity vs. b-tagged-jet multiplicity, the one-dimensional b-tagged-jet-multiplicity distribution, or the invariant-jet-muon-mass distribution in a sample enriched with jets from a leptonically decaying top-quark.

In order to compare scale factors of both measurements from QCD-multijet and \( t\bar{t} \) events, \( p_T \)- and \( \eta \)-dependent scale factors from QCD-multijet events are reweighted such that they match the \( t\bar{t} \) jet-\( p_T \) and jet-\( \eta \) spectra. The efficiency measurements obtained from \( t\bar{t} \) events result in scale factors that are consistent with the results obtained from QCD-multijet events, but have slightly smaller uncertainties. In measurements with QCD-multijet events, \( SF_b = 0.91 \pm 0.04 \) is obtained, and \( SF_b = 0.93 \pm 0.03 \) is obtained in measurements with \( t\bar{t} \) events. Here, \( SF_b \) refers to the average data-simulation-scale factor in the \( t\bar{t} \) phase space for jets from b-parton fragmentation that are tagged with the TCHP algorithm at the tight working point. The obtained scale factors are found to be stable over the whole data-taking period.

In this analysis, the jet-\( p_T \) dependent scale factors that are obtained from efficiency measurements with QCD-multijet events are used [173]. These measurements are inferred from events that are, in particular, uncorrelated to events in this analysis, whereas a subset of data events is the same between the \( t\bar{t} \) efficiency measurement and the analysis at hand. However, similar techniques as in the efficiency measurement with \( t\bar{t} \) events are exploited in this analysis. The b-tagging-efficiency uncertainty is constrained in-situ during the statistical inference, which is part of the analysis strategy (cf. sections 4.1, 4.6, and 6.1).

### 3.3.7 Missing Transverse Energy

The CMS detector is directly sensitive to particles that interact either via the electromagnetic or strong force with the detector material. Neutrinos are electrically neutral and weakly interacting particles, such that they escape the CMS detector without any response. The typical signature of such weakly interacting particles is instead inferred indirectly by a momentum-imbalance of all visible particles. This is possible since momentum is conserved in each direction. However, particles along the beamline escape undetected due to the limited detector acceptance, and the initial momentum of the colliding partons is unknown as they carry only fractions of the proton momenta. Thus, only the imbalance of the transverse momentum \( p_T^{\text{miss}} \) can be measured. \( p_T^{\text{miss}} \) is calculated as the negative vectorial sum of the transverse momenta of all visible particles that are reconstructed with the PF algorithm. The missing transverse energy \( E_T^{\text{miss}} \) is defined as the magnitude of the transverse-momentum-imbalance vector.

In this analysis, the signal cross section is measured in events with leptonically decaying W bosons, which stem from a top-quark decay (\( t \rightarrow bW \rightarrow bl\nu_l \), with \( l = e, \mu \)). Neutrinos from W-boson decays typically have a large transverse momenta due to the large mass of the mother particle. Therefore, \( E_T^{\text{miss}} \) is an important part of the event signature for \( t \)-channel events. Background processes like QCD-multijet production include neutrinos from subsequent hadron decays. These neutrinos are typically much softer than neutrinos from the hard-scattering
3 Experimental Setup

process, but can be important due to the large QCD-multijet-production cross section. Furthermore, $E_T^{\text{miss}}$ can be artificially induced by particle-momentum mismeasurements, particle misidentification, or detector malfunctions etc. (cf. [175]).

The precise determination of $E_T^{\text{miss}}$ plays also an important role in searches for physics processes beyond the SM. As an example, highly energetic neutrinos might be produced from heavy charged gauge-boson (W') production with Sequential SM couplings, and are searched for in events with lepton-plus-$E_T^{\text{miss}}$ event signatures [176, 177].jet-plus-$E_T^{\text{miss}}$ signatures can be interpreted in terms of large extra dimensions or canonical searches for Weakly Interacting Massive Particles (WIMPs) [178, 179]. WIMPs might serve as dark-matter candidates. $E_T^{\text{miss}}$ is also an important ingredient in searches for unparticles [180]. Moreover, a variety of hypothetical particles exist that itself interact only weakly with the detector material, e.g. due to supersymmetric extensions to the SM (cf. [11]).

$E_T^{\text{miss}}$ performance is studied in $pp$ collisions at $\sqrt{s} = 7$ TeV with minimum-bias and QCD dijet events. Data corresponding to an integrated luminosity of 11.7 nb$^{-1}$ is used to study effects from anomalous calorimeter signals or beam background [181]. A procedure is established to clean data events from remaining instrumental anomalies and beam-background contributions (cf. section 4.3).

The performance of the $E_T^{\text{miss}}$ reconstruction is also studied in events with low Pile-Up conditions and an integrated luminosity of 36 pb$^{-1}$ of data [175]. Events containing leptonically decaying Z bosons ($Z \rightarrow e^+e^-, Z \rightarrow \mu^+\mu^-$) and events with $\gamma$ bosons are used to determine both $E_T^{\text{miss}}$ response (i.e. scale) and resolution. In these events, a momentum imbalance is induced by removing the vector boson (resp. its decay products) before reconstructing $p_T^{\text{miss}}$. Transverse-momentum conservation is used to determine the $E_T^{\text{miss}}$ response and resolution. Here, the accurately measured transverse momentum of the boson is balanced with the transverse momentum of all other particles in the event, which often referred to as the “hadronic recoil”. Furthermore, the $E_T^{\text{miss}}$ performance is evaluated in events with leptonically decaying W bosons ($W \rightarrow e\nu, W \rightarrow \mu\nu$). These events contain intrinsic and large $E_T^{\text{miss}}$ due to the presence of a prompt neutrino. The $E_T^{\text{miss}}$ response is found to be well reproduced between data and simulation, although the $E_T^{\text{miss}}$ resolution is found to be degraded by 10% in data compared to simulation.

Furthermore, the $E_T^{\text{miss}}$ response and resolution are studied in a larger dataset which corresponds to an integrated luminosity of 4.6 fb$^{-1}$ with $Z \rightarrow \mu^+\mu^-$ events [182]. Events in this dataset have on average much more Pile-Up interactions than in the studies from reference [175]. Hence, the $E_T^{\text{miss}}$ response and resolution can be studied as a function of the number of additional interactions. The $E_T^{\text{miss}}$ response is expected to be unaffected by Pile-Up. This is confirmed in measurements with data [182]. An overall reasonable agreement of the $E_T^{\text{miss}}$ response between data and simulated events is found for the total dataset. The $E_T^{\text{miss}}$ response has variations up to $\pm 20\%$ between data and simulated events. The reason is that the $E_T^{\text{miss}}$ resolution in data is not perfectly described by the simulation.

When using the direction of the Z boson in the transverse plane as an axis, the $E_T^{\text{miss}}$ resolution (of the hadronic recoil) can be studied in both parallel and transverse directions to this axis. The $E_T^{\text{miss}}$ resolution can be further expressed as a function of the transverse momentum $q_T$ of the Z boson (fig. 3.15 run 2011A). The parallel component of the $E_T^{\text{miss}}$ resolution is $\approx 13$ GeV/c for $q_T < 50$ GeV/c, and it is increasing linearly to $\approx 25$ GeV/c for $q_T = 225$ GeV/c. The transverse component of the $E_T^{\text{miss}}$ resolution is on average $\approx 13$ GeV/c and mostly flat within $q_T$. Furthermore, the $E_T^{\text{miss}}$ resolution is expected to be degraded as a function of the number of additional interactions $N_{\text{Vertices}}$. The parallel component of the $E_T^{\text{miss}}$ resolution is $\approx 9$ GeV/c.
3.4 Collision Data and Luminosity Determination

for $N_{\text{Vertices}} = 1$ and is increasing linearly to $\approx 19\text{GeV}/c$ for $N_{\text{Vertices}} = 18$. The transverse component has the same behavior.

The transverse component has the same behavior.

Figure 3.15: $E_{\text{T}}^{\text{miss}}$ resolution as a function of the transverse momentum of the Z boson ($q_T$), when induced $E_{\text{T}}^{\text{miss}}$ in $Z \rightarrow \mu^+ \mu^-$ events by removing the vector boson \[182\]. The $E_{\text{T}}^{\text{miss}}$ resolution is shown in both parallel and transverse directions to the axis of the transverse momentum of the Z boson. In this analysis, data of run 2011A are used.

The observation is that the $E_{\text{T}}^{\text{miss}}$-resolution components are well simulated for events with more than five reconstructed vertices in both parallel and transverse directions. However, the resolution is worse in data for events with less than five vertices. The resolution of both components is found to be 7% worse for events with exactly one reconstructed primary vertex, and 4% worse for events with exactly three reconstructed primary vertices. The observed resolution in data as a function of $q_T$ agrees well with the simulated resolution by 5% for $q_T < 50\text{GeV}$ for both components. These findings are consistent with the result from studies with events at low Pile-Up conditions (cf. \[175\]).

In conclusion, $E_{\text{T}}^{\text{miss}}$ is found to be reasonably well understood and ready to be used in CMS high-precision analyses including top-quark analyses \[175\]. The varied $E_{\text{T}}^{\text{miss}}$ resolution between data and simulation, which also propagates to the $E_{\text{T}}^{\text{miss}}$ response, is covered with corresponding uncertainties and discussed in section 6.1.6.

3.4 Collision Data and Luminosity Determination

Recorded collision data are categorized into several datasets already at the [HLT] level. Each trigger belongs to a certain dataset. In this analysis, datasets are used that correspond to single-muon triggers, single-electron triggers, and triggers that require an electron and hadronic activity. Technically speaking, the SingleMu, SingleElectron, and ElectronHad datasets of the proton-proton run Run2011A are analyzed.

The CMS-data certification ensures that the subdetectors, subsystems, and physics objects reconstruction were operational and that they performed as expected. Thus, it ensures a good
Experimental Setup

quality of the data used for physics analyses. Data are certified in units of luminosity sections of a certain run; a “luminosity section” refers to a 23.3-s-long data-taking period in which the instantaneous luminosity remains nearly constant [183]. Only data that are officially certified by CMS are used in this analysis.

Many techniques for luminosity determination exist, including techniques that provide measurements either online for real-time performance monitoring or offline for an absolute normalization as used in high-level-physics analyses. In the following, a brief overview of the usual methods to determine the luminosity, which is recorded by the CMS detector, is given, and the luminosity determination as used in this analysis is briefly discussed.

One measure of luminosity is based on the HF calorimeter [153, 184, 185], since every interaction deposits energy in the HF calorimeter well above the noise level. The number of interactions itself is correlated to the luminosity. Two methods exist that make use of the HF calorimeter. The first method is based on counting the (average) number of “empty” HF towers, i.e. towers in which no energy is deposited. The number of occurring interactions in each bunch crossing follows Poisson statistics, and HF towers can be used to check whether an interaction occurred or not. However, the HF is not directly sensitive to the number of interactions that occurred, but can distinguish between zero or more interactions. Instead of counting the number of HF towers, the number of “empty” towers is used to determine the number of interactions with inverted Poisson probability. Additional sensitivity comes from the fact that all HF towers are statistically independent among each other. The second method makes use of the linear correlation between the deposited transverse energy $E_T$ in the HF towers and the number of interactions. Van der Meer (VdM) scans are used for an absolute calibration of the luminosity with minimal model dependence.

In a VdM scan [186], which was originally conceived by S. Van der Meer for the “Intersecting Storage Rings” at CERN in 1968, the absolute luminosity is deduced from machine parameters only. The two beams are displaced against each other in the transverse plane, while monitoring the relative beam-beam-interaction rate. Then, the instantaneous luminosity can be expressed as a function of the beam separation, the beam profile (i.e. the size and shape of the proton density in the beam), the bunch intensity, and the LHC-orbit frequency $f$. The bunch intensity is determined externally with measurements using Fast Beam Current Transformers [184]. The LHC-orbit frequency (also referred to as “revolution frequency”) is a general parameter of the LHC storage ring, which is $f = 11.245$ kHz [130]. Measuring the collision rate and fitting the observed beam-separation dependence allows the peak instantaneous luminosity to be determined, as well as the effective beam size of the two beams. A double-Gaussian-beam profile describes the observed VdM-scan curves well [184]. The parametrization of the beam sizes in both $x$ and $y$ directions needs subsequent (separate) scans. However, the beam width evolves with time, and the emittance needs to be corrected accordingly [184].

A luminosity measurement with the HF calorimeter has the advantage that it allows for an online measurement. However, it has a non-linear response w.r.t. the instantaneous luminosity and a large sensitivity to the beam intensity. The total systematic uncertainty of the measured integrated luminosity is 4.5% [184] when using the HF-based luminometer.

Alternative determinations of luminosity include normalizations to measured cross sections of “standard-candle processes” like W-boson or Z-boson production [153]. Furthermore, the primary-vertex-production rate can be counted to derive a luminosity measurement, either by simply counting vertices or by deriving it (with Poisson statistics) from events in which no vertex is found [184]. Moreover, the TOTEM experiment uses the optical theorem to provide a complementary luminosity measurement [153].
In this analysis, the integrated luminosity in each luminosity section is calculated offline using a method that is based on pixel-cluster counting. The pixel-cluster cross section itself is a function of the integrated luminosity, mean number of observed pixel clusters, and the revolution (orbit) frequency $f$ (see above). The absolute calibration of the pixel-cluster cross section is done with VdM scans. Here, the data are recorded using zero-bias triggers, i.e. using totally inclusive triggers. Since the number of pixels in the tracker is large, the probability that two tracks of the same bunch-crossing hit the same pixel is small. Therefore, the actual number of counted pixel clusters in a collision, divided by the pixel-cluster cross section, can be used as a measure of luminosity.

The estimated total uncertainty of the luminosity measurement is 2.2% [183]. The dominant uncertainties are the scan-to-scan variations (1.5%) and afterglow of the beam (1.0%) [183]. “Scan-to-scan” variations include two effects of the VdM scans. First, the measured pixel-cluster cross section spreads by 1.2% in a pair of two subsequent scans. Second, potential shape differences of the beam profile are covered, since the measured cross section depends on the shape of the interaction region. “Afterglow” effects are caused by late-arriving particles and activated material [183]. They induce a “ghost” response, even in unfilled runs, which has to be corrected for. The afterglow effect is negligible for a small number of bunches, but becomes significant if many bunches are combined in bunch trains [184].

The recorded integrated luminosity of the CMS experiment at $\sqrt{s} = 7$ TeV is shown as a function of time in fig. 3.16. Collisions corresponding to 6.13 fb$^{-1}$ were delivered to the CMS experiment in 2011, and the detector was fully operational to record data corresponding to 5.55 fb$^{-1}$.

Figure 3.16: Total integrated luminosity of proton-proton collisions at $\sqrt{s} = 7$ TeV delivered to and recorded by the CMS experiment in 2011 [183].
3.5 Simulation of Signal and Background Processes

In this section, the event generation of signal and background processes and their production cross sections are described. A summary of all event generators for the various processes and the process normalizations is given at the end of this section.

Event modeling is discussed in detail from a theoretical point of view in section 2.2.1. In this section, the software packages which are used to generate events are described. All event generators are interfaced to PYTHIA 6 [63], and the PYTHIA 6 Tune Z2 [70] is used for underlying-event modeling. The top-quark mass is set to $m_t = 172.5 \text{ GeV}/c^2$, the W-boson mass is set to $m_W = 80.419 \text{ GeV}/c^2$, and the Z-boson mass is set to $m_Z = 91.188 \text{ GeV}/c^2$ for event generation. $\tau$-lepton decays are simulated with TAUOLA [65].

Simulated events are normalized to the integrated luminosity of the analyzed data according to their predicted production cross sections. Here, theoretical calculations are used for the normalization of inclusive production cross sections. These predictions are usually available in higher orders of perturbative QCD than event generators can provide. Table 3.4 summarizes the predicted production cross section values for all processes. Uncertainties on the normalization of background processes, in the dedicated phase space in which this analysis is performed, are an essential part of the statistical inference and discussed in section 6.1.11.

Residual corrections to simulated events are described in section 3.5.4. These corrections account for variations of trigger efficiencies, Pile-Up conditions, jet-flavor-tagging efficiencies, and jet-transverse-momentum resolutions when comparing simulated events with measured data events. Jet-energy calibration is discussed in a previous section 3.3.5.

3.5.1 Signal-Event Generation and Normalization

The LO-Feynman diagrams for $t$-channel single-top-quark production in both 4-flavor scheme (2→3, $qq\rightarrow q't\bar{b}$, left) and 5-flavor scheme (2→2, $qb\rightarrow q't$, right) are shown in fig. 3.17. The 4-flavor scheme does not contain b partons in the PDF and the b quark has to be massive within the matrix-element calculation. In the 5-flavor scheme, logarithms ($\log(\mu^2/m_b^2)$) that arise from (collinear) initial-state-gluon splitting are resummed into the b-parton PDF [115, 187]. Here, the b-quark is assumed to be massless in the matrix-element calculation.

The $2\rightarrow 2$ ($qb\rightarrow q't$) process (fig 3.17, right) implies that an additional final-state-b parton is generated within the parton shower, since the initial-state-b parton can be produced only via off-shell-gluon splitting ($g^*\rightarrow b\bar{b}$) in the proton. Initial-State Radiation (ISR) modeling, e.g. with PYTHIA, provides a way to generate the second $b$ parton. However, the PYTHIA generator...
3.5 Simulation of Signal and Background Processes

accurately models the additional-b-quark kinematics only for low transverse momenta. The \(2 \rightarrow 3\) \((gg \rightarrow q't\bar{b})\) process (fig. 3.17 left) instead is suited most for modeling the hard-\(p_T\) (2\(^{\text{nd}}\) b) region. The additional final-state-b quark, which does not stem from the top-quark decay, is also referred to as “2\(^{\text{nd}}\)-b quark” or “spectator-b quark”.

The \(2 \rightarrow 3\) process can also be interpreted in terms of a NLO correction to the \(2 \rightarrow 2\) LO Feynman diagram. Event generation according to both diagrams gives an “effective” NLO description of the \(t\)-channel events. However, part of the phase space is double-counted when simply adding events from both \(2 \rightarrow 2\) and \(2 \rightarrow 3\) contributions. Moreover, different kinematics for the spectator-b quark, as well as different production cross sections, are obtained for events from both diagrams. Therefore, event generators implement special procedures to match events from \(2 \rightarrow 2\) and \(2 \rightarrow 3\) diagrams.

Event generators that simulate \(t\)-channel events in NLO+Parton-Shower (PS) accuracy use either the 4-flavor scheme (with massive initial-state-b quarks) or the 5-flavor scheme (with a massless-initial-state-b-quark approximation) for event simulation. A full NLO description for the second-b quark is obtained only in the 4-flavor scheme, which is, however, not available yet (cf. the discussion in the following paragraphs).

In this analysis, the default signal modeling is done with the POWHEG BOX event generator \[60–62, 188\] in the 5-flavor scheme with NLO+PS accuracy. A generator based on COMPPHEP \[189, 190\] (SINGLETOP) is further used to study the influence of the choice of the generator on the \(t\)-channel-signal modeling. This generator implements a matching of the \(2 \rightarrow 2\) and \(2 \rightarrow 3\) diagrams according to the transverse-momentum distribution of the spectator-b quark, and uses massive b quarks. This generator is simply referred to as “COMPPHEP” for unambiguousness in the following. COMPPHEP was also used as the central signal generator for the observation of single-top-quark production by the D0 collaboration in 2009 \[6\]. COMPPHEP generates events in an “effective NLO approximation” \[190\]. The CMS experiment provides events for these two generators only.

\(t\)-channel events are generated using the CTEQ6M NLO PDF set \[52\]. Moreover, both POWHEG BOX and COMPPHEP preserve correlations of the top-quark spin between its production and decay.

Signal modeling with these generators is explained briefly in the following paragraphs. Furthermore, an overview of event generators that produce \(t\)-channel events in LO accuracy, 4-flavor scheme in NLO+PS accuracy, 5-flavor scheme in NLO+PS accuracy, and effective NLO accuracy with matched \(2 \rightarrow 2\) and \(2 \rightarrow 3\) processes is given.

**Leading-Order-Event Generation** Event generators that are able to model \(t\)-channel events in the 4-flavor scheme \(2 \rightarrow 3\) or 5-flavor scheme \(2 \rightarrow 2\) in LO+PS accuracy include MADGRAPH \[59\], WHIZARD \[191, 192\], and PYTHIA \[63\]. MADGRAPH and WHIZARD take spin correlations between the top-quark production and decay into account. Furthermore, the WHIZARD event generator is able to generate events with anomalous \(Wtb\) couplings.

**4-flavor scheme \(2 \rightarrow 3\) at NLO** The event generation of the \(t\)-channel 4-flavor scheme \(2 \rightarrow 3\) with massive b quarks in the initial state is available at NLO with the POWHEG BOX \[187\] or aMC@NLO \[187\] generators. Event generation in the 4-flavor scheme with massive b quarks yields a more precise description (NLO) of the spectator-b quark \[187\], and is, in principal, the

\[\text{aMC@NLO is implemented within the MADGRAPH 5 framework and provides an automated matching between events generated in NLO QCD accuracy and parton-shower simulations.}\]
preferred choice to generate events. However, in the current implementations, the spin correla-
tions between top-quark production and decay are not yet preserved. The total production
cross section and differential distributions can also be obtained with MCFM [76, 115].

5-flavor scheme (2→2) at NLO POWHEG BOX [60, 62, 188] and MC@NLO [67, 193] generate
events in the 5-flavor scheme with the massless initial-state-b-quark approximation in NLO+PS
accuracy. Hard and wide-angle emissions are calculated in NLO Soft and collinear emissions
are subject to PS modeling. POWHEG BOX and MC@NLO use different schemes to avoid a double-
or under counting of phase space between matrix-element calculation and parton-
shower simulation.

The MC@NLO generator subtracts double-counted phase space by generating negatively-
weighted events. MC@NLO can be interfaced to HERWIG [64] for PS modeling. However,
the current implementation faces a technical feature that leads to an unphysical description
of the additional b-quark (cf. [187]). POWHEG BOX is interfaced to PYTHIA, which uses a $p_T$-
ordered shower to resum all remaining soft and collinear corrections. Renormalization-
and factorization scale are set to the transverse momentum (relative to the beam axis) of the hardest
emitted parton. Then, the first emission is always calculated by POWHEG BOX, and subsequent
radiation is performed by PYTHIA. Only positively-weighted events are obtained in this way.
In the matrix-element calculation, u, d, s, c, b quarks are assumed to be massless. However, the
quark masses are considered as lower thresholds ($p_T^{min}$) for parton radiation of a certain flavor
(cf. [188] for more details). The default t-channel signal in this analysis is modeled with the
POWHEG BOX generator.

Matching of the 2→2 and 2→3 contributions at “effective NLO” An alternative procedure
involves the matching of 2→2 and 2→3 contributions in order to avoid a double counting
of phase space. Matching procedures enable an “effective NLO” description with massive
b quarks. Automated matching procedures are implemented in COMPHEP (SINGLETOP)
[189, 190], ACERMC [194, 195], and PROTOS [13]. Furthermore, PROTOS and COMPHEP
are able to generate events with anomalous Wtb couplings.

The COMPHEP (SINGLETOP) event generator matches events from both diagrams such that a
smooth transverse-momentum distribution of the spectator-b quark is obtained. The matching
procedure is described in detail in ref. [190], and briefly summarized in the following
paragraphs.

For the 2→2 ($qb\rightarrow q't$) process (fig.3.17, right), the final-state-b parton is generated within
the PYTHIA parton shower, which provides a good approximation at low transverse momenta.
This process is referred to as $pp \rightarrow tq + b_{ISR, PYTHIA}$ in the following. For the 2→3 ($qg\rightarrow q'tb$)
process (fig.3.17, left), the final-state-b parton is modeled within the matrix-element calculation.
The matrix-element calculation provides a good estimate for large transverse momenta of the
additional b parton. This process is referred to as $pp \rightarrow tq + b_{LO, COMPHEP}$ in the following.

However, the kinematics of the spectator-b quark, i.e. the $p_T$ and $\eta$ distributions, differ sig-
nificantly between both processes $pp \rightarrow tq + b_{ISR, PYTHIA}$ and $pp \rightarrow tq + b_{LO, COMPHEP}$. A matching
of both processes is done w.r.t. the kinematics of the spectator-b quark in the final-state,

\[
\sigma_{NLO} = K \times \sigma_{pp\rightarrow tq+b_{ISR, PYTHIA}} \bigg|_{p_T(2^{nd}\ b)<Q} \\
+ \sigma_{pp\rightarrow tq+b_{LO, COMPHEP}} \bigg|_{p_T(2^{nd}\ b)>Q} .
\] (3.16)
Here, the matching threshold $Q$ is optimized such that a smooth distribution of the transverse momentum of the spectator-$b$ quark ($p_T(2\text{nd} \ b)$) is obtained. Then, also a smooth $\eta$-distribution is achieved. The overall normalization is kept constant at the total NNLO cross section, while events with soft second-$b$ quarks are enhanced by a factor $K$. The $K$-factor effectively resums higher-order-loop corrections [190]. A threshold of $Q = 28 \text{GeV}/c$ is used for the simulated event samples at $\sqrt{s} = 7 \text{TeV}$. In particular, this threshold is close to $p_T$ threshold for reconstructed jets as used in this analysis (30 $\text{GeV}/c$).

Another matching technique [195] is implemented in the AcErMC event generator [194]. Here, the full phase space is described by $2\rightarrow2$ and $2\rightarrow3$ diagrams, and the double-counted phase space is subtracted (eq. 3.17). Events of the $2\rightarrow3$-subtraction term (order $\alpha_s^{(1)}$) obtain negative event weights.

$$\sigma = 2 \rightarrow 2 + 2 \rightarrow 3 \oplus (2 \rightarrow 3)_{\text{(subtraction term)}} \quad (3.17)$$

The PROTOS event generator [13] implements two matching procedures. The first matching algorithm exploits a matching based on the $p_T$ of the $2\text{nd} \ b$ quark, similar to the one implemented in the COMPhEP (SINGLETOP) event generator. The second matching procedure performs a subtraction to the $b$-quark PDF, which only has an effect on the $2\rightarrow2$ contribution.

**Normalization** The inclusive cross sections for the single-top-quark processes [8–10] are calculated with a top-quark mass of $m_t = 172.5 \text{GeV}/c^2$, which corresponds to the value used in the event simulation. Calculations are available in approximate NNLO accuracy. Renormalization scale $\mu_r$ and factorization scale $\mu_f$ are set to a common scale $\mu \equiv m_t$, i.e. $\mu = \mu_f = \mu_r = m_t$. The MSTW2008 NNLO PDF set [75] is used. Theoretical uncertainties of the calculated cross section arise from scale variations and the parametrization of the PDF set. The scale $\mu$ is varied between $m_t/2$ and $2m_t$, and $\mu_f$ and $\mu_r$ are taken as fully correlated. The PDF uncertainty covers (eigenvector) variations within the MSTW2008 NNLO PDF set [75] at 90% CL. The $t$-channel-production cross section is predicted to be $(41.92^{+1.99}_{-2.01} \pm 0.83) \text{ pb}$ for events with top quarks and $(22.65^{+0.50}_{-0.50} + 0.68)$ pb for events with top-anti quarks.

### 3.5.2 Background-Event Generation and Normalization

Background processes include $s$-channel single-top-quark, $tW$-channel single-top-quark, $t\bar{t}$, $W$-boson-plus-jets, $Z$-boson-plus-jets, Diboson ($WW$, $WZ$, $ZZ$), and QCD-multijet production.

**$s$-channel and $tW$-channel single-top-quark production** The POWHEG BOX generator is used to simulate the $s$-channel [60–62, 188] and $tW$-channel [60–62, 89] single-top-quark processes. NNLO corrections to the $tW$-channel result in an interference [88, 89] between $tW$-channel and $t\bar{t}$ production. In this analysis, the diagram-removal technique is used to define the NNLO corrections. However, differences to the diagram-subtraction technique are negligible for this analysis, and the $tW$-channel is only a minor background for $t$-channel events.

Events are normalized to approximate NNLO calculations as described for $t$-channel events in the previous paragraph. The $s$-channel-production cross section is estimated to be $(3.19^{+0.06}_{-0.06} + 0.13)$ pb for events with top quarks and $(1.44^{+0.01}_{-0.01} + 0.06)$ pb for events with top-anti quarks. The $tW$-channel-production cross section is calculated to be $(7.87^{+0.20}_{-0.20} + 0.55)$ pb for both top-quark and top-anti-quark production.
Top-quark-pair production is modeled with the tree-level matrix-element generator MadGraph [59], which is interfaced to Pythia 6 [63]. A dynamical, combined scale $\mu_f^2 = \mu_r^2 = \mu_f^2 = m_t^2 + \sum p_T^2$ is used, in which the sum runs over all additional jets. The CTEQ6L1 PDF set [52] is used to generate $t\bar{t}$ events.

Diagrams with up to three additional partons are generated at the matrix-element level. Double- and under counting between jets generated by the matrix element and parton shower are avoided by using the MLM-matching prescription [66] with $k_T$-jets. Matching and $Q^2$-scale parameters are given in table 3.1. Matching threshold and matching scale are absolute values, and parameters for the $Q^2$ scale are multiplicative factors. The “nominal” sample is generated with the default values. The matching threshold $xqcut$, which refers to the minimum $k_T$-jet distance between partons at the matrix-element level, is chosen as recommended by the authors of MadGraph [59]. The matching scale $qcut$ is optimized such that a smooth differential-jet-multiplicity distribution is obtained. This optimization is separately performed for each physics process.

Individual samples are generated for variations of matching- and $Q^2$ scale in order to study systematic effects due to the choice of the particular value. Two samples are generated with up- or down variations of the matching scale, while keeping the $Q^2$ scale at its default value. Additional two samples with varied $Q^2$ scale are generated, while keeping the matching scale constant.

<table>
<thead>
<tr>
<th>Parameter Name</th>
<th>Default Value $^{\text{Up-variation}}$</th>
<th>Default Value $^{\text{Down-variation}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matching</td>
<td></td>
<td></td>
</tr>
<tr>
<td>threshold ($xqcut$)</td>
<td>$20^{+2}_{-0.5}$</td>
<td></td>
</tr>
<tr>
<td>scale ($qcut$)</td>
<td>$40^{+30}_{-20}$</td>
<td></td>
</tr>
</tbody>
</table>

$Q^2$-scale variations

<table>
<thead>
<tr>
<th>MadGraph</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_f^2$</td>
<td>$1.0^{+4}_{-0.25}$</td>
<td></td>
</tr>
<tr>
<td>$\mu_r^2$</td>
<td>$1.0^{+4}_{-0.25}$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pythia</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PARP(64)</td>
<td>$1.0^{+4}_{-0.25}$</td>
<td></td>
</tr>
<tr>
<td>PARP(72)</td>
<td>$0.25^{+0.5}_{-2}$</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Parameters of matching scale, renormalization-, and factorization scale as used to generate $t\bar{t}$ events.

Simulated top-quark-pair events are normalized to an inclusive cross section of $(157.5^{+18.0}_{-14.7} - 19.5^{+14.7}_{-19.5}) \text{ pb}$ as calculated with MCFM (v5.8) in NLO [76, 77]. A top-quark mass of $m_t = 172.5 \text{ GeV}/c^2$, the CTEQ6M NLO PDF set [52], and a combined scale $\mu = \mu_r = \mu_f = m_t$ are used for this prediction. The first uncertainty of the prediction is due to scale variations. The scales $\mu_f$ and $\mu_r$ are varied simultaneously between $\mu/2$ and $2\mu$. The second uncertainty is calculated by varying the PDF set within all possible parametrizations (as given by the 40 eigenvector sets) at 68% CL.
3.5 Simulation of Signal and Background Processes

W-boson-plus-jets production  W-boson-plus-jets production is modeled with the tree-level matrix-element-generator MADGRAPH [59], which is interfaced to PYTHIA 6 [63]. The CTEQ6L1 PDF set [52] is used to generate the events. A dynamical, combined scale \(\mu^2 = \mu_f^2 = \mu_r^2 = m_W^2 + \sum \vec{p}_T^2\) is used, in which the sum runs over all additional jets. Diagrams with up to four additional partons are generated at the matrix-element level. The MLM-matching prescription [66] with \(k_T\)-jets is used similar to the description for \(t\bar{t}\) in the previous paragraph. Matching and \(Q^2\)-scale parameters are summarized in table 3.2.

<table>
<thead>
<tr>
<th>Parameter Name</th>
<th>Default Value</th>
<th>Up-variation</th>
<th>Down-variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matching</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>threshold ((xqcut))</td>
<td>10 (\times) 2 (\times) 0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>scale ((qcut))</td>
<td>20 (\pm) 10</td>
<td></td>
<td>-10</td>
</tr>
<tr>
<td>(Q^2)-scale variations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MADGRAPH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\mu_f^2)</td>
<td>1.0 (\times) 4 (\times) 0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\mu_r^2)</td>
<td>1.0 (\times) 4 (\times) 0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PYTHIA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PARP(64)</td>
<td>1.0 (\times) 4 (\times) 0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PARP(72)</td>
<td>0.25 (\times) 0.5 (\times) 2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Parameters of matching scale, renormalization-, and factorization scale as used to generate W-boson-plus-jets events.

The jet-multiplicity spectrum of W-boson-plus-jets production is steeply falling towards higher jet multiplicities. In order to enlarge the statistical precision of the simulated event samples, both an inclusive W-boson-plus-jets sample and samples with exclusive parton multiplicities are generated. Event samples with exclusive parton multiplicities are generated with up to four additional partons. Events from the inclusive and exclusive samples are merged. In the following, “W boson + bX” denotes events with at least one jet within the acceptance that originates from fragmentation of a b parton. “W boson + cX” denotes events with at least one jet within the acceptance that originates from fragmentation of a c parton, but no jet within acceptance that originates from a b-parton fragmentation. “W + light jets” denotes the remaining events.

The relative cross sections for the exclusive production of the \(W \rightarrow l\nu\) + \(N\) additional partons processes are taken from the inclusive sample in LO+PS accuracy. These relative contributions are normalized to an inclusive NNLO cross section of \((31314 \pm 407 \pm 1504)\) pb [196, 197]. Predicted values are calculated using the CTEQ6 PDF set [52]. The resulting production cross sections are 5372 pb for the production of \(W \rightarrow l\nu\) + 1 additional parton, 1685 pb for the production of \(W \rightarrow l\nu\) + 2 additional partons, 498 pb for the production of \(W \rightarrow l\nu\) + 3 additional partons, and 201 pb for the production of \(W \rightarrow l\nu\) + 4 additional partons. Uncertainties on the W-boson-plus-jets production cross section as a function of the jet multiplicity, as well as flavor of the associated partons, are considered as a systematic uncertainty.
3 Experimental Setup

**Z-boson-plus-jets production**  Z-boson-plus-jets events are generated with MadGraph and Pythia, similar as for W-boson-plus-jets events. An invariant-dilepton mass $m_{l^+l^-}$ of at least 50 GeV/$c^2$ is required for event generation. Parameters for matching-, renormalization-, and factorization scale are chosen similar as for W-boson-plus-jets events. Z-boson-plus-jets events are generated with an inclusive-jet multiplicity, since Z-boson-plus-jets events are only a minor background contribution in this analysis.

Simulated Z-boson-plus-jets events are normalized to a production cross section of $(3048 \pm 34 \pm 128)$ pb ($m_{l^+l^-} > 50$ GeV/$c^2$) as calculated in NNLO \[196, 197\]. The CTEQ6 PDF set \[52\] is used.

**Diboson production**  WW, WZ, and ZZ processes are simulated with the Pythia 6 event generator. The CTEQ6 PDF set \[52\] is used. The simulated events are normalized to the predicted NLO cross section \[198\]. WW production is estimated to be $(47.04 \pm 2.02 \pm 1.51)$ pb, WZ production to $(18.57 \pm 1.04 \pm 0.30)$ pb, and ZZ to $(6.46 \pm 0.21)$ pb.

**QCD-multijet production**  QCD-multijet production is modeled with the Pythia 6 event generator. However, the overall cross section of QCD-multijet production is orders of magnitude higher than the signal cross section, and the probability to identify a well-isolated lepton within the processes is very low. Therefore, dedicated simulated event samples exist that are enriched with events that contain decays of b-flavored or c-flavored hadrons either into muons or electrons. Also samples that are enriched in events with electrons from photon conversions or “fake electrons” are simulated. Here, a “fake electron” refers to a photon which is matched to a track from a jet, and, thus, identified as electron. Fake-electron background is found to be negligible.

However, the statistical precision of the simulation becomes very low once lepton isolation and jet-flavor-tagging algorithms are applied, and the simulated event samples provide only a rough estimate of the kinematics of QCD-multijet events. Therefore, QCD-multijet kinematics and yield are estimated from data as described in section \[5.1\] and simulated event samples are used only to develop and cross check the QCD-multijet-estimation technique.

### 3.5.3 Summary of Simulated Events and Process Normalization

Simulated processes and generators are summarized in table \[3.3\]. The total amount of simulated events and the predicted production cross section are summarized in table \[3.4\]. All processes are generated inclusively if not marked otherwise. Additional samples are simulated with altered $Q^2$- and matching scales for single-top-quark, $t\bar{t}$, W-boson-plus-jets, and Z-boson-plus-jets processes. Furthermore, $t$-channel events are also simulated with CompHEP. These samples have varied statistical precision, but are not explicitly listed in table \[3.4\].
### 3.5 Simulation of Signal and Background Processes

<table>
<thead>
<tr>
<th>Process</th>
<th>Event Generator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single top quark</td>
<td></td>
</tr>
<tr>
<td>(t)-channel</td>
<td>POWHEG BOX [60–62, 188], COMPHEP [189]</td>
</tr>
<tr>
<td>(s)-channel</td>
<td>POWHEG BOX [60–62, 188]</td>
</tr>
<tr>
<td>(tW)-channel (Diagram Removal)</td>
<td>POWHEG BOX [60–62, 89]</td>
</tr>
<tr>
<td>(\bar{t}\bar{t})</td>
<td>MadGraph [59]</td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + jets</td>
<td>MadGraph [59]</td>
</tr>
<tr>
<td>(Z/\gamma^* \rightarrow l^+l^-) + jets ((m_{l^+l^-} &gt; 50\ GeV/c^2))</td>
<td>MadGraph [59]</td>
</tr>
<tr>
<td>Diboson</td>
<td></td>
</tr>
<tr>
<td>(WW)</td>
<td>PYTHIA 6 [63]</td>
</tr>
<tr>
<td>(WZ)</td>
<td>PYTHIA 6 [63]</td>
</tr>
<tr>
<td>(ZZ)</td>
<td>PYTHIA 6 [63]</td>
</tr>
<tr>
<td>QCD multijet</td>
<td>PYTHIA 6 [63], estimated from data</td>
</tr>
</tbody>
</table>

Table 3.3: Summary of simulated processes and used event generators. All processes are generated inclusively if not marked otherwise. QCD-multijet production is estimated from data in the measurement and simulated for cross checks.
### 3 Experimental Setup

#### Process  | \(N_{\text{Events}}\) | Cross section [pb]
---|---|---
**Single top quark**
- \(t\)-channel (\(t\)) | 3900171 | \(41.92^{+1.59}_{-0.21} +0.83_{-0.83}\) (approx. NNLO [8])
- \(t\)-channel (\(\bar{t}\)) | 1944826 | \(22.65^{+0.50}_{-0.50} +0.68_{-0.91}\) (approx. NNLO [8])
- \(s\)-channel (\(t\)) | 259971 | \(3.19^{+0.06}_{-0.06} +0.13_{-0.10}\) (approx. NNLO [9])
- \(s\)-channel (\(\bar{t}\)) | 137980 | \(1.44^{+0.01}_{-0.01} +0.06_{-0.07}\) (approx. NNLO [9])
- \(tW\)-channel (\(t\), Diagram Removal) | 814390 | \(7.87^{+0.20}_{-0.20} +0.55_{-0.57}\) (approx. NNLO [10])
- \(tW\)-channel (\(\bar{t}\), Diagram Removal) | 809984 | \(7.87^{+0.20}_{-0.20} +0.55_{-0.57}\) (approx. NNLO [10])
- \(t\bar{t}\) | 3701947 | \(157.5^{+18.0}_{-19.5} +14.7_{-14.7}\) (NLO [76, 77])

**W (\(\to l\nu\)) + inclusive jets**
- \(W (\to l\nu) + 1\) add. parton | 7604128 | \(5372^{\phantom{+0}_{-0}}\) (LO+PS [59])
- \(W (\to l\nu) + 2\) add. partons | 25424787 | \(1685^{\phantom{+0}_{-0}}\) (LO+PS [59])
- \(W (\to l\nu) + 3\) add. partons | 7685944 | \(498^{\phantom{+0}_{-0}}\) (LO+PS [59])
- \(W (\to l\nu) + 4\) add. partons | 13133738 | \(201^{\phantom{+0}_{-0}}\) (LO+PS [59])

**Z/\gamma^* (\(\to l^+l^-\)) + jets (\(m_{l^+l^-} > 50\) GeV/c^2)**
- \(Z/\gamma^* (\(\to l^+l^-\)) + jets (\(m_{l^+l^-} > 50\) GeV/c^2) | 36277961 | \(3048^{\phantom{+0}_{-0}} +34^{+128}_{-0}\) (NNLO [196, 197])

**Diboson**
- \(WW\) | 4225916 | \(47.04^{+2.02}_{-1.51}\) (NLO [198])
- \(WZ\) | 4265243 | \(18.57^{+1.04}_{-0.80}\) (NLO [198])
- \(ZZ\) | 4187885 | \(6.46^{+0.30}_{-0.21}\) (NLO [198])

**QCD multijet (\(\mu\) enriched)**
- \(p_T(\mu) > 15\) GeV/c, \(p_T > 20\) GeV/c | 25080241 | \(84679^{\phantom{+0}_{-0}}\) (LO [63])

**QCD multijet (\(e\) enriched)**
- \(b/c\)-hadrons \(\to e+X\)
  - \(p_T > 20\)–30 GeV/c | 2081560 | \(139299^{\phantom{+0}_{-0}}\) (LO [63])
  - \(p_T > 30\)–80 GeV/c | 2030033 | \(143845^{\phantom{+0}_{-0}}\) (LO [63])
  - \(p_T > 80\)–170 GeV/c | 1082691 | \(9431^{\phantom{+0}_{-0}}\) (LO [63])

**Conversions**
- \(p_T > 20\)–30 GeV/c | 35729669 | \(2502660^{\phantom{+0}_{-0}}\) (LO [63])
- \(p_T > 30\)–80 GeV/c | 70392060 | \(3625840^{\phantom{+0}_{-0}}\) (LO [63])
- \(p_T > 80\)–170 GeV/c | 8150672 | \(142814^{\phantom{+0}_{-0}}\) (LO [63])

Table 3.4: Summary of production cross sections and number of generated events. The first uncertainty is due to scale variations, and the second uncertainty due to the PDF parametrization. Combined uncertainties are quoted for diboson production. Uncertainties on the predicted cross section for QCD-multijet events are not quoted, since QCD-multijet production is estimated from data.
3.5 Simulation of Signal and Background Processes

3.5.4 Corrections to Simulated Events

Residual corrections to the simulated trigger efficiency for events with muons, the jet-transverse-momentum resolution, and a more involved trigger reweighting of events with an electron according to the trigger efficiency in data events are applied to simulated events. Furthermore, a reweighting according to the pile-up conditions in the analyzed data-taking period is performed. Each of these corrections is discussed in the following paragraphs.

**Trigger Efficiency in the Muon Channel**

Corrections to the simulated trigger efficiencies are applied. These corrections are $\eta$-dependent, but flat in $p_T$, and described in section 3.3.3.

**Cross-Trigger Efficiencies in the Electron Channel**

Events with muons are triggered if at least one isolated muon with a transverse momentum of more than 17 GeV/c can be reconstructed within the HLT. However, triggering events with prompt electrons is much more challenging. Prompt electron identification and reconstruction is less efficient and less pure as for muons, and ambiguities between jet and electron reconstruction can arise. In consequence, a large background due to QCD-multijet production occurs. Either high trigger rates or $p_T$ thresholds much higher than 30 GeV/c are required for inclusive single-isolated-electron triggers for luminosity conditions in 2011. Thus, single-isolated-electron triggers are either prescaled or have a small signal acceptance. Therefore, a trigger which requires an isolated electron ($p_T > 25$ GeV/c) and a b-flavored jet is used to record data with prompt electrons. This trigger is also referred to as “cross trigger” in the following. The cross trigger is used to record data corresponding to an integrated luminosity of $1.35 \text{fb}^{-1}$, which is about 87% of the analyzed data for events with an electron.

The HLT sequence for the cross trigger consists of three parts. One part requires an isolated lepton candidate (either a muon or an electron) with a certain transverse momentum. A second part requires at least one central jet with a corrected $p_T$ of more than 30 GeV/c. The jet is reconstructed from calorimeter deposits using the anti-$k_T$-clustering algorithm. A third part of the trigger requires that at least one jet originates from b-parton fragmentation. Here, the TCHE algorithm at a medium working point is used for jet-flavor tagging. The medium working point corresponds to a mistag (u, d, s, g partons) rate of 1%. Details of the trigger requirements and data-taking periods are part of the “event selection” and described in section 4.3.

The challenge for this analysis is that the cross trigger is modeled with a tighter jet $p_T$ threshold (40 GeV/c) in simulated events than it is actually applied in data (30 GeV/c). Therefore, the simulated cross-trigger is omitted. Instead, the cross-trigger probability is measured in data events and applied to simulated events. This procedure [199] is developed and carried out by the CMS single-top-quark group for the analyses described in ref. [24], and detailed in the following.

The general strategy to model the trigger efficiency as observed in data events is as follows. The HLT sequence is factorized into two components, a leptonic and a hadronic component. For the modeling of the leptonic component of the data-trigger sequence, a single-isolated-electron trigger is used. The simulated trigger is comparable to the lepton component of the cross trigger, and the trigger efficiency in data can be directly compared to the simulated trigger efficiency, as discussed in section 3.3.4. The hadronic component of the trigger sequence is treated in the following way. For each jet, a probability that the trigger accepted that particular jet is derived. The probability is expressed as a two-dimensional function, which depends on the b-tagging-discriminator value $\text{bid}_{\text{TCHP}}$ (as calculated offline with the TCHE algorithm) and...
3 Experimental Setup

the transverse momentum $p_T$ of the jet. The probability to accept a particular event is defined by the combinatorics over all jets in an event.

The technical details are outlined in the following paragraphs. Trigger efficiencies $\epsilon(bid_{TCHP}, p_T)$ are parametrized in 15 bins of $bid_{TCHP}$ as a Gompertz curve with

$$\epsilon(bid_{TCHP}, p_T) = a \cdot \exp \left[ b \cdot \exp \left( c \cdot p_T \right) \right]. \tag{3.18}$$

In a subset of the analyzed data, un-prescaled single-isolated-muon triggers and muon cross-triggers are enabled in parallel. This dataset corresponds to about $540 \text{ pb}^{-1}$. Therefore, the dataset with muons can be used to derive the two-dimensional parametrization in events with a muon in the final state. Since the hadronic part of both muon- and electron cross-triggers is the same, the parametrization can be applied to events with an electron in the final state.

![Figure 3.18](image.png)

**Figure 3.18**: Trigger efficiency of the hadronic part of the electron trigger, which is parametrized with a Gompertz curve in bins of jet-$p_T$ and jet-flavor-tagging-discriminator value $bid_{TCHP}$ (left plot). As an example, the obtained trigger efficiency for a jet with $3.2 < bid_{TCHP} < 3.6$ and the parametrization with a Gompertz curve is shown in the right plot. Both plots are done with values from ref. [199].

Figure 3.18 (left) shows the average trigger efficiencies of the hadronic part as a two-dimensional projection in bins of jet-$p_T$ and $bid_{TCHP}$. Offline-selection criteria are marked with (blue) straight lines. As an example, figure 3.18 (right) shows the hadronic-trigger efficiency for discriminator values around the tight working point ($bid_{TCHP} = 3.41$) and its parametrization with a Gompertz curve.

The HLT efficiency has a relatively broad turn-on for the hadronic part, and the offline criteria used to identify jets from b-quark fragmentation are within this turn-on. This broad turn-on essentially has three reasons. First, the TCHP algorithm is used for offline-tagging and parametrization, but the TCHE algorithm is used for online tagging in the HLT. Second, also the working point is different. The medium working point, which corresponds to a mistag rate of 1%, is used online. Tighter criteria with a mistag rate of 0.1% are applied offline (cf. the “event selection” which is described in section 4.3). Third, the trigger-jets are reconstructed from calorimeter deposits only. In the offline analysis, jets are reconstructed with the PF technique by combining calorimeter deposits and tracks. In particular, the jet-transverse-momentum response is worse for calorimeter jets [71].

For technical reasons, the bins do not correspond to the (asymmetric) binning as used for the parametrization in the measurement.
3.5 Simulation of Signal and Background Processes

Furthermore, an additional dependency on the number of b-flavored jets and highest discriminator value of all reconstructed jets in an event is observed [199]. The “true” trigger efficiency is found to be higher in data events than obtained in simulated events with parametrization. Therefore, additional multiplicative correction factors have to be applied for each jet. These residual corrections to the trigger-efficiency parametrization are derived with simulation [199].

The additional scale factors are approximately 15%–20% for discriminator values at the tight working-point threshold, and they are approximately 5% for jets with discriminator values sufficiently above the tight working point. Additional scale factors are below one for events with zero b-flavored jets, but larger than one for events with at least one b-flavored jet. However, they become huge for b-flavored jets with discriminator values well below the tight working-point threshold, and the b-tag-trigger efficiency cannot be sufficiently parametrized for events with exactly zero tagged jets. Hence, electron events with exactly 0-tagged jets are excluded from this analysis. Furthermore, a cross check of the additional scale factors among t̄t, t-channel, and W-boson-plus-jets events reveals that the same scale factors can be applied to every jet independently of the particular hard process.

Uncertainties on the trigger-efficiency modeling include uncertainties of trigger-efficiency parametrization and additional scale factors. The parametrization with the Gompertz curve has three free parameters. The uncertainty of this parametrization can be described by three linearly independent parametrizations which result from the eigenvectors of the covariance matrix. An upward and downward shift is derived from each of these three parametrizations. Thus, 3 ⊗ 2 parametrizations are considered. The resulting variations are less than 1%. The uncertainty of the additional scale factors is estimated to be ±5%. Here, 1 ⊗ 2 parametrizations are considered. In total, four independent parametrizations are considered as sources of systematic uncertainties for the trigger-efficiency modeling in the electron channel.

Pile-Up Reweighting  In addition to the hard interaction, which is of particular interest in this analysis, several soft interactions may take place in parallel. Those additional interactions, which are referred to as Pile-Up interactions, contaminate the event signature by adding jets, smearing the jet resolution, or worsening the lepton isolation among other effects. As described in sec. 3.3.2, CHS within the PF algorithm is used to subtract Pile-Up contributions on an event-by-event basis. Nevertheless, a simulation of Pile-Up interactions is important to account properly for the remaining particles.

The number of additional interactions are determined purely by a statistical process, i.e. luminosity dependent, which basically originates from two sources. On the one hand, additional interactions are generated in the same collision. This effect depends on the large number of protons in single bunch and the number of simultaneously colliding bunches, and is usually referred to as “In-Time Pile-Up”. On the other hand, events are contaminated by additional interactions from collisions in a bunch before or after the hard interaction of interest. This

These additional scale factors are derived by comparing the events from the default samples (SUMMER11 conditions) with a set of simulated events that belongs to a newer version of the CMS reconstruction software (FALL11 conditions). However, both versions have a different HLT menu. In particular, the studied cross trigger is simulated only in FALL11 conditions. For the comparison, trigger probabilities can be taken from simulation for FALL11 events, and probabilities of the SUMMER11 events are reweighted according to the parametrization as derived from data.

Nevertheless, the SUMMER11 conditions are optimized to the data-taking period used in this analysis, while the FALL11 conditions are optimized for the total 2011 data. In particular, the Pile-Up conditions and isolated-single-lepton triggers are different. That is why the usage of the FALL11 simulation is disfavored in this analysis, even if it the electron cross trigger is simulated in FALL11 conditions, which is in particular useful for this cross check.
effect depends on the bunch spacing, since the time interval between two bunches (50 ns) is rather small against the time resolution of the subdetector components. This process is usually referred to as “Out-Of-Time Pile-Up”. Moreover, both processes depend on the total inelastic proton-proton cross section. The CMS collaboration measured the total inelastic proton-proton cross section at $\sqrt{s} = 7$ TeV to

$$\sigma_{\text{inel.}}(pp) = (68.0 \pm 2.0 \text{ (Systematics)} \pm 2.4 \text{ (Luminosity)} \pm 4.0 \text{ (Extrapolation)}) \text{ mb}$$

(3.19)

Here, the extrapolation uncertainty is the dominant uncertainty. It refers to the model dependence that occurs when the fiducial measurement is extrapolated to an inclusive phase space. This model dependence is estimated from various simulation models and accounts for the extrapolation from events with at least two charged particles (each with $p_T > 200$ MeV/c and $|\eta| < 2.4$). Alternative measurements by the TOTEM experiment at $\sqrt{s} = 7$ TeV yield

$$\sigma_{\text{inel.}}(pp) = (73.5 \pm 0.6 \text{ (Statistics)} \pm 1.8 \text{ (Systematics)} \text{ mb}$$

(3.20)

$$\sigma_{\text{inel.}}(pp) = (73.7 \pm 0.1 \text{ (Statistics)} \pm 3.4 \text{ (Systematics)} \text{ mb}$$

(3.21)

and

$$\sigma_{\text{inel.}}(pp) = (72.9 \pm 1.5) \text{ mb}$$

(3.22)

The ATLAS experiment measures

$$\sigma_{\text{inel.}}(pp) = (69.4 \pm 2.4 \text{ (Statistics} \oplus \text{ Systematics)} \pm 6.9 \text{ (Extrapolation)} \text{ mb}$$

(3.23)

The measurements well agree within uncertainties, which are taken as systematic uncertainty into account (cf. section 6.1), since a systematic shift in the total inelastic proton-proton cross section linearly translates into a shift in the number of additional interactions.

The pixel-luminosity-calculation method (cf. section 3.4) is used to determine the instantaneous luminosity in each luminosity section. Finally, the expected number of additional interactions is calculated in each luminosity section of data.

Pile-Up interactions are simulated with minimum-bias events with the PYTHIA 6 event generator. They are mixed into simulated events of the hard process according to a pre-defined luminosity profile. A priori, it was unknown how the instantaneous luminosity of the LHC would evolve during the early data taking in 2011. A scenario with an average of ten interactions per collision was estimated before the start-up of data-taking period in 2011. Technically, simulated additional interactions are mixed into simulated events according to a certain probability distribution that is expected to match the instantaneous-luminosity conditions. In the simulated event samples that are used in this analysis, the simulated probability distribution is uniformly distributed between zero and ten additional interactions, and has a Poisson tail with mean $\mu = 10$ above ten interactions. On the one hand, this distribution still contains a significant amount of events with less than ten additional interactions, because the instantaneous luminosity decreases significantly during a fill. On the other hand, a tail with up to 25 additional interactions per event is simulated, which corresponds to a statistical process with mean $\mu = 10$.

A reweighting procedure, as implemented in the CMS Software (CMSSW) framework, is used to reweight the simulated in-time and out-of-time Pile-Up interactions to the calculated distribution of additional interactions in the analyzed data. The number of reconstructed vertices can be used as a measure of the number of additional interactions. An alternative approach of reweighting would be to fit the simulated distribution of the number of reconstructed vertices to the observed distribution in data. This approach is much more involved, and accompanied by additional systematic uncertainties, because the vertex reconstruction efficiency possibly is a non-linear function of the total number of vertices in an event.
Figure 3.19 shows the number of reconstructed vertices for events with at least two jets \( (p_T > 30 \text{ GeV}/c) \) and either an electron \( (p_T > 30 \text{ GeV}/c) \) or muon \( (p_T > 20 \text{ GeV}/c) \) in the final state. The distribution peaks at five vertices with a broad tail up to 14 or more reconstructed vertices in a single event. Both distributions are described well within uncertainties of the total inelastic proton-proton cross section (hatched area). The tail of the muon distribution shows a disagreement for a small fraction of events with very high numbers of reconstructed vertices, which can possibly be explained by a difference in the vertex-reconstruction efficiency between data and simulation for events with a high number of Pile-Up interactions. Nevertheless, this effect is accounted for within the systematic variations.

Jet-Transverse-Momentum Resolution As the jet resolution in simulated events differs from the resolution that is observed in data, corrections to the jet-transverse-momentum resolution are derived from data and applied to simulated events. This resolution correction is discussed in section 3.3.5.

Jet-Flavor-Tagging Efficiencies The jet-flavor-tagging efficiency for the TCHP tagging algorithm at the tight working point is about 10% lower in data than in simulated events (cf. section 3.3.6). Residual corrections are expressed as \( (p_T, \eta, \text{jet flavor}) \)-dependent data-simulation-scale factors

\[
SF(p_T, \eta, \text{jet flavor}) = \frac{\epsilon_{\text{Data}}}{\epsilon_{\text{Simulation}}}.
\]  

These corrections are applied to simulated events (following the idea and implementation as described in ref. [205]). Here, residual corrections are applied to b-tagged jets in simulated events. This procedure allows the analysis to profit from the excellent simulation of the jet-flavor-tagging efficiency w.r.t. jet kinematics, angular separation among jets, number of (tagged) jets in an event, and additional interactions among others, but except for the \( p_T, \eta \), and jet-flavor dependence. Each simulated event, with a total number of b-tagged jets \( N_{\text{b-tagged jets}} \),
obtains the weights (cf. [205])

\[ w_i(0 \text{ tags}) = \prod_{i=1}^{N_{\text{tagged jets}}} (1 - SF(p_T, \eta, \text{jet flavor})), \]

\[ w_i(1 \text{ tag}) = \sum_{j=1}^{N_{\text{tagged jets}}} SF(p_T, \eta, \text{jet flavor}) \prod_{i \neq j}^{N_{\text{tagged jets}}} (1 - SF(p_T, \eta, \text{jet flavor})), \]

\[ w_i(2 \text{ tags}) = \sum_{j=1}^{N_{\text{tagged jets}}} SF(p_T, \eta, \text{jet flavor}) \sum_{k=1}^{N_{\text{tagged jets}}} SF(p_T, \eta, \text{jet flavor}) \]

\[ \cdot \prod_{i \neq j, i \neq k} (1 - SF(p_T, \eta, \text{jet flavor})), \]

\[ w_i(N \text{ tags}) = \ldots, \]

which express the probability that an event contains either 0, 1, or 2 tagged jets when considering a varied jet-flavor tagging efficiency with data-simulation-scale factors $SF$.

However, in this analysis, the reconstruction of some high-level distributions depends on the assignment of one “b-tagged-jet hypothesis” among all reconstructed jets (cf. also section 4.4.1). The b-tagged-jet hypothesis represents the b-flavored jet from the top-quark decay. As an example, such distributions include the transverse momentum of the b-flavored jet, which is expected to show a distinct peak at high transverse momentum. Another example is the reconstruction of the top quark, which uses the four-vectors of b-flavored jet, charged lepton, and $E_T^{\text{miss}}$.

Ambiguities in the reconstruction arise if corrections to the jet-flavor-tagging efficiency in the simulation are applied. As an example, in case of simulated events with exactly two jets, which are both tagged, one of the tags has to be removed (“un-tagged” jet) if the jet-flavor-tagging efficiency is higher in simulation than in data. However, any criterion used to remove a tag from one of the jets can bias the reconstruction of kinematics and objects in categories with lower b-tagged-jet multiplicities. Moreover, the definition of the “b-tagged-jet hypothesis” can vary among categories with different exclusive b-tagged-jet multiplicities, i.e. the total number of b-tagged jets in an event.

Therefore, in this analysis, the reweighting algorithm is extended such that a bias-free reconstruction of kinematics (like the top-quark mass) is obtained when applying the data-simulation-scale factors for jet-flavor-tagging efficiencies. In particular, this procedure is also referred to as incorporating “shape effects” of the data-simulation-scale factors (and their uncertainties) for jet-flavor tagging.

For each simulated event, multiple event hypotheses are formulated, in which each hypothesis is weighted with $w_i$ according to the description in the previous paragraphs. The number of additional event hypotheses ($N_{\text{hypotheses}}$) for each simulated event is a function of the number
of b-tagged jets according to simulation ($N_{\text{simulation tags}}$), and is given by

$$N_{\text{hypotheses}} = \sum_{i=0}^{i_{\text{max}}} \frac{N_{\text{hypotheses}}(i)}{N_{\text{simulation tags}}(i)} \quad (3.26)$$

in which each event hypothesis has $i$ b-tagged jets and ($N_{\text{simulation tags}} - i$) jets whose tags are removed. Hypotheses with up to $i_{\text{max}} = 4$ tagged jets are considered. Each event hypothesis is weighted with an extra multiplicative weight

$$w_{\text{combinatorics}}(i) = \frac{1}{N_{\text{hypotheses}}(i)} \quad (3.27)$$

due to combinatorics, assuming that jet-flavor-tagging is statistically independent among tagged jets in simulation.

### 3.6 Software and Libraries

This section briefly describes the software tools and libraries that are used in this analysis. ROOT is used mostly for the presentation of data, the multivariate analysis is performed with the Toolkit for Multivariate Data Analysis (TMVA), CMSSW is used for event reconstruction among others, the Visual Physics Analysis (VISPA) package and the Physics eXtension Library (PXL) are used for the high-level data-analysis, and the theta framework is used for the statistical inference.

**ROOT**  
ROOT [206] is an object-oriented C++ large-scale-data analysis framework which is developed at CERN. ROOT is used in version 5.30/02.

**TMVA**  
The Toolkit for Multivariate Data Analysis (TMVA) [207] is a platform for machine-learning techniques within the ROOT framework. Machine-learning techniques for both classification and regression tasks in multivariate analyses are implemented in TMVA (version 4.1.2) is used for training and testing of BDTs, as well as for evaluating the classifier output of events.

**CMSSW**  
CMSSW [136, 153] is the general software framework of the CMS experiment. It is written in C++. CMSSW combines various aspects of data analyses in the context of high-energy physics. CMSSW (version 4.2.8) is used for event and physics-objects reconstruction in this analysis.

In general, the CMSSW framework provides an extensive interface for event simulation. CMSSW allows for a steering of event generators and parton-shower models in various programming languages, and it supports the mixing of events from the hard-scattering process with additional soft interactions. CMSSW features a full detector description and detector simulation based on GEANT4, as described in section 3.2.6, but also services for alignment and
calibration of detector components. Furthermore, event and detector visualization are provided. CMSSW also serves as a base for data-quality monitoring. Moreover, event selection, reconstruction of event-related properties, and reconstruction of physics objects, as well as the offline-high-level analysis with user code are an important part of the framework. The CMSSW framework is used also for online-HLT event filtering.

CMSSW uses the “Event Data Model”. Here, the Event is the central data container. Events may contain detector readouts such as calorimeter deposits or tracks, reconstructed physics objects, simulation truth information, trigger decisions, and HLT objects. CMSSW is a modular framework. Modules operate on Events, and they are permitted to either add information or solely read data from the event. Modules are written in C++-programming language and steered by configuration files written in PYTHON. Modules are organized in user-defined sequences which are referred to as “paths”, in which Events are processed sequentially. Modules cannot directly communicate with each other.

theta  theta is a plug-in-based framework [208] that is used for building the statistical model and the performing statistical inference. In theta, the statistical model is expressed with binned templates, i.e. as a sum of histograms. Here, an arbitrary number of bins is supported.

Several likelihood-based quantities can be calculated to evaluate the statistical model in terms of hypothesis tests or interval estimation with Bayesian or frequentist methods. A hypothesis test may have a Bayes factor or p-value as result, while interval estimation results in a Bayesian confidence interval (“credible interval”) or a confidence interval in terms of frequentist statistics. More details about the statistical methods itself are given e.g. in references [11, 209, 210] and in section 4.6.

The likelihood function usually depends on many parameters besides the parameter(s) of interest. theta supports additional nuisance parameters to the likelihood function, e.g. to incorporate the effect of systematic uncertainties into the statistical model. Diversified methods to treat (or eliminate) nuisance parameters are provided, e.g. marginalization or profiling. Integrals are solved numerically by using a Metropolis-Hastings Markov-chain-Monte-Carlo (MCMC) algorithm [211, 212]. Minimization tasks can be solved e.g. by using the MINUIT-minimization package as implemented in the ROOT framework [206].

The statistical model can be evaluated with either observed data or pseudo-experiments. Furthermore, theta facilitates an efficient way to generate many pseudo-experiments.

VISPA and PXL  The Visual Physics Analysis (VISPA) package [213-216] and the Physics eXtension Library (PXL) [213] are used as software tools for the high-level analysis. The “high-level” analysis includes any analysis step beyond the reconstruction of physics objects with loose quality criteria, i.e. the event selection, top-quark reconstruction, and categorization of events among others.

PXL is written in C++ programming language and is used as the physics library for the VISPA package. PXL provides a modular framework with clear interfaces, which ensures a re-usability of source code. The modularity simplifies the communication about an analysis structure. Even a complex analysis logic can be mapped into the module chain, in which modules can have several input or output channels. Modules are used to implement event filtering and event-flow-steering mechanisms, as well as for input and output of data. Modules operate on Events, which is the central data container for applications in the context of high-energy physics, and they are allowed to modify the content of an Event. Modules can be written in
either C++ or PYTHON programming languages. Events contain C++-objects of reconstructed physics-objects, simulation-truth particles, and vertices among others. Within PXL relations among objects can be managed. As an example, such relations can exist between generated and reconstructed particles. Another example are mother-daughter relations among particles in decay chains, e.g. two particle objects that represent identified muons can be related to a newly created particle object in order to reconstruct a Z boson. Furthermore, PXL provides a set of transformations and algorithms that are commonly used in high-energy-physics or astroparticle analyses. Interfaces are used to connect dedicated experiment-software packages (e.g. CMSSW) with PXL. Moreover, PXL provides a simple mechanism to customize the event content, which is referred to as the UserRecord. A user can adapt the event content to the needs of an experiment. Neither a recompilation of the library nor changes to the data format are required to (permanently) add a UserRecord to the event content.

The typical development cycle of a data analysis starts with the design of a prototype and continues with an execution of this analysis prototype. The next step usually is the review and verification of preliminary analysis results. This step is an intermediate step towards the final result, since the interpretation of the results may again lead to a modification of the analysis design. VISPA is a graphical development environment for large-scale-data analyses in the context of high-energy and astroparticle physics that supports, organizes, and visualizes such data-analysis cycles.

VISPA is independent of any particular experiment software. Instead, it is based on a general post-reconstruction-data format which is designed for high-level analysis. This data format is provided by the Physics eXtension Library (PXL). Due to the condensed event content of PXL objects, VISPA facilitates a fast turn-around of even complex analysis cycles. Therefore, many iterations of the analysis-development cycle are possible, while a general overview of the analysis logic is retained.

The VISPA user interface is designed as a platform-independent and plug-in-based framework. Plug-ins include analysis design and module-flow management, data-content browsing, or batch-job execution among others. User-coded plug-ins are supported such that the scope of VISPA can be extended beyond the typical tasks of high-energy or astroparticle physics. Figure 3.20 shows the VISPA user interface with one of its typical applications in the context of high-energy physics, the so-called “Analysis Designer” plug-in. The “Analysis Designer” is a graphical representation and steering interface for the underlying module chain of an analysis.

VISPA offers possibilities to save the analysis output either into a binary PXL format or to include any other library. The analysis logic is saved into a conventional XML output format, and the possibility to export entire analyses facilitates the collaboration among physicists.

Recently, VISPA is extended such that it runs in a standard Internet browser that has JAVASCRIPT enabled [216–218]. The VISPA web application implements a client-server model, in which the worker nodes need (only) standard PYTHON, and the server is realized with the CHERRYPy framework [219].
Figure 3.20: Analysis-Designer plug-in of the VISPA software package. The analysis logic is implemented with a module chain, which is shown in the central column. Each module implements either analysis algorithms, event filtering, some sort of event-flow-steering mechanism, or input/output of event data. A list of available modules is shown in the left column, and properties of the current selected module are shown in the right column.
4 Analysis Strategy

This chapter starts with a brief introduction (section 4.1) into the main characteristics of this analysis. This analysis is performed as a blind analysis in the sense that no data, but only simulated events, were used in the signal-enriched phase space during the design-phase of this analysis. However, the agreement between data and simulation was continuously checked in the signal-depleted phase space. A categorization of events is applied, in which events are classified according to their charged-lepton flavor, number of reconstructed jets, and number of b-tagged jets. A categorization of events makes it possible to reduce the impact of systematic uncertainties in situ within the statistical inference. Furthermore, categorization facilitates an enhancement of signal acceptance and an optimization of the analysis by using phase space regions with varied signal-to-background ratios. This analysis is a multivariate analysis. The signal kinematics and event topology are characterized with eleven distributions, which are simultaneously evaluated in order to discriminate between signal and background events.

The phenomenology of t-channel single-top-quark production from an experimental point of view is described in section 4.2. The analysis presented in this thesis is based on the selection of events with at least two jets and either an electron or muon. The event selection is adapted to the signal characteristics and described in detail in section 4.3. The reconstruction of the t-channel final state, i.e. the association of jets and the reconstructions of neutrino, W boson, and top quark are discussed in section 4.4. As the neutrino escapes undetected, only its transverse momentum components can be reconstructed from missing-transverse-momentum-imbalance vector. The z-component of the neutrino momentum is estimated by applying a kinematic constraint on the W boson and its decay products. In addition, an algorithm that minimally modifies the transverse-momentum-imbalance vector is applied. This algorithm further optimizes the neutrino reconstruction. Finally, the top-quark reconstruction is discussed.

The signal-to-background ratio is rather poor even after the event selection. W-boson-plus-jets and tt production are the dominating background processes due to their large cross sections. These processes mimic the signal signature quite well as they both contain either top quarks or W bosons.

Boosted Decision Trees (BDTs) are a multivariate-analysis technique. They are used to further discriminate between signal and remaining background events. The main principle of BDTs is presented in a nutshell in section 4.5. The discriminating variables, which are input to the classifier training, their separation performance, and correlations among them are analyzed. In order to avoid any bias on the final result due to the classifier-training procedure, simulated events are split into three statistically independent samples. The first sample is used for training the classifier. The second sample is used to test the separation performance of the classifier and to optimize it. The third sample, the so-called evaluation sample, is used to do the actual measurement. Furthermore, the resulting BDT-classifier distributions and their discrimination performance are presented in this section. The resulting BDT-classifier-output distributions of all 12 categories are input to the statistical inference.

Finally, this chapter concludes (section 4.6) with a definition of the statistical model that is used to infer the signal cross section from measured data. A Bayesian core method is used. The construction of the likelihood function is discussed. Systematic uncertainties are incorporated...
with additional nuisance parameters into the likelihood function. The nuisance parameters are eliminated by marginalizing them.

### 4.1 Blind Analysis and Categorization

![Diagram of event categorization](image)

Figure 4.1: Categorization of events according to their lepton flavor (e, µ), number of selected jets, and number of b-tagged jets. $t$-channel signal categories are shown in white, categories dominated by W-boson-plus-jets events are shown in brown, and categories that are dominated by $t\bar{t}$ events are shown in orange. The signal cross section is measured in 12 categories simultaneously (hatched box).

In this analysis, a blinding procedure is used to avoid a bias of the measurement result towards the expected SM cross section. This analysis was developed, prepared, and carried out in two steps. In a first step, it was checked if the simulated events describe the data in the signal-depleted phase space. Signal-depleted phase space regions include regions that are almost exclusively dominated by either W-boson-plus-jets or $t\bar{t}$ processes. It was observed that simulation and data are in reasonable agreement. Physics-objects kinematics and their properties, input variables for the multivariate classifier training, and the resulting classifier output itself were checked. Since this check was passed, the signal region was unblinded in a second step. The compatibility of simulated and observed event kinematics and properties was checked, too. No unphysical outliers were found in data. Finally, the measurements were performed.

All selected events are categorized into **12 orthogonal categories** based on the flavor of the charged lepton (e, µ) from the top-quark decay, the number of selected jets $n$-jets = 2, 3, ≥ 4, and the number of b-tagged jets $m$-b-tags = 1, ≥ 2. An overview is given in fig. 4.1.

Categories enriched in $t$-channel signal events are “2 jets, 1 btag” and “3 jets, 1 btag” (white boxes). Even if these categories are enriched in signal events, they are still dominated by background events. The dominant background processes are W-boson-plus-jets and $t\bar{t}$ events.

All other categories have few signal events only. Categories with 0-tag events are dominated by W-boson-plus-jets events (brown boxes). Categories with events with $\geq$ 2 tags and the “$\geq$ 4 jets, 1 btag” categories are dominated by $t\bar{t}$ events (orange boxes).

A second-charged-lepton veto ensures that events of electron and muon final state categories are orthogonal. The signal cross section is simultaneously measured in all 12 categories.
The signal-depleted categories are not only used to gain confidence in the background modeling, but also to constrain systematic uncertainties “in situ” in the statistical inference. As an example, a constraint on the $t\bar{t}$ production cross section is obtained by selecting events with at least two b-tagged jets. Moreover, categorization also enhances the signal acceptance, and facilitates an optimization of the analysis towards varied signal-to-background ratios. Furthermore, the 0-tag categories are used as a cross-check of both the modeling of the input variables and multivariate-discriminator output, but are not used for the inclusive cross-section measurement.

### 4.2 Phenomenology of $t$-Channel Single-Top-Quark Production

![Feynman Diagram](image)

Figure 4.2: $t$-channel single-top-quark Feynman diagram and corresponding event signature as reconstructed in the detector. The 2\textsuperscript{nd} b quark mostly is out of the detector acceptance.

The $t$-channel single-top-quark event signature (cf. fig. 4.2) is composed of three distinctive features. These characteristics distinguish $t$-channel production from $s$-, and $tW$-channel single-top-quark production, as well as from top-quark-pair production.

First, the single-top-quark-production cross section approximately is two times larger than the single-top-antiquark-production cross section (cf. sec. 2.2.3). In practice, the charge of the produced top (anti-)quark is tagged by measuring the charge of the lepton that stems from the top-(anti)quark decay. The ratio of both production modes is sensitive to the description of the proton content, i.e. the proton PDFs.

Second, the light jet that arises from u, d, s, or c parton fragmentation can be very close to the beamline, up to a pseudo-rapidity of $|\eta| \approx 4.5$. The light jet is also referred to as “spectator jet”. It balances the heavy top-quark. Thus, it is not only forward, but also has a large transverse momentum. Moreover, the light jet is often also the leading jet, i.e. the jet with the highest transverse momentum, and competes with the b parton, which usually carries a lot of transverse momentum due to the large top-quark mass. In fact, the light jet typically also carries even more $p_T$ than the lepton from the top-quark decay.

Third, the produced top-(anti)quarks are highly polarized (cf. sec. 2.2.3). This polarization leads to characteristic angular correlations among its decay products even on reconstruction level (cf. e.g. [24] [102] [110] [111]).

The b quark from the top-quark decay has a large transverse momentum and is central in the detector. It hadronizes to a jet which can be identified as originating from a b parton by using dedicated b-tagging algorithms.
The charged lepton from the W-boson decay, which itself stems from the decaying top quark, carries also a significant amount of transverse momentum. Usually, a $p_T$ threshold of at least 20 GeV/$c$ is required. The large $p_T$ motivates the requirement of a lepton with a certain $p_T$ threshold already at the trigger level. In background processes, leptons with less transverse momentum are involved, except leptons from top-quark-pair production.

The neutrino from the W-boson decay escapes undetected and induces an overall transverse-momentum imbalance in an event. Its momentum $z$-component is estimated by imposing the kinematic constraint of four-momentum conservation to the W boson and its two daughter particles.

A 2\textsuperscript{nd} b quark, which is also referred to as “spectator” b quark, arises from the gluon splitting in the initial state.\footnote{Technically speaking, if events are simulated in the 5-flavor scheme, in which the b quark is taken from the proton PDFs, the 2\textsuperscript{nd} b quark is generated within the parton shower (cf. also the discussion in section \ref{sec:parton}).} The 2\textsuperscript{nd} b quark is mostly collinear with the proton remnants or has a low transverse momentum. The jet that results from fragmentation of this quark is mostly out of the detector acceptance.

Most t-channel events have two jets after event selection (cf. fig. 4.3). This category has the best signal-to-background ratio. Events may have less than two jets due to acceptance effects or additional jets due to initial-state or final-state radiation.

### 4.3 Event Selection

The event selection initially starts with the decision if a $pp$ collision should be processed by the computing facilities and, finally, recorded to disk. The processing of events involves the reconstruction of physics objects from sub-detector readouts, which typically is a time-consuming process. Already at this level, a data selection is necessary since no existing computing and storage system is feasible to cope with such a huge amount of data being delivered by the LHC (cf. also section \ref{sec:computing}). Only data from luminosity sections, in which the corresponding trigger was un-prescaled, are used.

In the offline event selection, events with exactly one electron or muon and at least two jets are selected in order to enrich the event sample corresponding to the $t$-channel single-top-quark event topology (fig. 4.2). An algorithm to identify jets originating from b-quark fragmentation is used to suppress background contributions from QCD-multijet and W-boson-plus-light-jets events. In order to further suppress QCD-multijet events, a significant amount of $E_T^{\text{miss}}$ is required for events with an electron in the final state. $E_T^{\text{miss}}$ is present in signal events due to the undetected neutrino. Alternatively, a significantly large transverse mass of the (reconstructed) W boson is required for events with a muon in the final state. In addition, each event must satisfy cleaning, trigger, and vertex-compatibility criteria.

**Trigger** For events with muons, a trigger that initiates the processing of a collision if at least one isolated HLT muon with a transverse momentum of $\geq 15$ GeV/$c$\footnote{HLT_IsoMu15_v5} is used for simulated events with a muon final state. The trigger threshold is $\geq 17$ GeV/$c$ for data events. A summary of the triggers with corresponding primary datasets, run ranges, and resulting integrated luminosities for data events with muon final states is given in table \ref{tab: triggers}.

The trigger criteria are quite similar in data and simulation for events with muons. Differences in the simulated trigger efficiency compared to the efficiency measured in data are taken…
into account (cf. section 3.3.3). Systematic uncertainties of the trigger-efficiency estimation are discussed in section 6.1.9.

<table>
<thead>
<tr>
<th>Dataset / Trigger</th>
<th>( \int \mathcal{L} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>SingleMu May10ReReco (Runs 160431 – 163869 incl.)</td>
<td></td>
</tr>
<tr>
<td>HLT_IsoMu17_v5</td>
<td>46.4 pb(^{-1})</td>
</tr>
<tr>
<td>HLT_IsoMu17_v6</td>
<td>164.6 pb(^{-1})</td>
</tr>
<tr>
<td>SingleMu PromptReco v4 (Runs 165088 – 167913 incl.)</td>
<td></td>
</tr>
<tr>
<td>HLT_IsoMu17_v8</td>
<td>136.4 pb(^{-1})</td>
</tr>
<tr>
<td>HLT_IsoMu17_v9</td>
<td>560.1 pb(^{-1})</td>
</tr>
<tr>
<td>HLT_IsoMu17_v10</td>
<td>4.4 pb(^{-1})</td>
</tr>
<tr>
<td>HLT_IsoMu17_v11</td>
<td>253.7 pb(^{-1})</td>
</tr>
<tr>
<td>total</td>
<td>1165.6 pb(^{-1})</td>
</tr>
</tbody>
</table>

Table 4.1: Used datasets with corresponding triggers and integrated luminosities for events with muon final states.

The HLT strategy for events with **electron** final states is different between simulation and data. In simulation, the HLT fires for events if at least one isolated HLT electron with a transverse energy of \( \geq 27 \text{ GeV} \) is found. However, in data events, a similar trigger is used only for the initial data-taking period (215.7 pb\(^{-1}\)). For the remaining data corresponding to an integrated luminosity of 1345.0 pb\(^{-1}\), triggers that require at least one (isolated) electron with \( E_T \geq 25 \text{ GeV} \) and a b-tagged jet with \( E_T \geq 30 \text{ GeV} \), are used. The jet itself has to be identified as originating from b-quark fragmentation with the TCHE algorithm [173].

In order to apply the same trigger criteria in data and simulation for events with electrons, trigger efficiencies for the jet criteria are derived in data events and applied to simulated events (cf. section 3.5.4). Trigger efficiencies for electron identification and isolation are measured in data and simulated events are reweighted accordingly (cf. 3.3.4).

### Event cleaning

Event-cleaning filters are applied to both data and simulated events. Filters against beam background and anomalous calorimeter noise are applied.

Events that are contaminated by particles from **beam background** are vetoed using a dedicated filter. Here, “beam background” refers to any interactions that occur between parts of the beam and the storage ring, i.e. protons interacting with remaining gas particles or beam collimators. This machine-induced background consists of beam-halo particles with trajectories parallel to the beamline. Beam-halo particles can produce secondary particle showers if they interact with the detector. Only a very few events are vetoed for the dataset at hand.

Furthermore, a filter that vetoes events with **anomalous noise** from the barrel and endcap sub-detectors of the hadron calorimeter is applied. “Anomalous noise” refers to noise with up to \( \mathcal{O}(\text{TeV}) \) that occurs when converting the scintillation light to an electrical output. The
Characteristics of this type of noise are rather short pulses, while noise from electronics has a rather constant profile.

**Vertex selection** In each event, at least one primary vertex is required. If more than one primary vertex is reconstructed, the leading primary vertex has to meet all criteria to ensure its compatibility with the primary hard interaction. The “leading primary vertex” is the vertex that has the highest sum of squared transverse momenta of tracks associated to it.

A primary vertex is reconstructed from a track fit with a group of tracks corresponding to at least four degrees of freedom (cf. sec. 3.3.1). In addition, the primary vertex has to be within an absolute distance to the nominal interaction point of smaller than $24.0$ cm in $z$-direction, and a cylindrical transverse component $\rho$ of less than $2.0$ cm.

**Muons** Muons with a transverse momentum larger than 20 GeV/c and an absolute pseudorapidity $|\eta| < 2.1$ are considered as final-state muons.

Beam-position compatibility and relative isolation quality criteria ensure that muons from the primary hard interaction, i.e. from a leptonically decaying W boson or Z boson, are selected. Each muon has to have an absolute value of the impact parameter $d_{\text{Beamposition}} < 0.02$ cm. Here, $d_{\text{Beamposition}}$ is calculated w.r.t. the (average) beam position in the transverse plane. The relative isolation $\text{relIso}$ (cf. sec. 3.3.3) of the muon is calculated within a cone of $\Delta R = 0.4$, and must be less than 0.15.

A typical background to prompt muons are muons from secondary decays of hadrons, which itself stem from parton fragmentation. This process is also referred to as “hadron decay in flight”. Moreover, hadrons that pass the absorber material, or even interact with it weakly, up to the muon system can also mimic the signature of prompt muons (“hadron punch-through”).

---

Table 4.2: Used datasets with corresponding triggers and integrated luminosities for events with electron final states. The “⋆” translates to “CaloIdVT_CaloIsoT_TrkIdT_TrkIsoT” and the “†” to “CaloIdVT_TrkIdT”.

<table>
<thead>
<tr>
<th>Dataset / Trigger</th>
<th>$\mathcal{L}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SingleElectron May10ReReco (Runs 160431 – 163869 incl.)</td>
<td></td>
</tr>
<tr>
<td>HLT_Ele27_⋆_v1</td>
<td>6.3 pb$^{-1}$</td>
</tr>
<tr>
<td>HLT_Ele27_⋆_v2</td>
<td>40.7 pb$^{-1}$</td>
</tr>
<tr>
<td>HLT_Ele27_⋆_v3</td>
<td>168.7 pb$^{-1}$</td>
</tr>
<tr>
<td>ElectronHad PromptReco v4 (Runs 165088 – 165633 incl.)</td>
<td></td>
</tr>
<tr>
<td>HLT_Ele25_†_CentralJet30_BTagIP_v4</td>
<td>136.4 pb$^{-1}$</td>
</tr>
</tbody>
</table>

| ElectronHad PromptReco v4 (Runs 165634 – 167913 incl.)     |                  |
| HLT_Ele25_⋆_CentralJet30_BTagIP_v1                        | 541.5 pb$^{-1}$  |
| HLT_Ele25_⋆_CentralJet30_BTagIP_v2                        | 277.3 pb$^{-1}$  |
| HLT_Ele25_⋆_CentralJet30_BTagIP_v4                        | 389.9 pb$^{-1}$  |
| **total**                                                  | **1560.7 pb$^{-1}$** |
Therefore, the following track-quality criteria are applied to suppress these types of background.

The muon reconstruction algorithm must have classified the muon as a global prompt muon (cf. sec. 3.3.3). It must be classified as both a global muon and a tracker muon. The fit with the best quality, i.e. the lowest \( \chi^2 / \text{Number of degrees of freedom (n.d.o.f.)} \) is used. A global-track-fit quality of \( \chi^2 / \text{n.d.o.f.} < 10 \) is required for the best fit. Furthermore, more than ten hits in the inner tracking system, at least one pixel hit for the inner track, and matches to segments in at least two muon stations are required.

Electrons

Electrons with a transverse energy larger than 30 GeV for \( |\eta| < 2.5 \) are considered as final-state electrons. Electrons within the transition region between barrel and end cap in the ECAL corresponds, which corresponds to a pseudo-rapidity range of \( 1.4442 < |\eta| \) (supercluster) \( < 1.5660 \), are excluded from this analysis.

An absolute value of the impact parameter \( d_{\text{Beamposition}} \) of less than 0.02 cm is required. \( d_{\text{Beamposition}} \) is calculated w.r.t. the (average) beam position in the transverse plane. A final-state electron has to pass identification criteria, except isolation and conversion rejection criteria, at a working point with about 70% efficiency. Its relative isolation \( \text{relIso} \) within a cone of \( \Delta R = 0.4 \) must be less than 0.125.

In order to reject electrons from photon conversions, the number of missing inner layers must be zero. Furthermore, the absolute minimum distance to potential “partner” tracks \( |\text{convDist.}| \) has to be larger than 0.02 cm or the \( |\Delta \cot \theta| \) between the electron track and the partner track at a conversion vertex has to be larger than 0.02.

Second charged lepton veto

In order to reduce the background from \( t\bar{t} \) dilepton and Drell-Yan processes, events including electrons or muons besides the final-state charged lepton are rejected. A muon qualifies as an additional lepton if it has a \( p_T > 10 \) GeV/c within \( |\eta| < 2.5 \), \( \text{relIso} < 0.2 \), and if it is reconstructed as a global muon. An electron qualifies as an additional lepton if it has \( E_T > 15 \) GeV for \( |\eta| < 2.5 \) and \( \text{relIso} < 0.2 \). The relative isolation \( \text{relIso} \) is calculated using a cone of \( \Delta R = 0.4 \).

Jets

At least two \( p_T \) jets, each with a calibrated transverse momenta larger than 30 GeV/c for \( |\eta| < 4.5 \), are required in every event. The jet reconstruction is done with the anti-\( k_T \) algorithm [170] using a distance parameter of 0.5. Reconstructed jets are built from more than one constituent, and charged-hadron subtraction (cf. sec. 3.3.2) is used to reduce the influence of Pile-Up.

Further quality criteria are applied to reconstructed jets in order to suppress fake jets reconstructed from deposits in either HCAL or ECAL exclusively. As an example, fake jets might be reconstructed from prompt electrons from the hard scattering. A jet must have both neutral-electromagnetic and neutral-hadronic energy fraction smaller than 0.99. Furthermore, a jet with an axis reconstructed in \( |\eta| < 2.4 \) must have a charged-electromagnetic-energy fraction smaller than 0.99, and a charged-hadron-energy fraction and charged multiplicity greater than 0. Jets that are reconstructed from noise in the calorimeter are rejected, too.

Figure 4.3 (left) shows the normalized jet-multiplicity distribution for reconstructed jets with \( p_T > 30 \) GeV/c and \( |\eta| < 4.5 \). The distribution is exemplarily shown for \( t\bar{t} \)-channel events with an isolated muon. Approximately 50% of the events have two jets. Events with more than two
jets are accepted in this analysis, since events with less than two jets suffer from the enormous background contributions due to W-boson-plus-jets processes.

![Figure 4.3: Jet-multiplicity distribution for t-channel events for events with one isolated muon from the top-quark-decay (left) and b-tagged-jet-multiplicity distribution (right).](image1)

**Jets originating from b-quark fragmentation**  Jets originating from b-quark fragmentation are tagged using the discriminator $b_{idTCHP}$ of the TCHP algorithm [173] at a “tight” working point ($b_{idTCHP} > 3.41$, cf. also sec. 3.3.6). The tagging efficiency for signal events is $\approx 35\%$ at the chosen working point. Figure 4.3 (right) shows the b-tagged-jet-multiplicity distribution for events with a muon in the final state.

The tight working point corresponds to a mistag probability close to 0.1\% at an average jet $p_T$ of 80 GeV/c [173]. A “mistag” refers to a false positive decision of the b-tagging algorithm, i.e. tagging jets originating from u, d, c, and g partons instead of jets originating from b-quark fragmentation. A b-tagged jet has to meet all jet criteria but is limited to $|\eta| < 2.4$ due to the tracking acceptance. The tight working point is used to suppress events from W-boson-plus-jets production. Processes, in which W bosons are mainly produced in association with jets from u, d, s, c, and g partons, have a three magnitudes larger cross section than signal events.

**$E_T^{miss}$ and transverse W-boson mass**  For events with a leptonically decaying W boson, a significant amount of $E_T^{miss}$ is typically measured due to the undetected neutrino. Moreover, a Jacobian peak in the $M_T$(W boson) distribution is expected. QCD-multijet events accumulate at low values of reconstructed $E_T^{miss}$ and $M_T$(W boson). In order to reject QCD-multijet events, $E_T^{miss} > 35$ GeV is required in events with electron final states, while a reconstructed transverse W-boson mass of $M_T$(W boson) > 40 GeV/$c^2$ is required in events with muon final states.

The transverse mass of a particle that decays into two particles $i = 1, 2$ is defined as

$$M_T = [E_{T,1} + E_{T,2}]^2 - [(p_{x,1} + p_{x,2})^2 + (p_{y,1} + p_{y,2})^2] \leq M.$$  \hspace{1cm} (4.1)

as in ref. [11].
Event Yield  Figure 4.4 summarizes the categorization of $t$-channel events according to their jet- and b-tagged-jet multiplicities. Values are exemplarily determined for simulated events with an isolated muon from the top-quark decay. Events must have passed the full event selection, in which at least two reconstructed jets with at least one b-tagged jet are required. In other categories, the background contributions are overwhelming, but serve as control regions as explained in section 4.1.

![Figure 4.4: Categorization of accepted $t$-channel events in jet- and b-tagged-jet multiplicities. Values are determined for simulated events that pass the full event selection and have one isolated muon from the top-quark decay.](image)

Detailed summaries of the event yields after the full event selection are given for events with a muon in the final state in table 4.3. Table 4.4 summarizes the event yields for events with an electron in the final state. The QCD-multijet event yield is estimated from data events (cf. section 5.1). The dominating background contributions are $t\bar{t}$ production and W-boson production with at least one jet within the acceptance that originates from b-quark or c-quark fragmentation. QCD-multijet events are an important source of background events as well. Di-boson processes, Z-boson-plus-jets production, and other single-top-quark-production modes are a minor contribution among all background events. An excess of observed data events w.r.t. the SM prediction is observed for events with a muon in the final state. The central values of event yields for the electron channel are in good agreement among observed and expected values. The data excess is treated in situ in the statistical inference (cf. section 4.6). The agreement between simulation and observed data differs for events with muon and electron final states. These differences are attributed to the uncertainty in the trigger-efficiency estimation for events in the electron channel (cf. section 3.5.4). Furthermore, the composition of signal and background events is altered in both decay channels due to the b-tagging requirement within the trigger, and background processes are affected by different normalization uncertainties.
## 4 Analysis Strategy

<table>
<thead>
<tr>
<th>Category</th>
<th>2 jets, 1 btag</th>
<th>2 jets, 2 btags</th>
<th>3 jets, 1 btag</th>
<th>≥ 2 jets, 1 btag</th>
<th>≥ 4 jets, 1 btag</th>
<th>≥ 2 jets, ≥ 2 btags</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t)-Channel (top quark)</td>
<td>440.08</td>
<td>11.97</td>
<td>170.51</td>
<td>32.84</td>
<td>54.71</td>
<td>14.90</td>
</tr>
<tr>
<td>(t)-Channel (top antiquark)</td>
<td>249.45</td>
<td>8.21</td>
<td>88.10</td>
<td>20.41</td>
<td>27.95</td>
<td>7.98</td>
</tr>
<tr>
<td>(s)-Channel (top quark)</td>
<td>34.64</td>
<td>8.22</td>
<td>12.41</td>
<td>2.76</td>
<td>4.17</td>
<td>1.04</td>
</tr>
<tr>
<td>(s)-Channel (top antiquark)</td>
<td>17.30</td>
<td>5.51</td>
<td>5.15</td>
<td>1.75</td>
<td>2.02</td>
<td>0.63</td>
</tr>
<tr>
<td>(tW)-Channel (top quark)</td>
<td>81.34</td>
<td>2.54</td>
<td>92.24</td>
<td>8.88</td>
<td>67.49</td>
<td>11.55</td>
</tr>
<tr>
<td>(tW)-Channel (top antiquark)</td>
<td>77.51</td>
<td>2.10</td>
<td>89.28</td>
<td>6.20</td>
<td>67.02</td>
<td>11.84</td>
</tr>
<tr>
<td>(t\bar{t})</td>
<td>996.47</td>
<td>145.61</td>
<td>2076.15</td>
<td>507.72</td>
<td>3073.24</td>
<td>909.54</td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + light jets</td>
<td>179.62</td>
<td>5.43</td>
<td>81.57</td>
<td>0.29</td>
<td>21.49</td>
<td>0.01</td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + bX</td>
<td>955.82</td>
<td>53.42</td>
<td>390.39</td>
<td>30.02</td>
<td>147.40</td>
<td>20.07</td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + cX</td>
<td>669.78</td>
<td>3.78</td>
<td>218.12</td>
<td>0.21</td>
<td>70.45</td>
<td>1.17</td>
</tr>
<tr>
<td>(Z/\gamma^* \rightarrow l^+l^-) + light jets</td>
<td>10.74</td>
<td>0.00</td>
<td>3.59</td>
<td>0.00</td>
<td>0.38</td>
<td>0.00</td>
</tr>
<tr>
<td>(Z/\gamma^* \rightarrow l^+l^-) + bX</td>
<td>116.37</td>
<td>4.19</td>
<td>47.48</td>
<td>2.69</td>
<td>16.18</td>
<td>0.77</td>
</tr>
<tr>
<td>(Z/\gamma^* \rightarrow l^+l^-) + cX</td>
<td>19.40</td>
<td>0.00</td>
<td>8.54</td>
<td>0.00</td>
<td>3.50</td>
<td>0.00</td>
</tr>
<tr>
<td>WW</td>
<td>19.32</td>
<td>0.09</td>
<td>9.38</td>
<td>0.32</td>
<td>4.45</td>
<td>0.20</td>
</tr>
<tr>
<td>WZ</td>
<td>22.38</td>
<td>2.77</td>
<td>6.87</td>
<td>1.28</td>
<td>1.62</td>
<td>0.29</td>
</tr>
<tr>
<td>ZZ</td>
<td>1.86</td>
<td>0.27</td>
<td>0.55</td>
<td>0.10</td>
<td>0.30</td>
<td>0.07</td>
</tr>
<tr>
<td>QCD multijet</td>
<td>260.51</td>
<td>5.54</td>
<td>87.40</td>
<td>10.65</td>
<td>31.57</td>
<td>23.44</td>
</tr>
<tr>
<td><strong>Total expected</strong></td>
<td><strong>4152.59</strong></td>
<td><strong>259.66</strong></td>
<td><strong>3387.72</strong></td>
<td><strong>626.13</strong></td>
<td><strong>3593.95</strong></td>
<td><strong>1003.49</strong></td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td><strong>4776</strong></td>
<td><strong>290</strong></td>
<td><strong>3804</strong></td>
<td><strong>695</strong></td>
<td><strong>3817</strong></td>
<td><strong>1128</strong></td>
</tr>
</tbody>
</table>

Table 4.3: Event yields after the full event selection for events with a muon in the final state. Expected event yields are scaled to an integrated luminosity of \(1165.6 \text{ pb}^{-1}\).
### 4.3 Event Selection

<table>
<thead>
<tr>
<th></th>
<th>2 jets, 1 btag</th>
<th>2 jets, 2 btags</th>
<th>3 jets, 1 btag</th>
<th>3 jets, ≥ 2 btags</th>
<th>≥ 4 jets, 1 btag</th>
<th>≥ 4 jets, ≥ 2 btags</th>
</tr>
</thead>
<tbody>
<tr>
<td>(t)-Channel (top quark)</td>
<td>263.52</td>
<td>9.54</td>
<td>110.48</td>
<td>24.26</td>
<td>36.53</td>
<td>11.89</td>
</tr>
<tr>
<td>(t)-Channel (top antiquark)</td>
<td>151.29</td>
<td>6.41</td>
<td>61.61</td>
<td>13.53</td>
<td>23.92</td>
<td>6.80</td>
</tr>
<tr>
<td>(s)-Channel (top quark)</td>
<td>24.30</td>
<td>5.78</td>
<td>8.44</td>
<td>2.47</td>
<td>2.46</td>
<td>1.19</td>
</tr>
<tr>
<td>(s)-Channel (top antiquark)</td>
<td>9.69</td>
<td>2.07</td>
<td>3.00</td>
<td>1.02</td>
<td>1.74</td>
<td>0.33</td>
</tr>
<tr>
<td>(tW)-Channel (top quark)</td>
<td>59.49</td>
<td>2.38</td>
<td>72.74</td>
<td>6.07</td>
<td>54.32</td>
<td>10.80</td>
</tr>
<tr>
<td>(tW)-Channel (top antiquark)</td>
<td>60.64</td>
<td>2.57</td>
<td>70.48</td>
<td>7.27</td>
<td>54.19</td>
<td>10.54</td>
</tr>
<tr>
<td>(t\bar{t})</td>
<td>875.53</td>
<td>134.01</td>
<td>1730.38</td>
<td>435.31</td>
<td>2668.69</td>
<td>875.39</td>
</tr>
<tr>
<td>(W (\to l\nu) + \text{light jets})</td>
<td>109.35</td>
<td>0.08</td>
<td>47.38</td>
<td>0.04</td>
<td>18.96</td>
<td>0.00</td>
</tr>
<tr>
<td>(W (\to l\nu) + bX)</td>
<td>524.29</td>
<td>30.02</td>
<td>271.15</td>
<td>25.35</td>
<td>114.20</td>
<td>17.04</td>
</tr>
<tr>
<td>(W (\to l\nu) + cX)</td>
<td>326.57</td>
<td>1.31</td>
<td>113.63</td>
<td>1.30</td>
<td>45.70</td>
<td>0.35</td>
</tr>
<tr>
<td>(Z/\gamma^* (\to l^+l^-) + \text{light jets})</td>
<td>4.51</td>
<td>0.00</td>
<td>3.73</td>
<td>0.00</td>
<td>2.10</td>
<td>0.00</td>
</tr>
<tr>
<td>(Z/\gamma^* (\to l^+l^-) + bX)</td>
<td>18.58</td>
<td>0.00</td>
<td>12.49</td>
<td>1.09</td>
<td>8.42</td>
<td>1.88</td>
</tr>
<tr>
<td>(Z/\gamma^* (\to l^+l^-) + cX)</td>
<td>4.82</td>
<td>0.00</td>
<td>1.33</td>
<td>0.00</td>
<td>1.62</td>
<td>0.00</td>
</tr>
<tr>
<td>(WW)</td>
<td>10.47</td>
<td>0.10</td>
<td>5.28</td>
<td>0.54</td>
<td>3.18</td>
<td>0.21</td>
</tr>
<tr>
<td>(WZ)</td>
<td>12.35</td>
<td>1.74</td>
<td>4.68</td>
<td>1.00</td>
<td>2.22</td>
<td>0.26</td>
</tr>
<tr>
<td>(ZZ)</td>
<td>0.53</td>
<td>0.06</td>
<td>0.38</td>
<td>0.03</td>
<td>0.18</td>
<td>0.03</td>
</tr>
<tr>
<td>QCD multijet</td>
<td>239.94</td>
<td>4.49</td>
<td>183.77</td>
<td>36.40</td>
<td>90.96</td>
<td>95.77</td>
</tr>
<tr>
<td><strong>Total expected</strong></td>
<td><strong>2695.87</strong></td>
<td><strong>200.56</strong></td>
<td><strong>2700.95</strong></td>
<td><strong>555.69</strong></td>
<td><strong>3129.41</strong></td>
<td><strong>1032.49</strong></td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td>2682</td>
<td>232</td>
<td>2774</td>
<td>541</td>
<td>3101</td>
<td>1034</td>
</tr>
</tbody>
</table>

Table 4.4: Event yields after the full event selection for events with an electron in the final state. Expected event yields are scaled to an integrated luminosity of 1560.7 pb\(^{-1}\).
4.4 Single-Top-Quark Event Reconstruction

![Feynman diagram of a top-quark that decays into W boson and a b quark, whereas the W boson decays leptonically. The four-vector of the b-tagged jet, which evolves from b-quark fragmentation, and the four-vector of the lepton are directly measured in the detector. The neutrino escapes undetected. Its transverse momentum components are measured as transverse-momentum imbalance in an event.]

Figure 4.5: Feynman diagram of a top-quark that decays into W boson and a b quark, whereas the W boson decays leptonically. The four-vector of the b-tagged jet, which evolves from b-quark fragmentation, and the four-vector of the lepton are directly measured in the detector. The neutrino escapes undetected. Its transverse momentum components are measured as transverse-momentum imbalance in an event.

In each analyzed event, a top-quark candidate is reconstructed by adding the four-vectors of the (reconstructed) W boson \( \vec{P}_{W \text{ boson}} \) and a “b-tagged-jet” hypothesis \( \vec{P}_{b\text{-tagged jet}} \). The W boson itself is reconstructed from the final-state charged lepton \( \vec{P}_{\text{lepton}} \) and the transverse-momentum imbalance \( \vec{p}_{T\text{miss}} \), since the neutrino escapes undetected. A dedicated algorithm is used to estimate the momentum vector of the neutrino. The more often the “b-tagged-jet” hypothesis matches the “true” jet from the b quark out of the top-quark decay, and the better the reconstruction of the W boson is, the more precise is the reconstruction of the top-quark candidate. A precise top-quark reconstruction is expected to improve the signal-to-background separation.

The first paragraph of the current section explains the “b-tagged-jet” hypothesis. After that, the neutrino and W-boson reconstruction and, finally, the top-quark reconstruction are discussed.

4.4.1 “b-tagged-jet” Hypothesis

Among all final-state jets in an event, a “b-tagged-jet hypothesis” and a “light-jet hypothesis” are introduced. These hypotheses are assigned on an event-per-event basis and their assignments depend on the b-tagged-jet multiplicity as explained in the following.

In categories with exactly zero b-tagged jets, the jet that is closest to the beamline is used as the “light-jet” hypothesis. Among all other jets, the jet with the highest b-tagging discriminator value \( bid_{TCHP} \) is chosen as the “b-tagged-jet” hypothesis.

In categories with exactly one b-tagged jet, the tagged jet is used as the “b-tagged-jet” hypothesis. Here, the jet closest to the beamline, among all jets but the b-tagged jet, is used as the “light-jet” hypothesis.

Categories with two or more b-tagged jets are \( t\bar{t} \) enriched phase space and hardly include any signal events. The “light-jet” hypothesis is altered in these categories in order to separate \( t\bar{t} \) from W-boson-plus-jets events rather than optimizing a correct assignment according to the single-top-quark event topology. In these categories, the tagged jet with the highest \( bid_{TCHP} \) is used as the “b-tagged-jet” hypothesis, and the jet with the second highest \( bid_{TCHP} \) is used as the “light-jet” hypothesis.

Matching efficiencies between the b-tagged-jet hypothesis and the b parton from the top-quark decay are summarized in table 4.5. The uncertainty intervals, which are spanned by the lower and upper uncertainties, correspond to the lower and upper bounds of Clopper-Pearson
4.4 Single-Top-Quark Event Reconstruction

Confidence intervals \([220]\) at 68.3% CL. These efficiencies are determined for \(t\)-channel events with a muon final state in the “2 jets, 1 btag” category, and they are obtained using simulated events. As matching criterion, \(\Delta R(\text{parton, jet axis}) < 0.3\) is used.

The b-tagged jet matches in the majority of the events the b quark from the top-quark decay. In about 10% of the events, the b-tagged jet also matches the 2\(^{\text{nd}}\) b quark from the initial gluon splitting. A good assignment efficiency is obtained.

The light-jet hypothesis matches the spectator quark in most of the events. In about 7% of the events, the light jet stems from additional radiation or Pile-Up interactions. An overall good assignment efficiency is obtained for the light jet, too.

The b quarks from the top-quark decay and the 2\(^{\text{nd}}\) b quark itself are mostly well separated. On parton level, they are collinear \((\Delta R(\text{partons}) < 0.3)\) in \((0.72 \pm 0.04)\%\) of all events. The reconstructed b-tagged jet\(^4\) rarely matches both the b quark from top-quark decay and the 2\(^{\text{nd}}\) b quark from the initial gluon splitting, namely in \((1.04 \pm 0.04)\%\) of all events.

<table>
<thead>
<tr>
<th>b-tagged jet matches . . .</th>
<th>Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>b quark from top-quark decay</td>
<td>((87.93^{+0.13}_{-0.14})%)</td>
</tr>
<tr>
<td>2(^{\text{nd}}) b quark from initial gluon splitting</td>
<td>((9.76 \pm 0.12)%)</td>
</tr>
<tr>
<td>the light spectator quark</td>
<td>((0.46 \pm 0.03)%)</td>
</tr>
<tr>
<td>none of the above</td>
<td>((1.86 \pm 0.06)%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>light-jet matches . . .</th>
<th>Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>b quark from top-quark decay</td>
<td>((4.53 \pm 0.09)%)</td>
</tr>
<tr>
<td>2(^{\text{nd}}) b quark from initial gluon splitting</td>
<td>((4.91 \pm 0.09)%)</td>
</tr>
<tr>
<td>the light spectator quark</td>
<td>((83.67 \pm 0.15)%)</td>
</tr>
<tr>
<td>none of the above</td>
<td>((6.88^{+0.11}_{-0.10})%)</td>
</tr>
</tbody>
</table>

Table 4.5: Matching efficiencies for b-tagged-jet hypothesis and light-jet hypothesis. The efficiencies are determined for simulated \(t\)-channel events with a muon final state in the “2 jets, 1 btag” category.

4.4.2 Neutrino and W-Boson Reconstruction

In the analysis at hand, W bosons decaying into an electron or muon and a neutrino are considered,

\[
W^- \rightarrow \bar{\nu}_l l^- \quad \text{and} \quad W^+ \rightarrow \nu_l l^+ \tag{4.2} \\
\text{with } l = (e, \mu). \tag{4.3}
\]

A W boson decays into an electron or muon in \(\approx 20\%\) of all cases (cf. table 2.7).

As the neutrino escapes undetected, the transverse-momentum imbalance \(\vec{p}_T^{\text{miss}}\) is used as a starting point for the reconstruction of the neutrino-momentum vector. The conservation of four-momentum gives

\[
\vec{P}_{W \text{boson}} = \vec{P}_\nu + \vec{P}_l \tag{4.5} \\
\vec{P}_{W \text{boson}}^2 = m_W^2. \tag{4.6}
\]

\(^4\)A radius parameter of \(R = 0.5\) is used for jet clustering.
4 Analysis Strategy

Applying a W-boson-mass constraint $m_W = 80.4 \text{ GeV/c}^2$ \cite{11} leads to a quadratic equation for the momentum $z$-component. The neutrino-momentum vector is given by (cf. \cite{107})

$$\vec{p}_\nu = \left( \begin{array}{c} \frac{1}{E_{\text{lepton}} - p_{z, \text{lepton}}} (a p_{z, \text{lepton}} \pm E_{\text{lepton}} \sqrt{a^2 - E_{\text{lepton}}^2 (E_{T, \text{miss}})^2 + p_{z, \text{lepton}}^2 (E_{T, \text{miss}})^2}) \\
\end{array} \right),$$

with $a = \frac{M^2_W}{2} + E_{T, \text{miss}} p_{T, \text{lepton}} \cos (\phi_{p_{T, \text{lepton}}} - \phi_{p_{T, \text{miss}}})$, where $p_{T, \text{lepton}}$ refers to the transverse momentum of the charged lepton, and $p_{z, \text{lepton}}$ refers to the longitudinal momentum of the charged lepton.

Two classes of solutions for the longitudinal momentum of the neutrino exist. Either one complex solution in case the radicand becomes negative or two real solutions if the radicand becomes positive. After the full event selection in the muon “2 jets, 1 btag” category, but without the transverse-W-boson-mass criterion, $(77.50 \pm 0.15)\%$ of all signal events have real solutions and $(22.50 \pm 0.15)\%$ have complex solutions.

If two real solutions are obtained, the solution with the smaller absolute longitudinal momentum component $\min(|p_{z, \nu}|)$ is picked. A correct choice is obtained in $(63.36 \pm 0.20)\%$ of all events with real solutions. Here, the “correct choice” is defined as the reconstruction hypothesis that is closest in $\Delta R$ to the neutrino four-momentum as obtained from simulation truth.

Complex solutions arise if the radicand in eq. [4.4.2] becomes negative. They are obtained due to mismeasurements of the transverse-momentum-imbalance components $p_x$ and $p_y$, an underestimation of W-boson mass due to its final width, or due to a mismeasurement of the muon momentum. Simply omitting the square-root term if the radicand becomes negative, i.e. setting the radicand to zero, results in unphysically high reconstructed transverse-W-boson-mass values of $M_{T, W, \text{boson}} > 80.4 \text{ GeV/c}^2$.

In order to solve complex solution for the $z$-component of the neutrino momentum, a fit algorithm \cite{221} is used. This algorithm individually modifies the transverse momentum components $p_x$ and $p_y$. The parameters $p_x$ and $p_y$ are iteratively changed as long as the radicand is negative, i.e. as long as $M_{T, W, \text{boson}} > 80.4 \text{ GeV/c}^2$. The fit minimizes the distance between the modified transverse momentum of the neutrino $\vec{p}_{T, \nu}$ and the measured momentum imbalance $\vec{p}_{T, \text{miss}}$. Hence, the algorithm changes both the direction and magnitude of the transverse-momentum-imbalance vector $\vec{p}_{T, \text{miss}}$. Furthermore, scaling both magnitude and direction of the momentum-imbalance vector than solely its magnitude results in a more precise neutrino reconstruction \cite{107}. Although the algorithm aims for a reconstruction of the $z$-component of the neutrino momentum, it also improves the neutrino reconstruction in the transverse plane and the momentum direction of the reconstructed top quark \cite{107}.

Figure[4.6] shows the difference between the generated neutrino on particle level, i.e. from matrix element calculation, and the reconstructed neutrino at detector level. The distributions are obtained for simulated $t$-channel signal events with a muon final state in the “2 jets, 1 btag” category. Shown are the $\Delta \phi$, $\Delta \eta$, and $\Delta R$ distributions. A bias free reconstruction is obtained.

The mean of each distribution is referred to as $\mu$. The standard deviation $\sigma$ is calculated as the square root of the mean squared deviation,

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2},$$

in which $\bar{x}$ refers to the average $x$. 

102
Figure 4.6: Difference between generated neutrino and reconstructed neutrino at detector level for simulated $t$-channel signal events with a muon final state in the “2 jets, 1 btag” category. Shown are the $\Delta \phi$ (top left), $\Delta \eta$ (top right), and $\Delta R$ (bottom) distributions.
Figure 4.7 shows the difference between the generated and reconstructed W boson in $\Delta \phi$, $\Delta \eta$, and $\Delta R$ distributions. A bias-free reconstruction is obtained.

On generator (or parton) level, transverse-W-boson-mass values larger than 80.4 GeV/$c^2$ occur due to the finite width of the W boson. The width $\Gamma_{W\text{ boson}}$ of the W boson corresponds to the width parameter of a Breit-Wigner distribution. For simulated $t$-channel events, it is set to $\Gamma_{W\text{ boson}} = 2.124$ GeV/$c^2$, and the W-boson mass is set to $m_{W} = 80.425$ GeV/$c^2$. Complex solutions arise in $(9.37 \pm 0.11)\%$ of the simulated events even on generator level, when using the complete four-momentum of the generated neutrino, due to the finite width of the W boson.

Distributions for the mass and transverse mass of the W boson are shown in fig. 4.8. The reconstruction of the transverse-W-boson mass has a small bias of 2 GeV/$c^2$. Before applying the neutrino reconstruction, $(22.50 \pm 0.15)\%$ of the signal events have $M_{T, W\text{ boson}} > 80.4$ GeV/$c^2$ due to the mismeasured transverse-momentum-imbalance components. When applying the neutrino-reconstruction algorithm, finite W-boson mass effects are neglected and the momentum-imbalance vector is modified such that $M_{T, W\text{ boson}} \leq 80.4$ GeV/$c^2$.

Figure 4.7: Difference between generated and reconstructed W-boson distributions for simulated $t$-channel signal events with a muon final state in the “2 jets, 1 btag” category. Shown are the $\Delta \phi$ (top left), $\Delta \eta$ (top right), and $\Delta R$ (bottom) distributions.
4.4 Single-Top-Quark Event Reconstruction

Figure 4.8: Difference between generated and reconstructed W-boson mass (top left) and transverse mass (top right) distributions. Comparison of transverse-W-boson-mass distributions (bottom) on parton level with reconstruction level before and after applying the neutrino-reconstruction algorithm. All events are simulated \( t \)-channel signal events with a muon final state in the “2 jets, 1 btag” category, but without the transverse-W-boson-mass criterion.
4.4.3 Top-Quark Reconstruction

The reconstructed top quark is obtained by adding the four-momenta of the W-boson and the b-tagged jet hypothesis

$$\vec{P}_{\text{top quark}} = \vec{P}_{\text{W boson}} + \vec{P}_{\text{b-tagged jet}}.$$  \hspace{1cm} (4.9)

Figure 4.9 shows the difference between the generated top quark for simulated $t$-channel signal events with a muon final state in the “2 jets, 1 btag” category. Shown are the $\Delta \phi$, $\Delta \eta$, and $\Delta R$ distributions. A bias-free top-quark reconstruction is obtained.

Distributions for the differences between generated and reconstructed top-quark $p_T$ and mass are shown in fig. 4.10. The reconstructed top-quark-mass distribution is bias-free up to 1 GeV/c^2. It has a modest tail to higher reconstructed top-quark masses. The center of the distribution has a standard deviation of $\approx 22$ GeV/c^2. The corresponding relative mass resolutions are $\approx 20\%$ for the all reconstructed events and $\approx 13\%$ for events in the center of the distribution. The top-quark $p_T$ is on average 2.85 GeV/c lower on reconstruction level than on generator level.
4.5 Boosted Decision Trees

Boosted Decision Trees (BDTs) are a machine-learning method that are used in data analysis for classification or regression tasks. They belong to the class of non-linear classifiers. BDTs exploit the available phase space by simultaneously using multiple variables, and they combine them into one powerful discriminator. In high-energy physics, variables typically involve kinematics or properties of final state and composite objects, as well as correlations among them, and descriptions of the event topology.

A typical task in the context of high-energy physics, in which BDTs are applied in terms of regression, is to determine a reconstructed object momentum. As an example, an electron momentum could be determined from deposits in nearby calorimeter cells. The usage of a multivariate-regression technique may translate into an improved momentum resolution.

More likely, BDTs are applied as multivariate-analysis techniques for classification tasks. They may be used in the context of object identification, e.g. to identify prompt electrons from a list of candidates that also contains electrons from photon conversions and electrons from subsequent decays of hadrons from fragmenting b quarks. They may be also used in the context of object reconstruction, e.g. to discriminate physics signal from instrumental backgrounds.

In the analysis at hand, BDTs are used for a binary classification of events from physics processes into “signal” and “background” categories. They are used to separate $t$-channel signal events from events originating from background processes like $t\bar{t}$, $W$ boson plus jets, or QCD multijet. In the following, BDTs will be used in terms of classification tasks.

In the following subsection, the main principle of “Decision Trees” and “Boosting” are discussed. Furthermore, their configuration, which is tailored to the particular needs of this analysis, is presented. Simulated events are split into statistically independent samples, which are used for training, testing, and evaluation of the BDT classifier. The splitting of simulated events aims to avoid any bias in the training procedure that may result in a bias on the final result. Then, the input variables for the classifier trainings and their discrimination power between signal and background events are evaluated. Finally, the resulting BDT discriminant distributions and their separation power are addressed.
4.5.1 The Main Principle

Figure 4.11: Exemplary scheme of a Decision Tree. Each event passes or fails a sequence of selection criteria until it reaches a background- or signal-enriched leaf with a certain stopping criterion.

**Decision Trees**  A comprehensive description of Decision Trees is given e.g. in ref. [207] and summarized in the following paragraphs.

A Decision Tree (cf. fig. 4.11) is a binary tree in which each node represents a splitting criterion w.r.t. a set of input variables. Each branch represents a test of a particular hypothesis and is defined by a sequence of nodes. Each leaf represents the test result of a particular branch or hypothesis. In case of a categorization task, each leaf represents one of the existing categories, e.g. “signal-like events” or “background-like events”. In case of a regression task, each leaf represents a target value.

Decision Trees classify events according to learned signatures w.r.t. certain sets of input data. “Training” a single Decision Tree means to assign a splitting criterion to each node in the tree. The training initially starts with the root node, and splitting of nodes successively continues until a stopping criterion is reached. The assignment of a splitting criterion to each node, i.e. picking a particular variable from the set of available input variables and finding a selection value, is done in a way that it optimizes the increase in a particular quality criterion. The quality criterion is defined as

$$\max(N_{\text{parent node}} \cdot G_{\text{parent node}} - N_{\text{left}} \cdot G_{\text{left}} - N_{\text{right}} \cdot G_{\text{right}})$$

for a binary classification with a separation index $G$, sum of event weights $N$, and daughter nodes “left” and “right” (cf. [207]).

The Gini index [207], which evolves to $G = 2 \cdot p \cdot (1 - p)$ in case of a binary splitting node, is used as separation index in this analysis. Here, $p$ is the purity of a subsample, i.e. $p = 0.5$ refers to a fully mixed sample, and $p \rightarrow 0$ for fully separated signal and background subsamples. The Gini index is symmetric around $p = 0.5$ (cf. fig. 4.12) such that the quality criterion maximizes both the purity of signal events and the impurity with background events. Other separation indices may include cross entropy $(-0.5 \cdot [p \cdot \log_2(p) + (1 - p) \cdot \log_2(1 - p)])$, misclassification
error \((1 - \max(p, 1 - p))\), or the statistical significance \((S/\sqrt{S + B})\) (cf. \cite{207}). A comparison of the performance when alternating the choice of the separation index is given in section 4.5.4.

The splitting of a node stops if a full separation, a maximal tree depth, a maximal number of nodes, or a minimal number of events is reached.

![Figure 4.12: The separation indices Gini index, misclassification error, and cross entropy are shown as a function of the purity of a node.](image)

In this analysis, based on the majority of events passing or failing a particular sequence (branch) of selection criteria, a leaf is classified as “signal” or “background” like. In a single Decision Tree, a discriminator value of \(D(\vec{x}) = +1\) is assigned to signal-like events and \(D(\vec{x}) = -1\) is assigned to background-like events. “Evaluating” an event means to calculate the final discriminator value for that particular event. The weights of an already trained classifier are used for evaluation.

The classification result of a single Decision Tree might still contain a high number of misclassified events, and it might be sensitive to statistical outliers. Its performance can be improved when Boosting is applied, which is explained in the next paragraph.

**Boosting**  
“Boosting” is a technique to combine an ensemble of weak machine-learning classifiers, e.g. a single decision tree, into one highly accurate classifier. A “weak classifier” refers to a classifier that has a substantial fraction of incorrectly classified events, but at least performs better than a random guess. In this analysis, the Adaptive Boost (AdaBoost) algorithm \cite{222} is used. Here, “adaptive” is used in the sense that the algorithm does not need any prior knowledge about the performance of a certain weak classifier \cite{222}. Instead, the algorithm uses the misclassification rate, i.e. the probability of an incorrect classification, as obtained from a previous boosting cycle. Furthermore, the total number of trained weak classifiers does not need to be known a priori.

Technically, the training of the full ensemble of classifiers initially starts with training of a first classifier, e.g. a Decision Tree, on a set of events with user-defined event weights. In subsequent training cycles, each misclassified event is reweighted by a boost weight \(\alpha\). An event is identified as “misclassified” according to the results of the previous training cycle. The procedure is repeated until in total \(N_{\text{Trees}}\) classifiers are trained. In the TMVA implementation, which is used in this analysis, one predefined the dimension of the ensemble \(N_{\text{Trees}}\) as the stopping criterion for the boosting.
The boost weight $\alpha$ is defined as

$$\alpha = \left(1 - \frac{err}{err}\right)^\beta$$

(4.10)

with misclassification rate $err$ and AdaBoost parameter $\beta$ [207]. The sum of weights in each sample is conserved by renormalizing the reweighted training sample for each training cycle.

**Final Discriminant**  The weighted majority vote of the whole ensemble of Decision Trees yields one powerful discriminator [207],

$$y_{\text{Boost}}(\vec{x}) = \frac{1}{N_{\text{Trees}}} \sum_{i}^{N_{\text{Trees}}} \ln \alpha_i D_i(\vec{x}),$$

(4.11)

in which $N_{\text{Trees}}$ Decision-Tree discriminants $D \in \{-1, +1\}$ are linearly combined. The weight of each single classifier $D_i$ depends on its individual classification performance, which is expressed in the boost weight $\alpha$. The final discriminator distribution $y_{\text{Boost}}$ has a continuous spectrum with values $y_{\text{Boost}} \in \{-1, +1\}$, in which signal-like events are shifted to higher values, and background-like events to lower values.

In general, BDTs are well out-of-the-box performing classifiers in a sense that they neither require any preprocessing of input variables nor complex tweaking of parameters. They are also able to ignore non-separating variables which makes their classification performance quite robust. A potential caveat of the AdaBoost algorithm is that its performance might be sensitive to uniform noise [223] or outliers [207]. However, none of that is observed during the analysis at hand.

**4.5.2 TMVA Setup And Simulated-Sample Splitting**

Individual trainings are performed with TMVA for the muon “2 jets, 1 btag”, electron “2 jets, 1 btag”, muon “3 jets, 1 btag”, and electron “3 jets, 1 btag” categories. In all four trainings, events are weighted according to the SM prediction. Events are binarily classified into “signal” and “background” categories. Separate files with “training” and “testing” events are input to the TMVA algorithm. AdaBoost with a boost parameter of $\beta = 0.2$ is used as the boosting type. Pruning is not applied, since overtraining is not observed. An ensemble of 500 Decision Trees are trained in the muon “2 jets, 1 btag” category, and 400 Decision Trees are trained in all other categories. The Gini index is used as the separation criterion for node splitting. A TMVA algorithm is used to optimize the variable and selection value in each node w.r.t. the available parameter space of the training sample. The splitting of a node stops if a minimum of 150 events is reached, except for the trees built in the muon “2 jets, 1 btag” category in which the stopping criterion is 50 events. The number of nodes as well as the maximum depth of the whole Decision Tree itself are not limited.

Simulated events are randomly splitted into three orthogonal, i.e. statistically independent, samples used for training, testing, and evaluation of the BDT classifier. Each sample contains one third of the available simulated events of each signal or background process. The “training” sample is used to grow the forest of BDTs. The “testing” sample is used for parameter optimization and optimization of the list of input variables. The “evaluation” sample is used for the final statistical evaluation. Typically, more than one training cycle is done in a multivariate analysis. The splitting procedure is expected to avoid any bias on the result, e.g. a
4.5 Boosted Decision Trees

hidden optimization due to the choice of variables that are input to the classifier training, or an overtraining of the classifier. “Overtraining” refers to an effect in which a classifier memorizes certain features of individual events of the training sample such that it performs much better on the training sample than on a statistically independent sample. An overtraining of the classifier usually occurs if the training statistics is very low. Overtraining of the trained classifiers is evaluated by comparing the classifier distributions for both “training” and “testing” samples. Their compatibility is determined with a binned Kolmogorov–Smirnov (KS) test [224]. KS-test probabilities are expected to be uniformly distributed between zero and one if both distributions are compatible, and zero if both distributions are incompatible. This test is individually performed for both signal events and the sum of background events. KS-test probabilities well above zero are obtained for all classifier distributions, i.e. all trained classifier distributions pass the KS test. In conclusion, an overtraining is not observed for classifiers which are trained for the analysis at hand.

4.5.3 Discriminating Variables

Variables describing basic-object kinematics, composite-object kinematics, as well as angular correlations and event topology are considered as input variables for the BDT-classifier training. 37 variables have been used in a previous iterations of this analysis [21, 107], which themselves are partially inspired by the choice in [225, 226]. In this analysis, the set of variables is further reduced to 11 well discriminating variables. This subset is chosen such that the performance of the overall signal-to-background separation of the BDT classifier remains nearly constant while the analysis itself becomes less complex.

This subsection is organized as follows. In the first paragraph, a short introduction into an indicator of signal-to-background-separation power is given. After that, all 11 input variables which are used for the BDT training are discussed. Finally, all input variables and their separation power are summarized in table 4.6.

Definition of separation power The Receiver operating characteristic (ROC) curve (cf. fig. 4.13) illustrates the performance of a binary decision criterion at various thresholds of a particular variable. In the context of signal-to-background separation, the ROC is evaluated in terms of “signal efficiency” vs. “background rejection”. The binary decision criteria are “value is larger” or “value is less” a certain threshold.

The area under the receiver-operating-characteristic curve (AUC) serves as a key indicator that represents the “separation power” of a particular variable or classifier output. The values of the AUC lie within the interval [0.5, 1.0]. A value of 0.5, equivalent to the bisecting line in the signal-efficiency vs. background-rejection plot (cf. fig. 4.13), means no separation power at all, and AUC → 1.0 for well-separating variables.

The following should be kept in mind when using the AUC as a performance indicator. The AUC is insensitive to symmetric variables. However, BDTs is sensitive to symmetric variables during tree building. In order to calculate meaningful AUC values for the (symmetric) pseudo-rapidity distributions, the absolute distributions |η| are taken (cf. table 4.6). In the following figures, the AUC are always calculated w.r.t. the shown distribution, i.e. no additional transformations are applied. The AUC is calculated w.r.t. the sum of (weighted) background contributions. This makes it in particular insensitive to variables in which the dominant background contributions, t ¯ t and W-boson-plus-jets events, have shapes that “envelope” the signal contribution. As an example, the sum of the transverse energies of all jets (fig. 4.19) appears
to be a less powerful variable with an AUC of 50 – 55%. However, $t\bar{t}$ events have a harder spectrum than $t$-channel-signal events, and an even much harder spectrum than W-boson-plus-jets events. A variable with an overall low performing AUC might still separate very well in tree nodes in which the background contribution is dominated by one particular process.

**Discussion of input variables** The distinctive feature of $t$-channel production, the pseudo-rapidity of the light-jet hypothesis, is one of the most discriminating variables (fig. 4.14). Signal events have a jet that can be close to the beamline and that most probably has a pseudorapidity of $|\eta| \approx 2.5$. Jets from background events mostly are in the central part of the detector. Their $\eta$ distribution peaks at $\eta \approx 0$.

In $t$-channel-signal events, the light-jet candidate often also carries a large amount of transverse momentum, because it balances the heavy top quark. The jet originating from the $b$ quark of the top-quark decay usually has a huge amount of $p_T$ due to the large top-quark mass, too. Thus, the leading jet is still often close to the beamline as well (fig. 4.15). In combination with the light-jet $\eta$ distribution, the leading-jet $\eta$ adds valuable information about the $p_T$ ordering of all jets in an event.
4.5 Boosted Decision Trees

Figure 4.14: Shape comparison of the pseudo-rapidity $\eta$ distribution of the light jet in the electron (left) or muon (right) "2 jets, 1 btag" signal category. Simulated background contributions are normalized to the SM prediction, and signal events are normalized to the same area.

Figure 4.15: Shape comparison of the leading-jet $\eta$ distribution in the "2 jets, 1 btag" signal category for events with electron (left) or muon (right) final states. Simulated background contributions are normalized to the SM prediction, and signal events are normalized to the same area.
Another input variable is the cosine of the angle between the reconstructed $W$ boson and $W$-boson-plus-leading-jet system (cf. [227]),

$$
\cos^* (W \text{ boson, leading jet}) \equiv \cos \left( \frac{(P_W + P_{jet1})}{(\vec{p}_W + \vec{p}_{jet1})_{lab}} \right),
$$

(4.12)

Here, $jet1$ refers to the leading jet, and $W$ refers to the $W$ boson. $\vec{p}_i$ is the momentum vector of particle $i$. $P_i$ is the four-momentum of particle $i$. $lab$ refers to the laboratory frame, and $P_W + P_{jet1}$ is the rest frame of $W$-boson plus leading jet. If both particles are back-to-back, and $|\vec{p}_W| > |\vec{p}_{jet1}|$ then $\cos^* \to 1$, while $\cos^* \to -1$ for $|\vec{p}_W| < |\vec{p}_{jet1}|$. If the $W$-boson movement ($\vec{p}_W$), in the center-of-mass system of $W$ boson and leading jet, is orthogonal to the movement of the center-of-mass system in the lab frame, then $\cos^* \to 0$ [227]. Thus, this variable is sensitive not only to the relative directions of $W$ boson and leading jet, but also to their absolute-momentum ordering.

The angle between the between the reconstructed $W$ boson and $W$-boson-plus-leading-jet system is a kind of variable that describes the event topology. It well separates between signal and background processes. Figure 4.16 shows the resulting distribution for simulated events. The distribution shows a distinctive peak at $\cos^* \to -1$ for signal events and a much more even distribution for background events. About 33% of all $t$-channel events peak at $\cos^* \approx -1$ with a steeply falling spectrum to 0 and nearly constant spectrum between 0 and 1. The spectrum for $W$-boson-plus-jets events is much more smooth, with about 15% of all events peaking at $\cos^* \approx -1$, while the spectrum is nearly flat for $t\bar{t}$ events. The angular correlation mostly vanishes for $t\bar{t}$ events since two high-$p_T$ jets arise from the $b$ quark in the top-quark decay, as well as due to combinatorics of the two top quarks (and two $W$ bosons). A small “bias” to low values occurs due to the jet $p_T$ ordering used in the definition of $\cos^*$. 

Figure 4.16: Shape comparison of the cosine of the angle between the reconstructed $W$ boson and $W$-boson-plus-leading-jet system. Shown are events with electron final states (left) or muon final states (right) in the “2 jets, 1 btag” signal category.
Another variable that describes the topology of an event is the sphericity $S$. The following definition is given in ref. [226]. The sphericity tensor $S^{\alpha \beta}$ ($3 \times 3$) is defined as

$$S^{\alpha \beta} = \sum_{\text{jets, } l^\pm} \frac{p_\alpha^i p_\beta^i}{|\vec{p}_i|^2}, \quad \alpha, \beta = x, y, z,$$

(4.13)

in which the sums take into account all reconstructed (accepted) jets in an event and the charged lepton. $p_\alpha^i$ refers to the $\alpha$-component of the momentum vector of particle $i$. The normalized eigenvalues of the sphericity tensor, $\lambda_1$, $\lambda_2$, and $\lambda_3$, are calculated and sorted in descending order

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \quad \text{with } \lambda_1 + \lambda_2 + \lambda_3 = 1. \quad (4.14)$$

The sphericity $S$ is a linear combination of the eigenvalues $\lambda_2$ and $\lambda_3$ and is calculated as

$$S = \frac{3}{2} (\lambda_2 + \lambda_3) \quad \text{with } 0 \leq S \leq 1. \quad (4.15)$$

The expected sphericity distribution for simulated events at a center-of-mass energy of $\sqrt{s} = 7$ TeV is shown in fig. 4.17. The energy of $t$-channel events is mostly clustered in one direction, they are highly spherical. In background processes like $t\bar{t}$ production, the energy flow is more spherically and more regularly distributed in all three space dimensions than in $t$-channel events.

Figure 4.17: Shape comparison of the sphericity distributions in the “2 jets, 1 btag” signal category for events with electron (left) or muon (right) final states. Simulated background contributions are normalized to the SM prediction, and signal events are normalized to the same area.
The next category of input variables consists of variables that are related to the hadronic activity in the event. These are the **sum of energies of all jets**,

\[ H(\text{jets}) = \sum_{\text{jets}} E_i, \]

the **sum of transverse energies of all jets**

\[ H_{\text{T}}(\text{jets}) = \sum_{\text{jets}} E_{\text{T},i}, \]

the **angular separation between the two leading \( p_T \) jets**, 

\[ \Delta R(\text{jet}1, \text{jet}2) = \sqrt{(\eta_{\text{jet}1} - \eta_{\text{jet}2})^2 + (\phi_{\text{jet}1} - \phi_{\text{jet}2})^2}, \]

and the **mass of the Hadronic final state (HFS)** i.e. the mass of the composite \( N \)-jets system,

\[ \text{mass}(\text{HFS}) = \text{mass}(\sum_{\text{jets}} P_i). \]

The mass of the HFS is also referred to as “dijet mass” in case of categories with exactly two jets.

Since the light jet is close to the beamline in \( t \)-channel events, the separation power of the variables related to the hadronic activity is enhanced. Jets in signal events tend to have higher energies for signal events as described by the \( H(\text{jets}) \) distribution in fig. 4.18.

Furthermore, higher dijet masses (\( \text{mass}(\text{HFS}) \)) are generated in \( t \)-channel events than in background events (cf. fig. 4.21). The dijet mass is highly correlated to the other variables describing the hadronic activity, e.g. to the angular separation \( \Delta R \). These correlations are taken into account by the BDT training. Additional separation power is gained from these correlations, as long as those correlations are different between signal and background processes. This is the case for variables which are related to the hadronic activity, since their separation power is driven by the forward jet in \( t \)-channel events.

The angular separation between both jets (\( \Delta R(\text{jet}1, \text{jet}2) \)) is much broader for signal events than for background events (cf. fig. 4.20). The distribution peaks for both signal and background events at \( \Delta R \approx 3 \); the cut-off at 0.5 is due to the jet-clustering parameter. W-boson-plus-jets events have a smaller second peak at \( \Delta R \approx 0.8 \). These jets are expected to be two narrow jets originating from a radiated gluon.

Figure 4.19 refers to the \( H_{\text{T}}(\text{jets}) \) distribution, which is –at a first glance– not very well separating. However, jets in \( t\bar{t} \) events usually have more transverse energy than jets in \( t \)-channel events, and even much more transverse energy than jets in W-boson-plus-jets events. Thus, only the weighted background contribution is balanced against signal events, while individual processes can be separated from each other.
4.5 Boosted Decision Trees

Figure 4.18: Shape comparison of the sum-of-energies distributions of all jets for events in the electron (left) or muon (right) “2 jets, 1 btag” signal category.

Figure 4.19: Shape comparison of the sum-of-transverse-energies distributions for events in the electron (left) or muon (right) “2 jets, 1 btag” signal category.

Figure 4.20: Shape comparison of the angular separation $\Delta R$ between the two leading jets for events in the electron (left) or muon (right) “2 jets, 1 btag” signal category. Simulated background contributions are normalized to the SM prediction, and signal events are normalized to the same area.
Figure 4.21: Shape comparison of the dijet-mass distribution for events in the electron (left) or muon (right) “2 jets, 1 btag” signal category. Simulated background contributions are normalized to the SM prediction, and signal events are normalized to the same area.

Figure 4.22: Shape comparison of the lepton-$p_T$ distribution between signal and background events in the electron (left) or muon (right) “2 jets, 1 btag” signal category. Simulated background contributions are normalized to the SM prediction, and signal events are normalized to the same area.

The transverse momentum of the lepton (fig. 4.22) is much softer in $t$-channel events than in W-boson-plus-jets or $t\overline{t}$ events. Leptons from $t\overline{t}$ processes have the hardest transverse momentum spectrum, they are highly boosted in the transverse plane.

The reconstructed b-tagged-top-quark mass (fig. 4.23) peaks for both $t$-channel and $t\overline{t}$ events at $\approx 170\text{GeV}/c^2$, while it peaks much lower at $\approx 140\text{GeV}/c^2$ for W-boson-plus-jets events. The mass distribution of the reconstructed top-quark candidate has a broad tail to high reconstructed masses for $t\overline{t}$ and W-boson-plus-jets events. For $t$-channel events, the distribution is much more narrow, as mostly a correct combination of W boson and b-tagged-jet hypothesis is picked in these events.

The best-top-quark-mass (fig. 4.24) distribution is biased, by definition, to be close to the input value of $172.0\text{GeV}/c^2$ for all processes. The obtained distribution is more narrow than the mass distribution of the b-tagged-top-quark candidate. This best-top-quark candidate appar-
4.5 Boosted Decision Trees

Figure 4.23: Shape comparison of the reconstructed b-tagged-top-quark-mass distributions for events in the electron (left) or muon (right) “2 jets, 1 btag” signal category. Simulated background contributions are normalized to the SM prediction, and signal events are normalized to the same area.

Figure 4.24: Shape comparison of the reconstructed best-top-quark mass distributions, i.e. the mass reconstructed with the jet that yields a mass closest to $172.0\, \text{GeV}/c^2$, for events in the electron (left) or muon (right) “2 jets, 1 btag” signal category. Simulated background contributions are normalized to the SM prediction, and signal events are normalized to the same area.
Correlations among input variables  Figure 4.25 shows the linear correlation coefficients among all BDT input variables. Most variables are moderately correlated or even uncorrelated among each other. Meaningful correlations exist between the reconstructed top-quark-mass hypotheses and among the jet-separation variables ($\sum E_T$, $\sum E$, $\Delta R$). Furthermore, the sphericity is (anti)-correlated to the dijet mass, angular separation between both jets and sum of energies of both jets, and $\cos^* (W \text{ boson, leading jet})$. $\cos^* (W \text{ boson, leading jet})$ is also correlated to the dijet mass and sum of energies of both jets. Light-jet $\eta$ and leading-jet $\eta$ are correlated as well.

Linear correlations among variables for background events are in the same ballpark as for signal events, with a modestly diversified pattern, which partially can be explained due to the more complex alternation of jet hypotheses for background events.

Figure 4.25: Linear-correlations coefficients for BDT input variables for signal (left) and background events (right). The coefficients are exemplarily shown for events in the muon “2 jets, 1 btag” category.
4.5 Boosted Decision Trees

Summary of discussion of input variables  All 11 input variables and their separation power are summarized in table 4.6. The most discriminating variables, according to the measure in units of [AUC], are the pseudo-rapidity of the jet that is closest to the beamline in an event, the cosine of the angle between the reconstructed W boson and the W-boson-plus-leading-jet system \((\cos (W, \text{leading jet}))\), and the sphericity of the event. For both signal and background events, most of the used input variables are linearly uncorrelated among each other, or they have relatively low correlation coefficients. Input variables that are related to the hadronic activity in an event are correlated among each other, but they add significant discrimination power to the BDT training even due to these correlations.

<table>
<thead>
<tr>
<th>Input variable</th>
<th>Performance [AUC in %]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\eta) of the light jet</td>
<td>(e) 73.0 (\mu) 73.3</td>
</tr>
<tr>
<td>Cosine of the angle between the rec. W boson and W-boson-plus-leading-jet system</td>
<td>(e) 73.2 (\mu) 69.8</td>
</tr>
<tr>
<td>Sphericity</td>
<td>(e) 69.4 (\mu) 69.9</td>
</tr>
<tr>
<td>Sum of the energies of all jets</td>
<td>(e) 67.9 (\mu) 70.7</td>
</tr>
<tr>
<td>Dijet mass of the b-tagged-jet plus light-jet candidates</td>
<td>(e) 65.3 (\mu) 67.9</td>
</tr>
<tr>
<td>Angular separation (\Delta R) between leading two jets</td>
<td>(e) 66.6 (\mu) 66.8</td>
</tr>
<tr>
<td>(\eta) of the leading jet</td>
<td>(e) 60.2 (\mu) 59.1</td>
</tr>
<tr>
<td>Mass of the b-tagged-top-quark candidate</td>
<td>(e) 64.8 (\mu) 56.4</td>
</tr>
<tr>
<td>Lepton (p_T)</td>
<td>(e) 61.7 (\mu) 56.9</td>
</tr>
<tr>
<td>Sum of the transverse energies of all jets</td>
<td>(e) 55.0 (\mu) 50.8</td>
</tr>
<tr>
<td>Mass of the best-top-quark candidate</td>
<td>(e) 55.4 (\mu) 51.5</td>
</tr>
</tbody>
</table>

Table 4.6: Input variables used for the Boosted-Decision-Tree trainings. Separate Boosted Decision Trees are trained for the electron “2 jets, 1 btag”, muon “2 jets, 1 btag”, electron “2 jets, 1 btag”, and electron “3 jets, 1 btag” categories.

4.5.4 Boosted-Decision-Tree Classifier Distribution

Independent BDT trainings are performed with events in the electron “2 jets, 1 btag”, muon “2 jets, 1 btag”, electron “3 jets, 1 btag”, and muon “3 jets, 1 btag” categories. The “2 jets, 1 btag” trainings are used for events in categories with exactly two jets, and the “3 jets, 1 btag” trainings are used for categories with at least three jets. Shape comparisons of the resulting BDT classifier distributions for signal and background events are shown for trainings in the two-jet categories in figure 4.26 and for trainings in the three-jet categories in figure 4.27. The mean of both distributions significantly differ between the two event classes. As different trainings are performed, the shape of the classifier distribution is expected to vary among the categories due to the altered event kinematics, varied jet-hypotheses assignments, and diversified patterns of background contributions. The shape of the classifier distribution is also influenced by the BDT training parameters, i.e. the relative normalization of signal and background events. Signal and background events are normalized according to the SM prediction. Performance values, as measured with the [AUC] for the four trained BDT classifier in the signal categories are shown in table 4.7. A performance of above 80% [AUC] is reached in every category.
Figure 4.26: Shape comparison of the BDT-classifier distributions in the “2 jets, 1 btag” signal category for events with electron (left) or muon (right) final states. Simulated background contributions are normalized to the SM prediction, and signal events are normalized to the same area.

Figure 4.27: Shape comparison of the BDT-classifier distributions in the “3 jets, 1 btag” signal category for events with electron (left) or muon (right) final states. Simulated background contributions are normalized to the SM prediction, and signal events are normalized to the same area.

### Category Performance [AUC in %]

<table>
<thead>
<tr>
<th>Category</th>
<th>Performance [AUC in %]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e$, “2 jets, 1 btag”</td>
<td>82.8</td>
</tr>
<tr>
<td>$\mu$, “2 jets, 1 btag”</td>
<td>81.5</td>
</tr>
<tr>
<td>$e$, “3 jets, 1 btag”</td>
<td>81.7</td>
</tr>
<tr>
<td>$\mu$, “3 jets, 1 btag”</td>
<td>80.5</td>
</tr>
</tbody>
</table>

Table 4.7: Performance of the trained BDTs in the electron “2 jets, 1 btag”, muon “2 jets, 1 btag”, electron “2 jets, 1 btag”, and electron “3 jets, 1 btag” categories.
As an additional test of the stability of the classifier trainings, the separation index is changed from the “Gini index”, which is used as the default, to “cross entropy”, “misclassification error”, or “statistical significance” \((S/\sqrt{S + B})\). A comparison of all distributions of these separation indices is given in figure 4.12. Table 4.8 shows the resulting variations of the separation power (measured in [AUC]) for all BDT trainings. As nominal performance, the separation power as achieved with the Gini index is used. The change of the separation index results in insignificant changes to the BDT classifier performance, i.e. the performance of the BDT classifier remains robust against the alternation of the separation index.

<table>
<thead>
<tr>
<th>Separation index</th>
<th>“2 jets, 1 btag”</th>
<th>“3 jets, 1 btag”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gini index (default)</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Cross Entropy</td>
<td>+0.1%</td>
<td>−0.1%</td>
</tr>
<tr>
<td>Misclassification error</td>
<td>0.0%</td>
<td>+0.8%</td>
</tr>
<tr>
<td>(S/\sqrt{S + B})</td>
<td>−0.2%</td>
<td>−0.3%</td>
</tr>
</tbody>
</table>

Table 4.8: Relative changes in the separation power as measured in the area under the receiver-operating-characteristic curve (AUC) for various separation indices. Shifts are shown w.r.t. the separation as achieved with the Gini index, which is used as the default. Trainings in the electron “2 jets, 1 btag”, muon “2 jets, 1 btag”, electron “2 jets, 1 btag”, and electron “3 jets, 1 btag” categories are compared.
4.6 Statistical Inference

This section describes how conclusions about a certain parameter of interest can be drawn from measured data. A Bayesian approach is used as the core method for the statistical inference. The posterior-probability distribution for measured data contains the desired information about the parameter of interest. This information is summarized with two quantities. First, the median of the posterior-probability distribution is used as the estimate of the parameter of interest. Second, the central-68% interval of the posterior-probability distribution is considered as the uncertainty of this estimate. An interval estimated from the posterior-probability distribution is also specifically referred to as “Bayesian confidence interval” or “credible interval” in literature. In the context of this analysis, only Bayesian statistics are used for statistical inference and intervals estimated from posterior-probability distributions simply are referred to as “confidence intervals”.

The Bayesian inference requires two ingredients. First, Bayesian inference requires to formulate the prior-probability distributions for all parameters the posterior probability depends on. The prior-probability distribution of a particular parameter represents the knowledge about that parameter before the measurement is carried out (cf. [11, 209]). Second, Bayesian inference requires to specify the likelihood function, which is characteristic to the particular experiment or analysis.

Systematic uncertainties are included with additional nuisance parameters into the likelihood function. The nuisance parameters are eliminated by marginalization, i.e. the likelihood function is integrated over all nuisance parameters. This integration is numerically performed with a Markov-chain-Monte-Carlo (MCMC) method.

This section starts with an introduction to the Bayesian evaluation in section 4.6.1. The posterior-probability distribution is derived step-by-step from first principles. Then, the parameter of interest and the prior-probability distributions are defined. Finally, the construction of the likelihood function is discussed in section 4.6.2.

4.6.1 Bayesian Evaluation

The Bayesian approach to interpret measured data in terms of one or more parameters of interest is well described in the literature. In particular, this section refers to the descriptions given in ref. [11, 205, 208, 209, 228, 229].

**General formulation of Bayes’ theorem** Bayes’ theorem relates the conditional probability $P(A|B)$, i.e. the probability to observe $A$ given $B$,

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)},$$

(4.16)

with the conditional probability $P(B|A)$, i.e. the probability to observe $B$ given $A$, and the “degree of belief” in $A$, which is referred to as $P(A)$, and degree of belief in $B$, which is referred to as $P(B)$ [209]. $B$ refers to measured data and is fixed a priori with constant $P(B) \neq 0$. For a total number of events $S$, which are divided into $A_1, \ldots, A_n$ exclusive sets, and in which $B$ is any event or subset of $S$, the probability to observe any set $A_x$ given $B$ becomes [209, 228]

$$P(A_x|B) = \frac{P(B|A_x) \times P(A_x)}{\sum_i P(B|A_i) \times P(A_i)},$$

(4.17)
**Parametric, continuous distributions** In the following, Bayes’ theorem is applied to parametric, continuous distributions. The probability $P(A_x|B)$ is identified with the posterior-probability-density function $P(\mu|x)$ for the parameter of interest $\mu$ w.r.t. measured data $x$, and is given by \[ P(\mu|x) = \frac{P(x|\mu) \times \pi(\mu)}{\int P(x|\mu) \times \pi(\mu) \, d\mu}. \] \tag{4.18}

In particular, the outcome of the experiment with measured data $x$ depends on the parameter $\mu$, which is unknown a priori.

$x$ can be either a single measurement or a set of data points. The parameter $\mu$ is referred to as the “parameter of interest”, i.e. the parameter that is about to be estimated. In principle, $\mu$ can be a vector of parameters, however, only one parameter of interest is necessary for this analysis, and $\mu$ is of dimension one without loss of generality.

In case of continuous distributions, the addition in eq. 4.17 becomes an integral over the parameter of interest $\mu$. The probability $P(B|A_x)$ becomes a probability-density function $P(x|\mu)$ to obtain a certain measurement $x$ for a given parameter of interest $\mu$ (cf. [11, 209]). In particular, $P(x|\mu)$ encodes the outcome of the experiment or analysis under a set of known parameters.

For a fixed set of data points $x$, $P(x|\mu)$ becomes the “likelihood function” $L(\mu|x)$, which is no longer a probability-density function (cf. [209])

\[ L(\mu|x) = P(x|\mu). \] \tag{4.19}

The likelihood function $L(\mu|x)$ is a function of the parameter $\mu$ for a fixed $x$, and is characteristic to the experiment or analysis. Thus, the posterior-probability-density function $P(\mu|x)$ is given by (cf. [209, 228])

\[ P(\mu|x) = \frac{L(\mu|x) \times \pi(\mu)}{\int L(\mu|x) \times \pi(\mu) \, d\mu}. \] \tag{4.20}

The construction of the likelihood function is discussed in detail in the next subsection (4.6.2).

$\pi(\mu)$ refers to the prior-probability distribution for the parameter of interest $\mu$, i.e. the knowledge about $\mu$ before the actual measurement is performed (cf. [11, 209]).

More generally, the Bayes theorem relates the probability of a specific theory given measured data with the prior probability about the theory and the predicted outcome of the experiment based on a specific theory [11, 209]

\[ P(\text{theory}|\text{data}) \propto P(\text{data}|\text{theory}) \times \pi(\text{theory}). \] \tag{4.21}

The statistical model of this analysis, described by the likelihood function, depends not only on the parameter of interest $\mu$, but on a number of additional nuisance parameters $\bar{\theta} = (\theta_1, ..., \theta_n)$. Then, also the prior-probability and posterior-probability distributions depend on the nuisance parameters $\bar{\theta}$, and eq. 4.20 becomes

\[ P(\mu, \bar{\theta}|x) = \frac{L(\mu, \bar{\theta}|x) \times \pi(\mu, \bar{\theta})}{\int L(\mu, \bar{\theta}|x) \times \pi(\mu, \bar{\theta}) \, d\bar{\theta} \, d\mu} \]

\[ = \frac{L(\mu, \bar{\theta}|x) \times \pi(\mu) \times \pi(\bar{\theta})}{\int L(\mu, \bar{\theta}|x) \times \pi(\mu) \times \pi(\bar{\theta}) \, d\bar{\theta} \, d\mu}. \] \tag{4.22}

The arguments of $L(\mu|x)$ are interchanged w.r.t. $P(x|\mu)$ to emphasize that $L$ is a function of $\mu$. 

---

5The arguments of $L(\mu|x)$ are interchanged w.r.t. $P(x|\mu)$ to emphasize that $L$ is a function of $\mu$. 

125
Here, the prior-probability distributions for the nuisance parameters are referred to as $\pi(\vec{\theta})$. If $\vec{\theta}$ and $\mu$ are independent of each other, which is the case in this analysis, the joint prior-probability distribution of $\vec{\theta}$ and $\mu$ factorizes

$$\pi(\mu, \vec{\theta}) = \pi(\mu) \times \pi(\vec{\theta}).$$

(4.23)

A posterior-probability distribution $p(\mu|\vec{x})$ that is independent of $\vec{\theta}$ is obtained by integrating over all nuisance parameters $\vec{\theta}$ (cf. [209])

$$P(\mu|\vec{x}) = \frac{1}{C} \times \int L(\mu, \vec{\theta}|\vec{x}) \times \pi(\mu) \times \pi(\vec{\theta}) \, d\vec{\theta} \propto L_m(\mu|\vec{x}) \times \pi(\mu).$$

(4.24)

This integration is also referred to as "marginalization". In this analysis, the integration is numerically performed by using a Metropolis-Hastings Markov-chain-Monte-Carlo (MCMC) algorithm [211, 212].

The information that is contained in the posterior-probability distribution $P(\mu|\vec{x})$ is summarized with two quantities. The median $\hat{\mu}$ of the posterior-probability distribution is used as the estimate of the parameter of interest in this analysis. The median is an unbiased estimator for the measurements that are presented in this analysis (cf. sec. 6.4) and its calculation is numerically stable. The uncertainty of $\hat{\mu}$ is estimated with the Bayesian-central-68%-confidence interval $[\mu_1, \mu_2]$, which is constructed by [230]

$$\int_{-\infty}^{\mu_1} P(\mu|\vec{x}) \, d\mu = \frac{1 - C.L.}{2} = \int_{\mu_2}^{\infty} P(\mu|\vec{x}) \, d\mu,$$

(4.26)

in which the (Bayesian) confidence level is $C.L. = 0.68$.

The parameter of interest, the prior-probability distribution, and the construction of the likelihood function will be discussed in the following paragraphs.

**Parameter of interest** In this analysis, the parameter of interest $\mu$ corresponds to the signal strength, which is defined as

$$\mu = \frac{\sigma_{\text{meas.}}^{\text{t-channel}}}{\sigma_{\text{SM}}^{\text{t-channel}}}.$$

(4.27)

Here, $\sigma_{\text{meas.}}^{\text{t-channel}}$ refers to the measured cross section, and $\sigma_{\text{SM}}^{\text{t-channel}}$ refers to the SM prediction\(^6\)

\(^6\)The SM prediction is given in section 3.5.1.
4.6 Statistical Inference

Prior-probability distribution $\pi(\mu)$  The prior-probability distribution for the parameter of interest $\pi(\mu)$ is chosen such that it is uniformly distributed (“flat”) in the parameter of interest $\mu$ within the interval $[0, \infty]$ (cf. \cite{[11]})

$$\pi(\mu) = \begin{cases} 0 & \mu < 0 \\ 1 & \mu \geq 0. \end{cases}$$

(4.28)

In particular, this prior probability is flat in terms of the $t$-channel cross section in this analysis, and, therefore, flat for the Poisson means of $t$-channel events.

From a physics point of view, one could argue that it is more “natural” to use a prior-probability distribution that is flat in the fundamental parameter $|V_{tb}|$, rather than using a prior-probability distribution that is flat in the measured cross section $\sigma_{t\text{-channel}} \propto |V_{tb}|^2$. A prior-probability distribution that is flat in $|V_{tb}|$ can be defined as

$$\pi_{\text{cross check}}(\mu) = \begin{cases} 0 & \mu < 0 \\ \frac{1}{\sqrt{\mu}} & \mu \geq 0, \end{cases}$$

(4.29)

since $\mu \propto \sigma_{t\text{-channel}} \propto |V_{tb}|^2$.

This prior probability is used as a cross check for the $|V_{tb}|$ measurement. The $|V_{tb}|$ measurement is performed twice, once with a prior-probability distribution that is flat in $|V_{tb}|^2$ (eq. 4.28) and once with a prior-probability distribution that is flat in $|V_{tb}|$ (eq. 4.29). The comparison of the obtained posterior distributions gives information about the “objectiveness” of the used prior-probability distribution, i.e. how sensitive the observed result is under variation of the prior-probability distribution for the parameter of interest (cf. \cite{[11]}). The impact on the $|V_{tb}|$ measurement due to the choice of the prior-probability distribution is discussed in the section “Results” \cite{7.2}.

The prior-probability distributions for the nuisance parameters $\pi(\theta)$ are discussed in the next section. They are directly related to the interpretation of systematic uncertainties, which are incorporated as additional nuisance parameters to the likelihood function.

4.6.2 Construction of the Likelihood Function

In the following paragraphs, the construction of the likelihood function is described. In case of a simple counting experiment, the probability to observe $x$ events with $\lambda$ expected events follows the Poisson distribution, which is given by \cite{[209]}

$$p(x|\lambda) \equiv \text{Poisson}(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}.$$  

(4.30)

The joint-probability-density function of $N$ statistically independent measurements (or data points) is given by the product of all individual probability-density functions $p(x|\lambda)$ \cite{[209]}

$$P(\vec{x}|\lambda) = \prod_{i=1}^{N} p(x_i|\lambda),$$

(4.31)

The likelihood function for a fixed set of measurements $\vec{x}$ as a function of $\lambda$ is obtained as \cite{[209]}

$$L(\lambda|\vec{x}) = P(\vec{x}|\lambda).$$

(4.32)
Analysis Strategy

In this analysis, data are measured in twelve categories simultaneously (cf. sec. 4.1). For each event, the BDT-discriminator distribution is calculated. The resulting BDT-discriminator distributions are binned into one-dimensional histograms, separately for measured data and each simulated physics processes, and individually for each of the twelve categories. For the signal-enriched categories “2 jets, 1 btag” and “3 jets, 1 btag”, 30 bins are used. The histograms have 20 bins for the remaining categories, which are signal depleted. The total number of bins is given by

\[ N_{\text{total bins}} = \prod_{i=1}^{N_{\text{categories}}} N_{\text{bins}}(i). \]  

(4.33)

The expected number of signal events is expressed as

\[ N_{\text{expected signal}} = \mu \times s, \]

in which \( s \) refers to the number of signal events as predicted by the SM, and \( \mu \) to the signal strength. The signal strength \( \mu \) is defined by

\[ \mu = \frac{\sigma_{\text{meas.}}}{\sigma_{\text{SM} \ t\text{-channel}}}, \]  

(4.34)

in which the (constant) SM prediction of the \( t \)-channel cross section is referred to as \( \sigma_{\text{SM} \ t\text{-channel}} \). The number of background events is referred to as \( b \). The total expected number of events in each bin \( j \) of category \( i \) is the sum of the number of signal and background events

\[ N_{\text{total expected}}^{i,j} = \mu \times s_{i,j} + b_{i,j}. \]  

(4.35)

The joint-likelihood function of measured data in all categories of this analysis can be described as

\[ L(\vec{x}|\mu) = \prod_{i=1}^{N_{\text{categories}}} \prod_{j=1}^{N_{\text{bins}}(i)} \text{Poisson}(x_{i,j}|\mu \times s_{i,j} + b_{i,j}), \]  

(4.36)

since the measurements in each bin are statistically independent of each other. \( x_{i,j} \) refers to the measured number of events in category \( i \) and bin \( j \).

The expected number of signal (\( s_{i,j} \)) and background (\( b_{i,j} \)) events are obtained from the simulation after the full event selection. The simulated sample of events itself is normalized to the integrated luminosity of the analyzed data according to the SM prediction before any event selection (cf. sec. 3.5). The “simulation of events” refers not only to the generation of events according to a theoretical prescription, but also to the detector modeling. In particular, the expected event yield in this analysis is obtained as the convolution of the theoretically predicted event yield with the acceptance of the CMS detector, the efficiency of physics-objects reconstruction, and the efficiency of the selection criteria as used in this particular analysis.

The simulation of events implies assumptions on the particular acceptance and efficiency functions, which usually depend on certain parameters. Typically, the uncertainty of a particular acceptance or efficiency function are expressed using a few physical quantities.

As an example, uncertainties of the jet-energy scale (cf. sec. 3.3.5) or jet-flavor-tagging efficiency (cf. sec. 3.5.4) are typical sources of “systematic uncertainties” that depend on both \( \eta \) and \( p_T \) of a reconstructed jet. Furthermore, QCD-multijet events are normalized to the data-driven-event yield. The QCD-multijet templates (\( b_{i,j} \)) are derived from data-sideband regions (cf. sec. 5.1). However, the yield estimation of QCD-multijet events and its extrapolation from...
a data-sideband region to the signal-phase space requires an inclusion of further systematic uncertainties.

The identification of sources of “systematic uncertainties”, as well their quantification, are important for the measurement, since systematic uncertainties alter the expected yield of a physics process. In particular, systematic uncertainties can introduce a bias in the measurement. Systematic uncertainties can be quantified with a particular measurement in data, estimated by simulation (e.g. by varying simulation parameters), or estimated by the experimenter (cf. sec. 3.3, 3.5, and 6).

Systematic uncertainties are included in the likelihood function by introducing additional nuisance parameters \( \vec{\theta} \). Each systematic uncertainty corresponds to a nuisance parameter \( \theta_i \). The joint-likelihood function of measured data in all categories in this analysis is given by

\[
L(\mu, \vec{\theta}|\vec{x}) = \prod_{i=1}^{N_{\text{categories}}} \prod_{j=1}^{N_{\text{bins}}} \text{Poisson}(x_{i,j}|\mu \times s_{i,j}(\vec{\theta}) + b_{i,j}(\vec{\theta})).
\]

(4.37)

In the following paragraphs, the dependence of the signal prediction \( s(\vec{\theta}) \) and background prediction \( b(\vec{\theta}) \) on the nuisance parameters \( \vec{\theta} \) are addressed. From an experimental point of view, the full set of nuisance parameters \( \vec{\theta} \) is divided into two orthogonal classes. The first class represents rate-only-changing systematic uncertainties \( (\vec{\theta}_{\text{flat}}) \), and the second class represents rate-and-shape-changing systematic uncertainties \( (\vec{\theta}_{\text{shape}}) \)

\[
\vec{\theta} = \begin{pmatrix} \vec{\theta}_{\text{flat}} \\ \vec{\theta}_{\text{shape}} \end{pmatrix}.
\]

(4.38)

Flat or “rate-only-changing” systematic uncertainties only have an effect on the normalization of a particular physics process and conserve the shape of a distribution. They have an effect on the predicted event yield that is independent of a particular category \( i \) or bin \( j \) of a distribution. The change of the predicted event yield of a particular physics process \( p \) in each bin \( j \) of the BDT discriminator distribution is relative to the nominal predicted yield. Usually, these uncertainties can be parametrized using continuous probability distributions, such as a normal, log-normal, or gamma distribution.

Rate-and-shape-changing systematic uncertainties distort the shape of a distribution and, in addition, may change the normalization of a particular process. These uncertainties have an effect on the normalization of a particular process in each category \( i \) and bin \( j \), but vary the predicted yield in each category \( i \) and bin \( j \) by a different amount.

The expected number of signal events in category \( i \) and bin \( j \) including systematic uncertainties is given by

\[
s_{i,j}(\vec{\theta}_{\text{flat}}, \vec{\theta}_{\text{shape}}) = \max \left\{ 0, N_{i,j,k=0}(\vec{\theta}_{\text{flat}}) \times \left[ 1 + \sum_{s=1}^{\dim(\vec{\theta}_{\text{shape}})} \sum_{s=1}^{\Delta s_{i,j,k=0}(\vec{\theta}_{\text{shape}})} \right] \right\}.
\]

(4.39)

The individual terms are discussed in the following paragraphs. The background prediction \( b(\vec{\theta}) \) is the sum of the event yield of all individual background processes \( (k) \) and is given by
As a side note, the sum runs over all different physics processes (cf. table 6.3), and an individual template is implemented for each background process. For the sake of convenience, all background processes are usually summarized into four groups in all figures of this thesis.

The individual terms in eq. 4.39 and 4.41 are

- \( N_{i,j,k}(\vec{\theta}_{\text{flat}}) \) is a function of all nuisance parameters, which represent flat systematic uncertainties and which have an impact on the signal process. \( N_{i,j,k}(\vec{\theta}_{\text{flat}}) \) is given by

\[
N_{i,j,k}(\vec{\theta}_{\text{flat}}) = N_{i,j,k}^{SM} \times \prod_{s=1}^{\text{dim}(\vec{\theta}_{\text{flat}})} \theta_{s,\text{flat}}^{s},
\]

in which \( N_{i,j,k}^{SM} \) refers to the nominal event yield for process \( k \) in category \( i \) and bin \( j \) as predicted by the SM. Here, “nominal” means that all nuisance parameters are set to their central or most-probable values. \( \text{dim}(\vec{\theta}_{\text{flat}}) \) is the total number of flat systematic uncertainties, which are incorporated as nuisance parameters.

- \( \text{dim}(\vec{\theta}_{\text{shape}}) \) is the number of shape-changing systematic uncertainties that are incorporated as nuisance parameters.

- \( \Delta_{i,j,k}^{s}(\theta_{s,\text{shape}}) \) is the relative change of the event yield due to the systematic uncertainty \( s \). In this analysis, only one process contributes to the signal prediction (\( k = 0 \)). The relative change of the event yield is a function of the nuisance parameter \( \theta_{s,\text{shape}} \), and given by

\[
\Delta_{i,j,k}^{s}(\theta_{s,\text{shape}}) = \frac{N_{i,j,k}^{s}(\theta_{s,\text{shape}}) - N_{i,j,k}^{SM}}{N_{i,j,k}^{SM}} \bigg|_{\text{SM-cross-section prediction}}.
\]

Technically, \( \Delta_{i,j,k}^{s}(\theta_{s,\text{shape}}) \) is constructed at the SM-cross-section prediction, i.e. at fixed \( \vec{\theta}_{\text{flat}} \). In particular, the expected event yield for a systematic variation \( s \) is given by \( N_{i,j,k}^{s}(\theta_{s,\text{shape}}) \), which is a function of the nuisance parameter \( \theta_{s,\text{shape}} \).

Systematic uncertainties are usually estimated by the experiments for two dedicated working points besides the expected values. These so-called “↑” and “↓” variations represent the change in efficiency, resolution, or acceptance of a specific parameter (or set of parameters) at 68% C.L. i.e. \( \pm 1\sigma \) in terms of Gaussian uncertainties.

Technically, the ↑ and ↓ variations of the systematic uncertainties \( s \) are independently applied to the simulation, i.e. the reconstruction of physics objects or the analysis selection criteria are varied by the corresponding efficiency, resolution, or acceptance terms. Then, the full analysis
4.6 Statistical Inference

is repeated. \( N_{i,j,k}^s(\theta_s^{\text{shape}}) \) is obtained from simulation for three working points, namely

\[
\begin{align*}
N_{i,j,k}^s(\theta_s^{\text{shape}} = \theta_s^{\text{shape}} \text{ (no variation)}), \\
N_{i,j,k}^s(\theta_s^{\text{shape}} = \theta_s^{\text{shape}} \uparrow \text{ (variation)}), \\
\text{and } N_{i,j,k}^s(\theta_s^{\text{shape}} = \theta_s^{\text{shape}} \downarrow \text{ (variation)}).
\end{align*}
\] (4.44)

However, the expected event yield \( N_{i,j,k}^s(\theta_s^{\text{shape}}) \) is required to be a continuous distribution of the nuisance parameter \( \theta_s \), such that it can be included in the likelihood function. Thus, \( N_{i,j,k}^s(\theta_s^{\text{shape}}) \) has to be interpolated within the interval \( \theta_s^{\text{shape}} \downarrow < X \leq \theta_s^{\text{shape}} \uparrow \), and values outside this interval have to be extrapolated. This procedure is also referred to as “template morphing”.

All available working points (eq. 4.44) are applied as constraints to the template-morphing algorithm. The template morphing is independently performed in each category \( i \), bin \( j \) of a distribution, as well as for each process \( k \) and systematic uncertainty \( s \). In this analysis, a “cubic-linear”-template morphing is used. In this algorithm, the interpolation is done with a cubic spline, i.e. a polynomial of order three. The extrapolation beyond the \( \pm 1\sigma \) interval is done using a linear function. In particular, the three obtained working points (eq. 4.44) lie on the interpolated curve \( N_{i,j,k}^s(\theta_s^{\text{shape}}) \). The cubic-linear-template morphing is implemented in [208]. The algorithm itself is described in more detail in ref. [205, 229].

The prior-probability distributions for nuisance parameters that correspond to rate-only-changing systematic uncertainties are modeled as log-normal distributions [209]

\[
\pi(\theta_s^{\text{flat}}) = \begin{cases} 
0 & \theta_s^{\text{flat}} < 0 \\
\frac{1}{\sqrt{2\pi}\sigma\theta_s^{\text{flat}}} \exp\left(-\frac{1}{2\sigma^2} (\ln \theta_s^{\text{flat}} - \lambda)^2\right) & \theta_s^{\text{flat}} \geq 0.
\end{cases}
\] (4.45)

Here, the parameter \( \lambda \) is chosen such that the median values of the log-normal distributions corresponds to the predicted SM cross section of the particular process (\( \lambda = 0 \)). The parameter \( \sigma \) is set to the relative amount of variation of the corresponding systematic uncertainty \( s \), i.e. the variation that corresponds to a \( 68\% \) CL when assuming a normal distribution. In order to restrict all physics processes to positive values in terms of cross sections and event yields, the prior-probability distributions are constrained to the interval \([0, \infty]\).

Log-normal distributions have the advantage that they are not truncated at zero. Instead, normal distributions become truncated at zero if \( \sigma \) is sufficiently large (cf. fig. 4.28). If truncated distributions are used as prior-probability distributions for the normalization of background processes, the event yields of background processes are systematically overestimated. Thus, a bias can be introduced into the measurement when using truncated normal distributions.

The prior-probability distributions for nuisance parameters that correspond to rate-and-shape-changing systematic uncertainties are modeled as standard normal distributions (\( \lambda = 0, \sigma = 1 \)) [209]

\[
\pi(\theta_s^{\text{shape}}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\theta_s^{\text{shape}} - \lambda)^2}{2\sigma^2}\right].
\] (4.46)

\( \theta_s^{\text{shape}}(\pm\sigma) \) corresponds to the \( \uparrow \) variation of the systematic uncertainty \( s \). \( \theta_s^{\text{shape}}(-\sigma) \) corresponds to the \( \downarrow \) variation of the systematic uncertainty \( s \).
Figure 4.28: Comparison of normal and log-normal distributions. Distributions are shown for $\sigma = 0.3$ and $\sigma = 1.0$. The normal distribution is truncated at zero, while the log-normal distribution is not truncated. All prior-probability distributions that correspond to rate-only-changing systematic uncertainties have median values of 1, which corresponds to the SM-cross-section prediction.

Additional nuisance parameters are introduced due to the limited statistical precision of the simulated templates, and the template for QCD-multijet events. This uncertainty, as well as its implementation in the likelihood function, is discussed in sec. 6.1.1 and not explicitly included in the likelihood function as discussed in these paragraphs.
5 Background Estimation and Evaluation of Simulated Event Modeling

The first section of this chapter (sec. 5.1) refers to the modeling and normalization of QCD-multijet events from data sidebands. In section 5.2, the observed jet-multiplicity distributions are discussed.

The next two sections discuss the compatibility between observed and simulated distributions in more detail. In section 5.3, the modeling of kinematic distributions is discussed for the main background processes. Two dedicated control regions either enriched in W-boson-plus-jets production or \( t\bar{t} \) production are shown. In section 5.4, the observed kinematic distributions in the “2 jets, 1 btag”-signal category are discussed.

**Plot Style**  In all figures, the background events are grouped into three distinct classes. One class consists of processes in which top quarks are involved. This class includes \( t\bar{t} \) events, \( s^- \), and \( tW \)-channel single-top-quark events, and is dominated by \( t\bar{t} \) events. A second class refers to vector-boson production. This class is dominated by W-boson-plus-jets and furthermore includes Z-boson-plus-jets and Diboson events. QCD-multijet production, which is estimated from data sidebands, is plotted separately. The individual composition and event yield of all processes are given in detail in tables 4.3 and 4.4.

The main part of each figure shows the comparison of the stacked simulated events and data. A second, smaller part on the bottom shows the distribution of the residuals between observed and expected distributions. The residuals are defined as \( (\frac{N_{\text{Data}} - N_{\text{MC}}}{N_{\text{MC}}}) \), in which \( N_{\text{Data}} \) refers to the number of events in data, and \( N_{\text{MC}} \) refers to the number of simulated events.

All simulated signal and background contributions are normalized to the integrated luminosity of the analyzed data according to their SM-production cross sections, except for QCD-multijet processes, which are estimated from data. The first bin of each histogram includes the underflow bin, and the last bin of a histogram includes the overflow.

Each figure has a label in the top-left corner that identifies the event category. Events are categorized according to the lepton flavor (\( e, \mu \)), jet multiplicity, and b-tagged-jet multiplicity (cf. sec 4.1). The integrated luminosity of the analyzed data, as well as the center-of-mass energy, are reported in a supplementary label, which is located in the top-right corner. In most figures, the compatibility between the simulated stacked distribution and the data distribution is expressed in terms of a binned \( \chi^2 \) test [231] or a \( \chi^2 \) test [231]. The test-probability values are located in the top-right corner of the figure.
5 Background Estimation and Evaluation of Simulated Event Modeling

5.1 QCD-Multijet-Background Estimation

The simulation of QCD-multijet events usually has large uncertainties. The number of simulated QCD-multijet events corresponds to a small integrated luminosity, since the cross section of QCD-multijet production is high and has diversified processes contributing, but the number of computing resources is limited. Moreover, tight lepton-identification and -isolation criteria and tiny jet-flavor-tagging-misidentification probabilities ($O(10^{-3})$) reject QCD-multijet events by several orders of magnitudes. Thus, simulated QCD-multijet events have large event weights in the dedicated phase space in which this analysis is performed.

Therefore, QCD-multijet are extrapolated from data-sidebands to the signal region rather than relying on event simulation. The aim of this section is to describe how the QCD-multijet-event yield is normalized and how the shape of the BDT-discriminant distribution is derived for QCD-multijet events that are obtained from a data-sideband region. In this analysis, one QCD-multijet model is derived for events with muons and one model is derived for events with electrons. Before both models are discussed in the following two subsections, the general strategy is explained.

The strategy to derive the shape of the BDT-discriminant distribution for QCD-multijet events is as follows (cf. fig. 5.1). First, a data sample that is enriched in QCD-multijet events is selected by inverting an event-selection criterion. This “inversion-criterion” is chosen such that it is mostly uncorrelated to the input distributions that are used for the BDT-classifier training. This avoids possible biases when extrapolating the shape of the BDT-discriminant distribution from the data-sideband to the signal region. For events with muons, the relative muon isolation ($relIso$) is used as the inversion criterion. For events with electrons, either the electron-identification criteria, except for the conversion rejection, or the conversion-rejection criteria itself are inverted.

Second, the BDT-discriminant distribution is calculated from QCD-multijet events in the data sideband. The (normalized) distribution is used to model QCD-multijet events in the signal region. Furthermore, it is demonstrated that the inversion criteria, which are used to enriched the data sample with QCD-multijet events, do not bias the BDT-discriminator distribution in the signal region.

![Figure 5.1: Concept of the BDT-discriminator-shape estimation from data-sideband regions for QCD-multijet events. Different inversion criteria are used for events with electrons (left) and events with muons (right).](image)
For the **normalization of QCD-multijet events**, a template fit to the observed $E_{T}^{\text{miss}}$ or $M_{T}(W \text{ boson})$ distributions with two components is used. One template consists of QCD-multijet processes. The other template consists of processes that include at least one W boson or Z boson. The fit is performed in each category after applying the nominal event selection, except for the $E_{T}^{\text{miss}}$ or $M_{T}(W \text{ boson})$ criteria.

Both templates have distinct shapes in $E_{T}^{\text{miss}}$ and $M_{T}(W \text{ boson})$ distributions (fig. 5.2). QCD-multijet events cluster at low values of $E_{T}^{\text{miss}}$ and $M_{T}(W \text{ boson})$ due to the absence of prompt neutrinos. Processes with a W boson have a significant amount of $E_{T}^{\text{miss}}$ due to the prompt neutrino from the W-boson decay, and the $M_{T}(W \text{ boson})$ distribution has a Jacobian peak. Events with Z bosons give only minor contributions.

The shape of the $E_{T}^{\text{miss}}$ or $M_{T}(W \text{ boson})$ distributions are estimated from data-sidebands similar as for the distribution of the BDT discriminant. The shape of the distribution for processes with W bosons is taken from simulation with a normalization according to the SM prediction. The normalization of the individual processes is varied in order to test the stability of the fit results. In particular, it is checked that the number of signal events does not bias the number of estimated QCD-multijet events. As an example, figure 5.2 shows the obtained templates to estimate the QCD-multijet normalization for events in the “2 jets, 1 btag” category.

The number of observed events $N_{\text{Data}}$ is given by

$$N_{\text{Data}} = N_{\text{QCD}} \cdot T_{\text{QCD}} + N_{W\text{-boson processes}} \cdot T_{W\text{-boson processes}},$$

in which $T_{\text{QCD}}$ and $T_{W\text{-boson processes}}$ are the normalized templates for QCD-multijet events and events with W (or Z) bosons, $N_{\text{QCD}}$ is the number of fitted QCD-multijet events, and $N_{W\text{-boson processes}}$ is the fitted number of remaining events. $T$ is either the $E_{T}^{\text{miss}}$ distribution in case of events with electrons or the $M_{T}(W \text{ boson})$ distribution in case of events with muons. The number of QCD-multijet events after the full event selection is given by

$$N_{\text{QCD, selection}} = N_{\text{QCD}} \cdot \left( \sum_{i=1}^{N_{\text{bins}}} T_{\text{QCD}}^{i} \right)^{X}.$$
in which $T_{\text{QCD}}^i$ is the relative fraction of QCD-multijet events in bin $i$ of the normalized template $T$ and the sum is meant to run over all bins that satisfy the criteria $X$ with $X := MT(W\text{ boson}) > 40\text{ GeV}/c^2$ for events with muons and $X := E_T^{\text{miss}} > 35\text{ GeV}$ for events with electrons.

5.1.1 QCD-Multijet Events with Muons

In order to obtain a QCD-multijet model with inverted muon-isolation criteria, events triggered by a single-muon trigger without any isolation criteria are used ($\text{HLT}_\mu15_v2$). The recorded dataset corresponds to an integrated luminosity of $45.9\text{ pb}^{-1}$. The data were recorded at the beginning of the data-taking period in 2011. In particular, this period had low-Pile-Up conditions. However, the overall impact on this measurement due to the number of additional interactions in QCD-multijet events is expected to be negligible, since uncertainties of the QCD-multijet yield are much larger.

Event-Yield Estimation In order to almost exclusively enrich the data sample with QCD-multijet events, events with loosely isolated muons are selected. These muons are required to have a relative isolation of $0.35 < \text{relIso} < 1.0$. The resulting event sample is enriched with QCD-multijet events with a purity of more than 95%. The following uncertainties of the yield estimation are taken into account.

- The statistical uncertainty of the fit is considered. This uncertainty is $\pm4.5\%$ in the “2 jets, 1 btag” category and $\pm13.5\%$ in the “3 jets, 1 btag” category. The statistical uncertainty is at least $\pm60\%$ for fits in all other categories, since the QCD-multijet yield is small in these categories. In categories with at least two b-tagged jets, the QCD-multijet event yield is small due to the small mistagging efficiency. In categories with at least four jets with at least one b-tagged jet, the $t\bar{t}$ background is much larger than the QCD-multijet yield.

- The predicted SM $t\bar{t}$-channel cross section is varied by multiplicative factors of 0.5 and 2 in order to check the sensitivity of the QCD-multijet-yield estimation w.r.t. the signal contamination. It is found that a variation of the signal cross section has a negligible effect on the fitted QCD-multijet yield.

- The expected SM $t\bar{t}$ cross section is varied by multiplicative factors of 0.5 and 2 in order to account for template-shape variations due to the background composition, and due to number of dileptonic-$t\bar{t}$ events. The resulting uncertainty is well below the statistical uncertainties of the fit.

- The definition of the anti-isolation region is varied to $0.35 < \text{relIso} < 0.5$. This uncertainty is the dominating uncertainty in “2 jets, 1 btag” ($\pm39\%$) and “3 jets, 1 btag” ($\pm38\%$) categories, and is of about the same order as the statistical uncertainty in all other categories.

The total uncertainty of the QCD-multijet-yield estimation is given by the square-root of all quadratically summed contributions.

Figure 5.3 (left) exemplarily shows the obtained transverse-W-boson-mass distribution for QCD-multijet events. Moreover, the distribution is shown for varied definitions of the data-sideband region. The filled-area histogram represents the nominal template with $\text{relIso} > 0.35$. The obtained shape significantly depends on the relative isolation of the muon. Therefore, the estimated QCD-multijet yield in the region $MT(W\text{ boson}) > 40\text{ GeV}/c^2$ has large uncertainties.
The obtained QCD-multijet-event yield and corresponding uncertainties are summarized in table 5.1.

<table>
<thead>
<tr>
<th>Category</th>
<th>Nominal Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$, “2 jets, 1 btag”</td>
<td>261 ± 39%</td>
</tr>
<tr>
<td>$\mu$, “2 jets, 2 btags”</td>
<td>6 ± 100%</td>
</tr>
<tr>
<td>$\mu$, “3 jets, 1 btag”</td>
<td>87 ± 50%</td>
</tr>
<tr>
<td>$\mu$, “3 jets, ≥ 2 btags”</td>
<td>11 ± 100%</td>
</tr>
<tr>
<td>$\mu$, “≥ 4 jets, 1 btag”</td>
<td>32 ± 100%</td>
</tr>
<tr>
<td>$\mu$, “≥ 4 jets, ≥ 2 btags”</td>
<td>23 ± 120%</td>
</tr>
</tbody>
</table>

Table 5.1: Estimated QCD-multijet yield for events with muons after the full event selection.

Figure 5.3: QCD-multijet model that is obtained from data with non-isolated muons. The left plot shows the transverse-W-boson-mass distribution, which is used to estimate the event yield. The right plot shows the BDT-discriminator distribution for events with muons in the “2 jets, 1 btag”-signal category ($M_T(W \text{ boson}) > 40 \text{ GeV}/c^2$). The red and blue distributions are obtained by varying the definition of the data-sideband region. The filled-area histogram is the nominal template with $\text{relIso} > 0.35$.

BDT-discriminator-shape estimation The BDT-discriminator distributions that are obtained from the varied isolation regions are compatible among each other within statistical uncertainties. As an example, the BDT-discriminator distribution for events with muons in the “2 jets, 1 btag”-signal category is shown in figure 5.3 (right).

A cross check is performed in order to test the extrapolation of the QCD-multijet kinematics from events with anti-isolated muons to events with isolated muons. In particular, this check does not rely on any information from simulation. Figure 5.4 shows the BDT-discriminator distribution for events with isolated muons in the “2 jets, 0 btags” category, including the QCD-multijet model that is obtained from data with non-isolated muons. The events are normalized to the fit results.

- Figure 5.4 (left) shows the BDT-discriminator distribution for a region that is enriched in QCD-multijet events ($M_T(W \text{ boson}) < 40 \text{ GeV}/c^2$).
- Figure 5.4 (right) shows the discriminator distribution for a region that is QCD-multijet depleted ($M_T(W$ boson) $> 40 \text{ GeV}/c^2$).

The extrapolated QCD-multijet model reasonably reproduces the QCD-multijet events with isolated muons. The extrapolated model slightly is more signal-like than expected from anti-isolation region, which results in an excess of data in the left tail of the BDT-discriminator distribution (fig. 5.4 left). These residual discrepancies are covered by the uncertainties of the normalization, i.e. composition of the W-boson-plus-jets and QCD-multijet events.

5.1.2 QCD-Multijet Events with Electrons

In order to obtain the QCD-multijet model for electron events from data, the same triggers as for the default event selection can be used (cf. table 4.2). Using the electron-cross trigger, which requires b-tagged jets, has the advantage that the QCD-multijet model is enriched with jets from B hadrons that mimic the signal signature quite well. Moreover, the QCD-multijet model is derived with similar Pile-Up conditions as the nominal events. The data sample is enriched in QCD-multijet events with isolated electrons by either inverting the electron-identification criteria, except for the conversion rejection, or inverting the conversion-rejection criteria. In contrast to QCD-multijet events with muons, electrons in QCD-multijet processes mostly stem from photon conversions. Electrons from conversions are well isolated.

Event-yield estimation  The estimated event yields for the full event selection are summarized in table 5.2. The following uncertainties of the yield estimation are taken into account.

- The statistical uncertainty of the fit is $\pm 6.5\%$ for fits in the “2 jets, 1 btag” category, about $\pm 15.0\%$ for fits in the “3 jets, 1 btag” category, and up to $\pm 50\%$ in other categories. The statistical uncertainty is increased for fits in categories with at least four jets or at least two b-tagged jets, since the QCD-multijet yield is low in these categories (cf. table 5.2).
5.1 QCD-Multijet-Background Estimation

![Graph showing Emiss T distribution and BDT output for QCD-multijet model](image)

Figure 5.5: QCD-multijet model as obtained from data with electrons, either with inverted conversion rejection (filled) or inverted electron-identification criteria (red). The left plot shows the $E_{\text{miss}}^{T}$ distribution, which is used to estimate the yield. The right plot shows the BDT-discriminator distribution for events with electrons in the “$2\text{jets, 1 btag}$”-signal category ($E_{\text{T}}^{\text{miss}} > 35\text{ GeV}$).

- It is checked that the number of $t$-channel-signal events does not bias the estimation of the QCD-multijet yield.

- The expected SM $tt$ cross section is varied by multiplicative factors of 0.5 and 2 in order to check the sensitivity of the yield-estimation procedure w.r.t. the background composition. The uncertainty resulting due to this variation is about $\pm 20\%$ for fits in the “$2\text{jets, 1 btag}$” category, about $40\%$ for fits in the “$2\text{jets, 2 btags}$” category, about $10\%$ for events with two b-tagged jets, and of the order of the statistical uncertainty in all other categories.

- The definition of the data-sideband region is varied. The $E_{\text{miss}}^{T}$ distribution obtained from events with electrons from conversions is softer than the $E_{\text{miss}}^{T}$ distribution obtained from electrons with inverted electron-identification. Fits with templates from both data-sideband regions are performed. This is the dominating uncertainty.

The average of the results of both fits with varied data-sideband definitions is used as the nominal yield estimate. The total uncertainties are estimated to be $\pm 100\%$ such that they cover both fit results.

Figure 5.5 (left) shows the comparison of the nominal $E_{\text{T}}^{\text{miss}}$ distribution (filled area), which is enriched in electrons from conversions, with the distribution that is obtained by inverting all other electron-identification criteria (red line). As an example, the “$2\text{jets, 1 btag}$” category is shown. The shapes are significantly different, especially in the tails of the distribution. Therefore, the estimated yield for events with $E_{\text{T}}^{\text{miss}} > 35\text{ GeV}$ has large uncertainties.

**BDT-discriminator-shape estimation** Figure 5.5 (right) shows the BDT-discriminator distribution in the “$2\text{jets, 1 btag}$” category. Here, the QCD-multijet model enriched with electrons from conversions (filled area) is shown in comparison to the distribution obtained from events by inverting the electron-identification criteria, except for the conversion rejection. Both distributions are in agreement within statistical uncertainties. Residual uncertainties are covered by the statistical precision of the simulated events.
Table 5.2: Estimated QCD-multijet yield for events with electrons after the full event selection.

<table>
<thead>
<tr>
<th>Category</th>
<th>Nominal Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e$, “2 jets, 1 btag”</td>
<td>$240 \pm 100%$</td>
</tr>
<tr>
<td>$e$, “2 jets, 2 btags”</td>
<td>$5% \pm 100%$</td>
</tr>
<tr>
<td>$e$, “3 jets, 1 btag”</td>
<td>$184 \pm 100%$</td>
</tr>
<tr>
<td>$e$, “3 jets, $\geq 2$ btags”</td>
<td>$36% \pm 100%$</td>
</tr>
<tr>
<td>$e$, “$\geq 4$ jets, 1 btag”</td>
<td>$91% \pm 100%$</td>
</tr>
<tr>
<td>$e$, “$\geq 4$ jets, $\geq 2$ btags”</td>
<td>$96% \pm 100%$</td>
</tr>
</tbody>
</table>

Figure 5.6 confirms the extrapolation of the QCD-multijet kinematics from the data-sideband to the “signal” region. Here, the sample with QCD-multijet events enriched with electrons from photon conversions is used to model QCD-multijet events, which pass the nominal event selection. The observed BDT-discriminator distribution of events with electrons that passed all electron-identification and conversion-rejection criteria is reasonably described by events with electrons that are isolated, but enriched in conversions. The BDT-discriminator distribution is shown in both the QCD-multijet enriched region with $E_T^{\text{miss}} < 35$ GeV (left) and for the nominal event selection with $E_T^{\text{miss}} > 35$ GeV (right). As an example, figure 5.6 shows the cross check for events in the “2 jets, 0 btags” category. Both distributions are normalized to the fit results.
In the analysis at hand, all simulated signal and background contributions are normalized to the integrated luminosity of the analyzed data according to their SM-production cross sections, except for QCD-multijet processes, which are estimated from data. The general strategy is to incorporate the normalization of individual processes as nuisance parameters to the likelihood function and fit them during the statistical inference (cf. sec. 4.6).

Uncertainties of the background-event yield are sizable with up to $O(100\%)$, e.g. $\pm 50\%$ on W-boson-plus-light-jets events, $\pm 100\%$ on W-boson-plus-heavy-flavored-jets events, and $\pm 15\%$ on $t\bar{t}$ events. These uncertainties of the background normalization reflect the knowledge of process-normalization in the dedicated phase space in which this analysis is performed, and are discussed in section 6.1.11. The uncertainty of the process normalization is not explicitly shown in the following plots.

Figure 5.7 shows the obtained jet-multiplicity distribution for events with electrons (left) and muons (right). Each bin refers to events with a given number of jets. Observed and expected distributions are compatible with each other at the level of a few percent. Residual corrections are covered by the uncertainties on the process normalization. Events with electrons contain more $t\bar{t}$ and $t$-channel events relative to W-boson-plus-jets events due to the jet-flavor-tagging requirement within the trigger. More QCD-multijet events remain in categories with electrons than in categories with muons.

Figure 5.7: Jet-multiplicity distribution for events with electron (left) or muon final states (right).

In figure 5.8, the jet-multiplicity distribution is divided into several categories with an exclusive number of b-tagged jets. Events with electrons are shown in the left column, and events with muons in the right column. The first row shows the jet-multiplicity distributions for events with 0-tagged jets, the second row for events with one b-tagged jet, and the third row for events with $\geq 2$ b-tagged jets.

The majority of the observed events do not have any b-tagged jet. The jet-multiplicity distributions for events with zero b-tagged jets (fig. 5.8 first row) show the same trend as for events with an inclusive number of tagged-jets. Here, the observed number of events is in good agreement with the simulation for events with both electrons and muons. Events with electrons and at least seven jets are underestimated, but only few events remain in these categories.

In figure 5.8, the jet-multiplicity distribution is divided into several categories with an exclusive number of b-tagged jets. Events with electrons are shown in the left column, and events with muons in the right column. The first row shows the jet-multiplicity distributions for events with 0-tagged jets, the second row for events with one b-tagged jet, and the third row for events with $\geq 2$ b-tagged jets.

The majority of the observed events do not have any b-tagged jet. The jet-multiplicity distributions for events with zero b-tagged jets (fig. 5.8 first row) show the same trend as for events with an inclusive number of tagged-jets. Here, the observed number of events is in good agreement with the simulation for events with both electrons and muons. Events with electrons and at least seven jets are underestimated, but only few events remain in these categories.
The jet-multiplicity distributions for events with electrons and with exactly one b-tagged jet (fig. 5.8, second row) are compatible between data and simulation, except for events with a high number of jets, which are overestimated by the simulation. For events with muons, up to 20\% more events are observed than predicted, while the agreement is better for higher jet multiplicities. In particular, the number of observed events in the electron-“2 jets, 1 btag”-signal category is compatible with the prediction (second row, right plot, first bin). Instead, the muon-“2 jets, 1 btag”-signal category shows an excess of data (second row, right plot, first bin).

The agreement between events with electrons and muons can differ due to varied sources of systematic uncertainties in both decay channels. These include the lepton-trigger efficiencies, the lepton-reconstruction and -identification efficiencies, and the QCD-multijet estimation. The uncertainties of the jet-flavor-tagging efficiencies have an altered impact on events with electrons or muons, since the electron channel uses jet-flavor tagging already at the trigger level. Moreover, the background-composition differs between both categories due to the varied lepton-$p_T$ cut and jet-flavor-tagging efficiency in the electron-jet trigger. Among others, all these sources of systematic uncertainties (cf. chapter 6) are treated within the statistical inference (cf. sec. 4.6). Furthermore, the $E_T^{\text{miss}}$-selection criterion is used in the electron channel, and $M_T(W \text{ boson})$-selection criterion in the muon channel. The observed distributions are found to be stable against alternation of $E_T^{\text{miss}}$ and $M_T(W \text{ boson})$ selection criteria.

The observed jet-multiplicity distribution for events with two or more b-tagged jets (fig. 5.8, third row) is compatible between data and simulation for events with electrons. The simulation underestimates the event yield for events with muons, up to five jets, and two or more b-tagged jets.

Moreover, several sources of systematic uncertainties have an effect on the predicted distributions for jet multiplicity and b-tagged-jet multiplicity. As an example, the jet-flavor-tagging efficiency shifts b-tagged-jet-multiplicity distributions, and the $Q^2$ scale and jet-energy scale shift the total-jet-multiplicity distribution.

The jet-multiplicity distributions do not give a conclusive answer which processes are over- or underestimated by simulation. All background processes ($t\bar{t}$, W boson plus jets, QCD multijet) have large uncertainties that potentially explain an excess or deficit of events.

The next two sections discuss the shape of the kinematic distributions in $t\bar{t}$-enriched and W-boson-plus-light-jets-enriched categories, as well as the signal region. The idea is that if the kinematics of the individual processes are reasonably described, the statistical inference will be able to disentangle the normalization of individual processes by using the multivariate discriminator output. Hence, uncertainties of predicted process normalizations are incorporated as nuisance parameters in the statistical inference.

---

1For events with at least one b-tagged jet, the dominant contribution of the W-boson-plus-jets background are W bosons that are produced in association with heavy-flavored jets. Their normalization has even larger uncertainties (±100\%) and is assumed to be independent of the normalization of W-boson-plus-light-jets events, which can be estimated with events zero b-tagged jets.
5.2 Event Yield and Jet-Multiplicity Distribution

Figure 5.8: Jet-multiplicity distribution for events with exactly zero b-tagged jets (first row), one b-tagged jet (second row), and \( \geq \) two b-tagged jets (third row) for events with electron (left) or muon final states (right).
5 Background Estimation and Evaluation of Simulated Event Modeling

5.3 Top-Quark-Pair- and W-Boson-plus-Jets-Enriched Phase Space

In the following section, the agreement between data and simulation in the background-enriched phase space is checked. Events in the “2 jets, 0 btags” category are enriched with W-boson-plus-jets production, while the “≥ 4 jets, 1 btag” category is used as a control region for \( t\bar{t} \) production. The W-boson-plus-jets-enriched control regions, i.e. events with zero b-tagged jets, are not explicitly used within the statistical inference. However, the “2 jets, 0 btags” category serves as a control region for the modeling of the kinematics of the W-boson-plus-jets events, which are an important background to the \( t\bar{t} \)-channel signal in the “2 jets, 1 btag” category.

This section is organized as follows. First, basic kinematic quantities (\( p_T, \eta, \) charge) of charged leptons (\( e, \mu \)), leading jet and second-leading jet, \( E_T^{\text{miss}} \), as well as the reconstructed W-boson mass are considered. Then, the kinematics (\( p_T, \eta, \) and mass) of the reconstructed top-quark hypotheses are discussed. The distribution of the azimuthal angle \( \phi \) is exemplarily shown for W-boson-plus-jets events with muons. All remaining distributions that are input to the BDT training (cf. table 4.6) are addressed in the following, as well as the resulting BDT-discriminator distributions. The evaluation of the BDT-discriminator output in the background-enriched categories is used as a “litmus test” that probes the agreement between data and simulated events. In particular, the proper descriptions of kinematic distributions and correlations among them are tested.

In particular, the observed normalization differs by a few percent between \( t\bar{t} \) events with electrons and muons in the “≥ 4 jets, 1 btag” category (fig. 5.8, second row), which is attributed to the parametrization of the electron cross-trigger. The compatibility between simulated and data distributions is quantified with binned KS tests [224] or \( \chi^2 \)-tests [231] and discussed in the text.

The KS test probability is expected to be uniformly distributed between zero and one if both distributions are compatible, and zero if both distributions are incompatible. The KS test only accounts for the compatibility of the shape of two distributions. However, in this analysis, the KS test only is a “weak” classifier to quantify the compatibility between two distributions. This has two reasons.

First, the KS test does not take uncertainties of the shape of a distribution into account, but shape-changing uncertainties are taken into account during the statistical inference and final measurement. As an example, the uncertainty of the jet-energy-scale typically shifts the jet-\( p_T \) distribution.

Second, the KS test does not take uncertainties of the process normalization into account. In contrast, the statistical inference does take these uncertainties into account. The uncertainties of the normalization of individual processes are different among physics processes. As an example, the normalization of W-boson-plus-jets events typically has larger uncertainties than the normalization of events from \( t\bar{t} \) production. The normalization of individual processes, however, defines the composition among simulated processes, and can have a significant impact on the obtained shape of the (stacked) simulated-event distribution.

The \( \chi^2 \) test is used as a second compatibility test, which additionally takes the normalization into account. The \( \chi^2 \) test is, however, also a “weak” classifier in case of systematic uncertainties (due to the same reasons as described for the KS test).
5.3 Top-Quark-Pair- and W-Boson-plus-Jets-Enriched Phase Space

Figure 5.9: Muon $p_T$ (top left), $\eta$ (top right), $\phi$ (bottom left), and charge (bottom right) distributions in the "2 jets, 0 btags" category.
5.3.1 Basic Kinematic Quantities

Charged lepton The distributions of transverse momentum, the pseudo-rapidity, and azimuthal angle of the muon for events in the “2 jets, 0 btags” are shown in fig. The transverse-momentum distribution peaks at $p_T \approx 35$ GeV/c and steeply falls towards higher momenta. The pseudo-rapidity distribution for muons is flat, except for binning effects in the endcaps, and the muon-reconstruction efficiency is slightly higher in the central region of the detector. The CMS detector fully covers the azimuthal-angle ($\phi$) range. The resulting $\phi$ distributions of muon and electron are uniformly distributed around the beamline. The observed and simulated $\phi$ distributions well agree in all categories. As an example, the $\phi$ distribution for muons in the “2 jets, 0 btags” category is shown in the bottom-left plot in figure The simulation well describes the kinematic distributions over the whole spectrum. Some minor residual corrections can be attributed to the description of the QCD-multijet contribution, and are covered by the normalization uncertainties of both QCD-multijet and W-boson-plus-jets production. Moreover, figure confirms that the $p_T$ and $\eta$ distributions are well described for muons and electrons in the $t\bar{t}$ enriched phase space (“$\geq 4$ jets, 1 btag” category).

![Figure 5.10: Distributions of the electron $p_T$ (top left), electron $\eta$ (bottom left), muon $p_T$ (top right), and muon $\eta$ (bottom right) in the “$\geq 4$ jets, 1 btag” category.](image)

The observed distribution of the muon charge (fig. bottom right) in W-boson-plus-jets events is poorly described. This is surprising, since dedicated measurements observe a good agreement between data and prediction. First, the inclusive ratio of $W^+\to W^-\nu$ production well agrees between data and simulation (POWHEG BOX NLO+PS) description with three different
5.3 Top-Quark-Pair- and W-Boson-plus-Jets-Enriched Phase Space

PDF sets), as well as with NNLO calculations [165]. Second, the differential distribution of the $|\eta|$-dependent-charge asymmetry, which is defined by [232]

$$A_{W^\pm \text{production}}(\eta) = \frac{d\sigma}{d\eta}(W^+ \rightarrow l^+ \nu_l) - \frac{d\sigma}{d\eta}(W^- \rightarrow l^- \bar{\nu}_l)$$

shows a reasonable agreement between data and simulation [232, 233].

The lepton charge is expected to be asymmetric for W-boson-plus-jets production (cf. eq. 5.3). Positively charged leptons are preferred for similar reasons as for $t$-channel production (cf. sec. 4.2). The production of $W^+$ and $W^-$ bosons is mostly determined by the valence quarks of the colliding protons, and is sensitive to the PDFs. $W^+$ bosons are mostly produced by valence-up quarks, and $W^-$ bosons are mostly produced by valence-down quarks. The charge asymmetry is expected to be $\eta$-dependent, since valence-up quarks carry on average a higher fraction of the proton momentum than valence-down quarks [233]. Instead, $t\bar{t}$ production is lepton-charge symmetric due to $q\bar{q}$ or $gg$ initial states.

Figure 5.12 (left) shows the differential distribution of the W-boson-charge asymmetry (eq. 5.3) as measured in data and predicted by simulation. The distributions are shown for events with a muon in the “2 jets, 0 btags” category. The observed charge asymmetry is lower than the simulated distribution over the full $|\eta|$ acceptance. Since dedicated measurements [232, 233] show a reasonable agreement between data and simulation, the observed differences in the description of the charge asymmetry are expected to be either due to the choice of the PDF set (CTEQ6L1 [52]) or due to the modeling of the W-boson events with additional partons with MADGRAPH and PYTHIA.

Figure 5.12 (right) shows the observed difference in the charge asymmetry between data and simulation. It is mostly independent of $|\eta|$. The difference is slightly larger for higher $|\eta|$ values. The residual difference \[ \frac{\text{data} - \text{prediction}}{\text{prediction}} \] in the charge asymmetry between data and simulation can be parametrized as

$$f(A_{W^\pm \text{production}}(\eta)) = (-0.437 \pm 0.034) + (0.056 \pm 0.025) \cdot |\eta|. \quad (5.4)$$

The inclusive measurement of the $t$-channel-production cross section is uncorrelated w.r.t. the lepton charge, and independent of the observed discrepancy in $A_{W^\pm}$ for W-boson-plus-jets events. The measurement of the ratio of top-quark-to-top-antiquark production in $t$-channel events takes this mismodeling into account.

The distribution of the lepton charge in the $t\bar{t}$-enriched phase space for events in the “$\geq 4$ jets, 1 btag” category confirms that W-boson-plus-jets events with a negatively-charged lepton are underestimated for both events with electrons and muons (fig. 5.11). The disagreement is less pronounced in this category due to the overall lower yield of W-boson-plus-jets events and due to the fact that $t\bar{t}$ is charge symmetric. Moreover, the overall yield of W-boson-plus-jets events is underestimated by a few percent for W-boson-plus-jets events with muons (fig. 5.11 left plot).

Jets The transverse-momentum and pseudo-rapidity distributions for the leading jet and second-leading jet of W-boson-plus-jets events are shown in figure 5.13. A reasonable agreement between data and simulation is observed, even if the transverse-momentum distribution for jets is softer in data than predicted by the simulation. The jet-\eta distributions show a good agreement between data and simulation, except for jets within $3 < |\eta| < 4$. Here, significantly more events are observed in data than predicted by the simulation. The excess in data
Figure 5.11: Distributions of the lepton charge in the “≥ 4 jets, 1 btag” category for events with electron (left) or muon final states (right).

Figure 5.12: Charge asymmetry of the W boson as measured in events with a muon in the “2 jets, 0 btags” category (left), and residuals between data and simulation (right).
is most prominent at $|\eta| \approx 3.5$, and is about 10% when considering shape-effects only. Two sources of systematic uncertainties have an effect on the shape of the jet-$p_T$ and jet-$\eta$ distributions, namely the jet-energy calibration and the $Q^2$ scale. The shape of the observed jet-$p_T$ distributions is compatible with the $Q^2$-scale variations at 1\sigma level. Furthermore, the $Q^2$-scale variations predict ±(5-10)% more (less) jets close to the beamline than in the central-detector region. The jet-energy-calibration uncertainties cover the observed discrepancies within 2\sigma in both jet-$p_T$ and jet-$\eta$ distributions. Furthermore, the uncertainty of the normalization of the W-boson-plus-jets events covers even larger variations.

Figure 5.13: Distributions of the leading-jet $p_T$ (top left), leading-jet $\eta$ (top right), second-leading-jet $p_T$ (bottom left), and second-leading-jet $\eta$ (bottom right) in the muon “2 jets, 0 btags” category.

The leading-jet $p_T$ and $\eta$ distributions for jets from t\bar{t} in the “≥ 4 jets, 1 btag” category are shown in figure 5.14. Events with electrons (left column) and muons (right column) are separately shown, since events in both lepton categories are selected with varied criteria. All kinematic distributions are also shown for the second-leading jet in figure 5.15.

The observed jet-$p_T$ spectrum is softer than predicted by the simulation. The same trend is observed for W-boson-plus-jets events. However, W-boson-plus-jets events are a minor contribution to events in the “≥ 4 jets, 1 btag” category, and W-boson-plus-jets events only partially explain the disagreement. Furthermore, the $Q^2$-scale variations have a smaller effect on the jet-$p_T$ spectra of t\bar{t} processes. The same is true for the jet-energy-scale variations, since jets from t\bar{t} production mostly are central in the detector, in which the jet-energy-scale uncertainties are small. Thus, the disagreement is expected to be a “real” effect of the event simulation. It is correlated to the description of the top-quark-$p_T$ spectrum and is discussed in more detail in

5.3 Top-Quark-Pair- and W-Boson-plus-Jets-Enriched Phase Space
one of the following paragraphs. The altered selection efficiency of $t\bar{t}$ events between data and simulation is covered by the normalization uncertainties.

Jets from $t\bar{t}$ production are more central in the detector than jets from W-boson-plus-jets production. The predicted jet-pseudo-rapidity distributions well agree with the observed distributions, except for residual corrections to the $t\bar{t}$ normalization (or other systematic uncertainties that affect the process normalization, e.g. jet-flavor-tagging efficiencies). The excess of jets around $|\eta| \approx 3.5$, which is observed for W-boson-plus-jets events, cannot be confirmed for $t\bar{t}$ production, but the statistical precision is low in that region.

Figure 5.14: Distributions of the leading-jet $p_T$ (top row) and leading-jet $\eta$ (bottom row) in the "$\geq 4$ jets, 1 btag" category for events with electron (left) or muon final states (right).
Figure 5.15: Distributions of the second-leading-jet $p_T$ (top row) and second-leading-jet $\eta$ (bottom row) in the “≥ 4 jets, 1 btag” category for events with electron (left) or muon final states (right).
$E_T^{\text{miss}}$ and transverse W-boson mass  The $E_T^{\text{miss}}$ is shown in both W-boson-plus-jets-enriched and $t\bar{t}$-enriched control regions in figure 5.16. An overall good agreement of the $E_T^{\text{miss}}$ response is observed. The observed $E_T^{\text{miss}}$ distribution is slightly broader than predicted by simulation and the tails of the $E_T^{\text{miss}}$ distribution are overestimated by simulation. The low-$E_T^{\text{miss}}$ region (top plot in fig. 5.16) is sensitive to both normalization and $E_T^{\text{miss}}$-distribution shape of QCD-multijet events. The QCD-multijet model is estimated from data with large uncertainties and might explain the observed discrepancies in the left tail of the distribution.

The $E_T^{\text{miss}}$ is an important input to the reconstruction of the transverse W-boson mass, whose distributions are shown in figure 5.17 for W-boson-plus-jets events and in figure 5.18 for $t\bar{t}$ events. The reconstructed W-boson mass is shown either before (left figure) or after (right figure) applying the neutrino-reconstruction algorithm. The QCD-multijet yield is underestimated at low transverse-W-boson-mass values (both plots in fig. 5.17), but covered by the uncertainties of the QCD-multijet prediction.

The tails of the $M_T(W \text{ boson})$ distributions ($M_T(W \text{ boson}) \geq M(W \text{ boson})$) show an excess of data w.r.t. the prediction (right plot in fig. 5.17). The data excess originates from residual corrections as observed in the $E_T^{\text{miss}}$-response distribution (fig. 5.16). The reconstructed transverse-W-boson mass can be larger than the invariant-W-boson mass if the $E_T^{\text{miss}}$ is mismeasured. If the $E_T^{\text{miss}}$ is mismeasured, complex solutions for the z-component of the neutrino momentum arise while applying the neutrino-reconstruction algorithm. These complex solutions are then resolved by varying $E_T^{\text{miss}}$ (cf. sec. 4.4.2).

As a result of the neutrino-reconstruction algorithm, events with values $M_T(W \text{ boson}) > M(W \text{ boson})$ obtain a reconstructed transverse-W-boson mass that is equal to the invariant-W-boson mass $M_T(W \text{ boson}) = M(W \text{ boson}) = 80.4 \text{ GeV}/c^2$. Figure 5.17 (right) shows that a reasonable description of the observed transverse W-boson mass is obtained after applying the neutrino-reconstruction algorithm. Small residual differences of about 7% remain due to the $E_T^{\text{miss}}$ resolution. Uncertainties are covered by jet-energy-scale and $Q^2$-scale uncertainties within 1.5σ. However, both $E_T^{\text{miss}}$ and $M_T(W \text{ boson})$ are not a direct input to the BDT-classifier training.
Figure 5.16: Distributions of the $E_T^{\text{miss}}$ in the muon “2 jets, 0 btags” category (top plot) and “$\geq 4$ jets, 1 btag” category for events with electron (bottom left) or muon (bottom right) final states.

Figure 5.17: Reconstructed transverse-W-boson mass distributions before (left) and after (right) applying the neutrino reconstruction algorithm. The distributions are shown in the “2 jets, 0 btags” category for events with a muon final state.
Figure 5.18: Reconstructed transverse-W-boson mass distributions in the “≥ 4 jets, 1 btag” category for events with electron (left) or muon final states (right). Distributions are shown before (top row) and after (bottom row) applying the neutrino-reconstruction algorithm.
5.3 Top-Quark-Pair- and W-Boson-plus-Jets-Enriched Phase Space

5.3.2 Top-Quark Reconstruction

The distributions of the transverse momentum ($p_T$) of the top quark in $t\bar{t}$ events are shown in figure 5.19 (top row). The top-quark $p_T$ is significantly lower in data than predicted by the simulation. Variations of jet-energy scale and $Q^2$ scale only have minor effects on the predicted top-quark-$p_T$ distribution. The observed disagreement is confirmed in dedicated measurements of the differential top-quark-pair production cross sections [167]. About 15% more events are observed at a top-quark $p_T$ below 50 GeV/c, and about 15% less for a top-quark $p_T$ of about 225 GeV/c. The observed difference remains stable against the choice of the event generator (MADGRAPH +PYTHIA, POWHEG BOX +PYTHIA, and MC@NLO +HERWIG). In particular, the observed data are compatible with approximate NNLO predictions [79].

The softer top-quark-$p_T$ spectrum translates into softer jet-$p_T$ spectra, which explains that the observed jet-$p_T$ distributions are softer than predicted by the simulation (cf. fig. 5.14 and 5.15). The simulation is not reweighted according to the observed top-quark-$p_T$ spectrum, since the top-quark-$p_T$ is only weakly correlated to the BDT output as it is not directly used for the BDT training.

The distributions of the top-quark pseudo rapidity ($\eta$) show a good agreement between data and simulation (fig. 5.19 bottom row). A disagreement between simulation and data is observed at $-4 < \eta_{\text{top quark}} < -3$ for events with electrons, but cannot be confirmed in events with muons.

The reconstructed top-quark mass in $t\bar{t}$ events is shown in figure 5.20 (top row). A good agreement between data and simulation is obtained.

Figure 5.20 (bottom row) shows the mass of the top-quark-reconstruction hypothesis (“pseudo-top quark”) in W-boson-plus-jets events, which do not contain a top quark. The distribution peaks at significantly lower mass values as for $t\bar{t}$ events. In general, a good agreement between data and simulation is observed. The observed distribution of the pseudo-top-quark mass is slightly higher compared to simulation due to discrepancies in the basic-physics objects (leptons, jets, $E_{\text{T}}^{\text{miss}}$). Mainly the uncertainties of the $Q^2$ scale, W-boson-plus-jets normalization, and QCD-multijet normalization cover these differences.

An alternative top-quark reconstruction is the combination of the four-vectors of reconstructed W boson with one of the jets, but independent of jet-flavor-tagging algorithms. Here, the jet is chosen such that the mass of the top-quark candidate is close to 172.0 GeV/c$^2$. The mass distribution of the “best-top-quark hypothesis”, which is also used as an input for the BDT discriminator, is shown in figure 5.21. The best-top-quark-mass distribution has a significantly larger width for W-boson-plus-jets events than for $t\bar{t}$ events. A reasonable agreement is observed both in W-boson-plus-jets-enriched and $t\bar{t}$-enriched phase space.
Figure 5.19: Transverse-momentum distributions (top row) and pseudo-rapidity distributions (bottom row) of the b-tagged-top-quark reconstruction hypothesis in the “≥ 4 jets, 1 btag” category for events with electron (left) or muon final states (right).
Figure 5.20: Mass distributions of the b-tagged-top-quark reconstruction hypothesis in the “≥ 4 jets, 1 btag” category for events with electron (top left) or muon (top right) final states. Furthermore, the top-quark-reconstruction hypothesis is evaluated for events in the muon “2 jets, 0 btags” category, which do not contain a top-quark (bottom).
5 Background Estimation and Evaluation of Simulated Event Modeling

Figure 5.21: Reconstructed best-top-quark mass, i.e. the mass reconstructed with the jet that yields a mass closest to 172.0 GeV/c², for events in the muon “2 jets, 0 btags” category (top), and for events with electron (bottom left) or muon (bottom right) final states in the “≥ 4 jets, 1 btag” category.
5.3 Top-Quark-Pair- and W-Boson-plus-Jets-Enriched Phase Space

5.3.3 BDT-Training Variables

In total eleven variables are used to train the BDT classifier (cf. table 4.6). The charged-lepton-$p_T$ distributions are shown in figures 5.9 and 5.10. The pseudo-rapidity distributions of the leading jet are shown in figures 5.13 and 5.14. The distributions of the mass of the b-tagged-top-quark hypothesis are shown in figures 5.19 and 5.20. The distributions of the mass of the alternative top-quark reconstruction “best-top-quark hypothesis” are shown in figure 5.21. The distributions of the remaining variables that are input to the BDT-classifier training are described in the following.

The pseudo-rapidity of the most-forward (light) jet is shown in figure 5.22. An excess of events is observed for jets that are close to the beamline. This excess is are covered by jet-energy scale, $Q^2$ scale, and normalization uncertainties as discussed for the leading-jet and second-leading-jet distributions in one of the previous paragraphs (sec. 5.3.1). Moreover, the distribution has a distinct peak for jets within the HCAL-endcap-HF-transition region ($|\eta| \approx 3$), which is compatible between data and simulation. For $t\bar{t}$ production, the observed distributions are in good agreement with the prediction.

The distribution of $\cos^* (W$ boson, leading jet), which is the cosine of the angle between the reconstructed W boson and W-boson-plus-leading-jet system, is shown in figure 5.23.

Figure 5.22: Pseudo-rapidity $\eta$ distribution of the most forward jet for events with a muon in the “2 jets, 0 btags” category (top), and for events with an electron (bottom left) or muon (bottom right) in the “≥ 4 jets, 1 btag” category.

Figure 5.23: Pseudo-rapidity $\eta$ distribution of the most forward jet for events with a muon in the “2 jets, 0 btags” category (top), and for events with an electron (bottom left) or muon (bottom right) in the “≥ 4 jets, 1 btag” category.
The agreement between data and simulation is reasonable for all distributions. In the “2 jets, 0 btags” category, the simulation underestimates the number of events in which the reconstructed W-boson is back-to-back to the leading jet ($\cos^\ast(W \text{ boson, leading jet}) \rightarrow (-1, 1)$). Moreover, the region in which $\cos^\ast(W \text{ boson, leading jet}) \rightarrow -1$ is covered by the uncertainties of the QCD-multijet normalization, while the tail of the distribution is sensitive to the $Q^2$ scale of the event.

Figure 5.23: Distributions of $\cos^\ast(W \text{ boson, leading jet})$, which is the cosine of the angle between the reconstructed W boson and W-boson-plus-leading-jet system. The distributions are shown for events in the muon “2 jets, 0 btags” (top), electron “≥ 4 jets, 1 btag” (bottom left), and muon “≥ 4 jets, 1 btag” (bottom right) categories.

Figure 5.24 shows the sphericity distributions in the W-boson-plus-jets-enriched and $t\bar{t}$-enriched control regions. An overall good agreement is observed between data and simulation, except for the total normalization of both $t\bar{t}$ and W-boson-plus-jets events as discussed in the introduction of this section.

The distributions of the sum of the energies and transverse energies of all jets are presented for the background-enriched categories in figure 5.25. The sum of energies are well described across the studied phase space for W-boson-plus-jets and $t\bar{t}$ events.

The distributions of the sum of transverse energies are softer in data than predicted by the simulation for both W-boson-plus-jets-enriched and $t\bar{t}$ enriched categories, which is a conse-
sequence of the softer \( p_T \) spectra (cf. sec. 5.3.1). The discrepancies in the sum-of-transverse-energy distributions for W-boson-plus-jets events are covered by the \( Q^2 \) scale at 1\( \sigma \) level. The observed events in the \( t\bar{t}\)-electron channel (fig. 5.25, center row, left) are still compatible with the simulation within the statistical precision. Instead, the observed distribution in the \( t\bar{t}\)-muon channel (fig. 5.25, center row, right) is merely compatible with the prediction. However, the sum-of-transverse-energy distributions are of minor importance for the BDT training (cf. the ranking of the input variables in table 4.6), and the overall impact of this disagreement w.r.t. the shape of the BDT classifier output is expected to be small.

The distributions of the dijet mass of the b-tagged-jet plus light-jet hypothesis are shown in figure 5.26 (first row and left plot in the third row). A good agreement is obtained. Figure 5.26 (second row) shows the distribution of the angular separation \( \Delta R \) between the two leading jets. A good agreement is observed for both distributions for \( t\bar{t} \) events (first and second row), which is confirmed by W-boson-plus-jets events (third row).

However, the simulation underestimates the number of close-by jets whose four-vector sum yields low invariant masses. Moreover, jets are spatially more separated in data than predicted by simulation (figure 5.26, third row, right). Both effects are covered by the \( Q^2 \)-scale variations at a 1\( \sigma \) level.
Figure 5.25: Distributions of sum of energies (first row) and sum of transverse energies of all jets (second row) for events in the “≥4 jets, 1 btag” category with electrons (left) or muon final states (right). Both kinematic distributions are also shown for events in the muon “2 jets, 0 btags” category (third row).
Figure 5.26: Distributions of the mass of the dijet system (first row) and angular separation $\Delta R$ between the two leading jets (second row) for events with electron (left) or muon final states (right) in the “$\geq 4$ jets, 1 btag” category. Both distributions are also shown for events in the muon “2 jets, 0 btags” category (third row).
5 Background Estimation and Evaluation of Simulated Event Modeling

5.3.4 BDT-Discriminator Distribution

The distributions for the BDT-classifier output as obtained in the W-boson-plus-jets-enriched categories without any tagged jet (“2 jets, 0 btags”, “3 jets, 0 btags”, and “≥ 4 jets, 0 btags”) are shown in figure 5.27. These categories serve as cross check for the W-boson-plus-jets-event modeling, but are not used for the t-channel-cross-section determination. The simulated distribution reproduces the shape of the data distribution well in all three categories, which is confirmed by the KS-test probabilities. All processes are normalized to the SM prediction, except for QCD-multijet events. The predicted process normalizations are compatible with the observed data when keeping uncertainties of the normalization in mind. However, these uncertainties are not considered when calculating the $\chi^2$-test probabilities, which results in low p-values for this particular test.

![Figure 5.27: Distributions of the BDT discriminator for events with a muon in the “2 jets, 0 btags” category (top plot), “3 jets, 0 btags” category (bottom left), or “≥ 4 jets, 0 btags” category (bottom right). All categories are used as control regions for the W-boson-plus-jets-event modeling only, but are not used for the t-channel-cross-section determination.](image)

For the sake of completeness, the BDT-discriminator distributions for events in the “2 jets, 2 btags” category are shown in figure 5.28. The event statistics are low in these categories. The normalization of simulated events is not well reproduced, resulting in low $\chi^2$-test p-values, since the compatibility tests do not take systematic uncertainties into account.
5.3 Top-Quark-Pair- and W-Boson-plus-Jets-Enriched Phase Space

Figure 5.28: Distributions of the BDT discriminator for the “2 jets, 2 btags” category for events with electron (left) or muon final states (right). Both distributions are input to the statistical inference and shown prior to the statistical inference.

The distributions for the BDT-classifier output in the $t\bar{t}$-enriched phase space are shown in figure 5.29. The $t\bar{t}$-enriched categories include events with at least three jets and at least one tagged jet. These categories are included in the statistical inference. Events with electrons (left column) and muons (right column) are separately shown. In general, the simulation reproduces the data well over the full studied phase space.

Since separate BDTs are trained for events with electrons and muons, the shape of the (predicted) discriminator distribution differs between events with electrons and muons. Small differences can also be obtained when the discriminator shape among categories with different jet multiplicity. The reason is that the object kinematics themselves are varied. Moreover, the assignment of object candidates to composite objects may be altered among categories with a different number of jets.
Figure 5.29: Distributions of the BDT discriminator in the “3 jets, ≥ 2 btags” (first row), “≥ 4 jets, 1 btag” (second row) and “≥ 4 jets, ≥ 2 btags” (third row) categories for events with electron (left) or muon final states (right). All distributions are input to the statistical inference and shown prior to the statistical inference.
5.4 Signal Region

The “2 jets, 1 btag” category is used as a representative for the signal-enriched phase space. In particular, it is demonstrated in this section that the excess of data events with muons cannot be attributed to a particular small piece of phase space, e.g. caused by a mismodeling of the detector, but that the observed events are of physics origin.

Kinematic quantities ($p_T$, $\eta$, charge) of charged leptons ($e$, $\mu$), leading jet and second-leading jet, $E_T^{\text{miss}}$, as well as the reconstructed W-boson mass are discussed first. Afterwards, the $p_T$, $\eta$, and mass distributions of the reconstructed b-tagged-top-quark and best-top-quark hypotheses are discussed. Then, all other distributions, which are used as input distributions to train the BDT classifier (cf. table 4.6), are discussed. Finally, the BDT discriminator is evaluated for events in the signal category, and the resulting distributions are discussed.

5.4.1 Basic Kinematic Quantities

Charged lepton The lepton-charge distributions for events in the “2 jets, 1 btag” signal categories are shown in figure [5.30] before any charge-asymmetry reweighting. The charge asymmetry is smaller in data than predicted by simulation. One reason might be that the charge asymmetry for W-boson-plus-jets events is lower in data than predicted by simulation, as discussed in section 5.3.1. The $t\bar{t}$-charge asymmetry is found to be compatible between data and simulation. However, the breakdown of the origin of the charge asymmetry is subject to the dedicated measurement in this analysis, and discussed in section 7.3.

Charged-lepton $p_T$ and $\eta$ distributions are shown in figure [5.31] (first and second row). The predicted electron-$p_T$ distribution agrees well with data (first row, left plot), but the observed muon-$p_T$ distribution (first row, right plot) is significantly harder than expected from the simulation. This indicates that the number of background events, mainly W-boson-plus-jets and $t\bar{t}$, is underestimated for events with muons, since these processes have a significantly harder muon-$p_T$ distribution than $t$-channel events. The charged-lepton-$\eta$ distributions are compatible between data and simulation. The observed distributions of the azimuthal angle $\phi$ for the charged leptons well agree with the simulated distributions (third row).

Figure 5.30: Distributions of the lepton charge in the “2 jets, 1 btag” signal category for events with electron (left) or muon final states (right).
Figure 5.31: Distributions of the lepton $p_T$ (first row), lepton $\eta$ (second row), and lepton $\phi$ (third row) in the “2 jets, 1 btag” signal category for events with electron (left) or muon final states (right).
5.4 Signal Region

Jets The leading-jet $p_T$ and $\eta$ distributions are shown in figure 5.32 and the second-leading-jet kinematics ($p_T$, $\eta$) are shown in figure 5.33. The shapes of the observed distributions are well compatible with the simulation within statistical precision. In particular, the low-jet-$p_T$ regions have large uncertainties due to the QCD-multijet modeling. The observed $\eta$ distributions for the leading jet and second-leading jet show a small excess in the interval $\eta \in [-2, -1]$ for events with a muon in the final state (fig. 5.32 (bottom right) and fig. 5.33 (bottom right)). However, the disagreement cannot be confirmed in other categories, i.e. neither for electron events in the signal category, nor for electron and muon events in the $t\bar{t}$-enriched and W-boson-plus-jets-enriched control regions. The KS-test probabilities for all distributions are well above zero in the signal categories.

![Figure 5.32: Distributions of the leading-jet $p_T$ (top row) and leading-jet $\eta$ (bottom row) in the "2 jets, 1 btag" signal category for events with electron (left) or muon final states (right).]
5 Background Estimation and Evaluation of Simulated Event Modeling

Figure 5.33: Distributions of the second-leading-jet $p_T$ (top row) and second-leading-jet $\eta$ (bottom row) in the “2 jets, 1 btag” signal category for events with electron (left) or muon final states (right).

Figure 5.34: Distributions of the $E_{\text{miss}}^T$ in the “2 jets, 1 btag” signal category for events with electron (left) or muon final states (right).
$E_T^{\text{miss}}$ and transverse W-boson mass The $E_T^{\text{miss}}$ distributions show a good agreement between data and simulation for events with electrons or muons in the “2 jets, 0 btags” category (fig. 5.34). The excess of data is uniformly distributed over the full $E_T^{\text{miss}}$ range.

A good agreement of the shape of the distribution is observed as well for the $M_T(W \text{ boson})$ distribution before the reconstruction of the neutrino (fig. 5.35, top row) and after the reconstruction of the neutrino (fig. 5.35, bottom row). The data excess is uniformly distributed over the full $M_T(W \text{ boson})$ range. The Jacobian peak is well described, which demonstrates that events with W bosons are selected. The left tails of the $E_T^{\text{miss}}$ and $M_T(W \text{ boson})$ distributions probe the description of QCD-multijet shape and yield, since they are enriched in QCD-multijet events. While residual uncertainties remain in the description of the shape, the QCD-multijet contributions are well normalized.

Figure 5.35: Reconstructed transverse-W-boson mass distributions in the “2 jets, 1 btag” signal category before (top row) and after (bottom row) neutrino reconstruction for events with electron (left) or muon final states (right).
5.4.2 Top-Quark Reconstruction

The distributions of the transverse momentum of the top-quark are softer in data than expected from simulation (fig. 5.36, top row), similar to the conclusions obtained from the $t\bar{t}$ enriched phase-space. A disagreement especially is visible in the tails of the distributions with large transverse momenta. A good agreement is observed for the distributions of the top-quark pseudo-rapidity (fig. 5.36, bottom row).

The observed top-quark-mass distributions are in reasonable agreement with the simulation for events with electrons (fig. 5.37, top row). For events with muons, slightly more events are observed than predicted in the tail of the distribution with large top-quark masses. The top-quark-mass distributions are expected to be significantly broader for W-boson-plus-jets events than for $t$-channel or $t\bar{t}$ events, while the distribution for $t\bar{t}$ events is between the distributions for $t$-channel and W-boson-plus-jets events. Thus, the observed distributions for events with muons indicates that events from either $t\bar{t}$ or W-boson-plus-jets production are missing in the prediction of the “2 jets, 1 btag” category. For events with electrons, the simulation well reproduces the data.

The best-top-quark-mass distributions are shown in figure 5.37 (bottom row). A good agreement between data and simulation is observed. However, the distributions are less discrim-
nent between signal and background processes than the b-tagged-top-quark-mass distributions.

Figure 5.37: Mass distributions of the b-tagged-top-quark-reconstruction hypothesis (top row) and best-top-quark-reconstruction hypothesis (bottom row) in the "2 jets, 1 btag" signal category for events with electron (left) or muon final states (right). The reconstructed best-top-quark mass is the mass reconstructed with the jet that yields a mass closest to $172.0 \text{ GeV}/c^2$. 
5.4.3 BDT-Training Variables

The $p_T$ and $\eta$ distributions of the charged leptons are shown in figure 5.31. The pseudo-rapidity distribution of the leading jet is shown in figure 5.32. The mass distributions of the b-tagged-top-quark-reconstruction and best-top-quark-reconstruction hypotheses are shown in figure 5.37. The distributions of the remaining variables, which are used as input to the training of the BDT classifier, are described in the following.

The distributions of the pseudo-rapidity of the most-forward jet are shown in figure 5.38. The observed and predicted distributions are compatible with each other. For events with muons, an excess of jets with $\eta \in [-2, -1]$ is observed. This excess cannot be confirmed in events with electrons, and is classified as statistically insignificant according to the high KS-test probability.

The dijet-mass distributions are shown in figure 5.39 (top row). The observed dijet-mass distributions are softer than expected from simulation. Since the dijet-mass distributions are softer for W-boson-plus-jets and $t\bar{t}$ events than for $t$-channel events, mainly the background processes are underestimated. For events with electrons, the observed distribution indicates altered normalizations of $t$-channel, $t\bar{t}$, and W-boson-plus-jets events, since the left tail of the distribution is well modeled, while the tail is overestimated.

Figure 5.39 also shows the angular separation ($\Delta R$) between the two jets. A reasonable agreement between data and simulation is obtained. Since signal-like events have widely separated jets, the tail of the distribution for electron events indicates that the signal contribution with electrons might be overestimated by the simulation.

Figure 5.40 shows the distributions of the sum of energies (top row) and sum of transverse energies (bottom row). These distributions indicate a possible excess of W-boson-plus-jets or $t\bar{t}$ events with altered shapes in the “2 jets, 1 btag”-muon category. The agreement between data and simulation is reasonable in all distributions. The KS-test probabilities indicate that data and simulation are statistically compatible, except for the sum of transverse energies in the muon category (bottom right). However, uncertainties of the normalization are not taken into account by the KS-test, and the sum-of-transverse-energies distribution remains compatible for the electron channel.
Figure 5.39: Distributions of the mass of the dijet system (top row) and angular separation $\Delta R$ between the two leading jets (bottom row) for events for events with electron (left) or muon final states (right) in the “2 jets, 1 btag” signal category.
Figure 5.40: Sum of energies (top row) and sum of transverse energies (bottom row) of all jets for events with electron (left) or muon final states (right) in the “2 jets, 1 btag” signal category.
The distribution of $\cos^\circ$ (W boson, leading jet), which is the cosine of the angle between the reconstructed W boson and W-boson-plus-leading-jet system, is shown for the signal region in figure 5.41. The agreement between data and simulation is good for events in the electron-“2 jets, 1 btag” category. The observed distribution in the muon category (fig. 5.41 right plot) indicates that the amount of background-like events is underestimated, since the tail of the distribution ($\cos^\circ (W \text{ boson, leading jet}) > -0.2$) essentially is free of signal events.

The observed sphericity distribution is in good agreement with the prediction (fig. 5.41, bottom row). The data excess in the muon channel is uniformly distributed.

Figure 5.41: Distributions of $\cos^\circ (W \text{ boson, leading jet})$ (top row) and sphericity (bottom row) for events with electron (left) or muon final states (right) in the “2 jets, 1 btag” signal category.

### 5.4.4 BDT-Discriminator Distribution

Figure 5.42 shows the BDT discriminator distributions for the signal “2 jets, 1 btag” (top row) and “3 jets, 1 btag” (bottom row) categories. These distributions are input to the statistical inference, as well as the discriminator distributions in the background-enriched control regions as discussed in the previous section. All distributions are shown prior to the statistical inference.

The right tails of the discriminator distributions for events in the “2 jets, 1 btag”-category are almost pure signal regions ($\text{bdt} > 0.2$). Instead, the left tail of the distribution almost exclusively...
Background Estimation and Evaluation of Simulated Event Modeling

consists of background events. Signal events with electrons are slightly overestimated by the simulation, while the overall agreement between data and simulation is good. Data and simulation are well compatible for events with muons, except for the normalization of background processes. In the muon channel, background events are underestimated, while the pure signal region indicates that the number of signal events is predicted well.

The “3 jets, 1 btag” category is dominated by $t\bar{t}$ processes, but still have a significant amount of signal events. For signal events, the third jet stems either from radiation or from the fragmentation of the spectator-$b$ quark.

The agreement of the discriminator distribution in the “3 jets, 1 btag” category is reasonable in both electron and muon distributions. A small shift in the distribution can be obtained for events with muons. However, this shift is not significant and can be explained with diversified discriminator distributions for W-boson-plus-jets, $t\bar{t}$, and signal events, as well as by the normalization of the individual processes.

Figure 5.42: Distributions of the BDT discriminator for the “2 jets, 1 btag” (top) and “3 jets, 1 btag” (bottom) categories for events with electron (left) or muon final states (right). All distributions are input to the statistical inference and shown prior to the statistical inference.
6 Sources of Systematic Uncertainties and Expected Sensitivity

This chapter starts with a discussion about the sources of systematic uncertainties that are related to the measurement of the $t$-channel cross section. All systematic uncertainties are classified into the categories “experimental uncertainties” (section 6.1) and “generator-related uncertainties” (section 6.2). The impact of each particular systematic uncertainty is elaborated in terms of rate-changing and shape-changing effects. Then, the impact that each systematic uncertainty has on the cross-section measurement is quantified (section 6.3). The expected sensitivity of the cross-section measurement and the linearity of the signal-extraction procedure are discussed in section 6.4. Finally, the characteristics for the measurement of the ratio of top-quark and top-antiquark production are described in section 6.5.

Experimental uncertainties (section 6.1) include uncertainties due to the limited statistical precision of the simulated events, detector-related uncertainties, and uncertainties due to the normalization of background processes. Experimental uncertainties are fully correlated among all simulated processes. They are also fully correlated between events with electrons and muons, except for uncertainties that involve lepton reconstruction or identification efficiencies.

The uncertainty of the BDT-discriminator distribution due to the limited statistical precision of simulated and QCD-multijet events is discussed first. Then, detector-related sources of experimental uncertainties are described. These include jet-flavor-tagging efficiencies, jet-energy resolution and scale, luminosity determination, amount of un-clustered energy, number of Pile-Up interactions, identification and reconstruction of charged leptons, and trigger efficiencies. Moreover, the charge ratio of W-boson-plus-jets production is discussed in section 6.1.10.

Uncertainties of the background normalization (section 6.1.11) can be quite large with up to 100%. One strategy would be to normalize the background processes in data-sideband regions a priori, i.e. before performing the statistical inference. However, this analysis follows another approach, in which the normalization of background processes is constrained in situ. Background-enriched categories are used to constrain background processes within the statistical inference. Correlations between systematic uncertainties are properly taken into account in this way. Mainly W-boson-plus-jets events and $t\bar{t}$ events can be constrained in dedicated categories. QCD multijet, diboson, $s$-channel-, and $tW$-channel-single-top-quark processes are only minor backgrounds.

Generator-related uncertainties (section 6.2) include the modeling of additional partons with the multi-leg-generator MADGRAPH combined with PYTHIA, uncertainties in the parametrization of the PDFs, renormalization and factorization scales, and signal modeling.
6 Sources of Systematic Uncertainties and Expected Sensitivity

6.1 Experimental Uncertainties

In this section, sources of experimental uncertainties are discussed. Experimental uncertainties are fully correlated between events with electrons and muons, and fully correlated among all simulated processes. Uncertainties due to the finite statistical precision of the simulated and QCD-multijet templates are uncorrelated among all processes, lepton flavors, and bins of the BDT-discriminator distribution. Uncertainties of lepton identification and reconstruction, as well as trigger efficiencies, are implemented as uncorrelated between events with electrons and muons, but correlated among all processes. For QCD-multijet events, only the uncertainties due to finite statistical precision of the template are considered, in addition to the normalization uncertainties as discussed in section 5.1. All experimental uncertainties are incorporated as nuisance parameters in the statistical model. Their prior-probability distributions are defined as Gaussian distributions. The central-68\% confidence interval of the prior-probability distribution, i.e. the ±1σ interval in terms of Gaussian standard deviations, corresponds to the ↓ and ↑ variations of a particular uncertainty. The −1σ variation is referred to as “↓ variation” in the following, and the variation corresponding to +1σ is referred to as “↑ variation”. The template-morphing algorithm, i.e. the interpolation between the nominal templates and templates that correspond to the ↓ and ↑ variations and the extrapolation to templates that correspond to larger variations, is discussed in section 4.6.

6.1.1 Statistical Precision of Templates for Simulated Events and QCD-Multijet Events

The number of simulated physics collisions usually is limited due to finite computing resources. Moreover, events from signal processes usually are generated with significantly higher corresponding integrated luminosities than background processes. A limited number of simulated events basically has two effects. First, the predicted number of events (of a particular process) after the full event selection is known up to the statistical uncertainty described by the Poisson distribution. Second, the number of events in each bin of a distribution are subject to statistical fluctuations. While the first effect usually is incorporated in analyses, the second effect becomes important if a template fit is performed. The result of a template fit can be biased if the templates are derived from distributions using only a limited number of events.

In this analysis, the finite statistical precision of simulated events are incorporated by using the method proposed in references [234, 235]. Uncertainties due to the finite statistical precision of the templates are uncorrelated among all bins of a distribution. The implementation of this uncertainty for the theta framework [208] is described in detail in reference [229]. The general idea is to introduce one additional, independent nuisance parameter for each bin of a simulated distribution, as well as for the QCD-multijet template. Each nuisance parameter describes the uncertainty due to Poisson fluctuations according to the number of simulated events in each bin. However, this approach becomes ineffective due to a large number of additional nuisance parameters, since all simulated processes have separate templates and are treated independently. Furthermore, the Poisson probabilities need to be adapted to event weights, which are different among all processes. Therefore, a robust, effective parametrization is used. Here, one nuisance parameter is introduced for each bin of distribution, but only for the total expected yield from simulation (plus QCD-multijet events). Each nuisance parameter is parametrized with a Gaussian distribution. The squared width of the Gaussian distribution corresponds to the sum of squared event weights of all contributing processes in a particular bin. This effective parametrization leads to an equation that can be analytically solved when maximizing the likelihood w.r.t. the contributions of finite statistical precision of the templates. A quadratic
equation that has one physical solution is obtained, i.e. the numerical computation of this uncertainty can be simplified significantly.

6.1.2 Jet-Flavor-Tagging Efficiency

Data-simulation-scale factors and corresponding uncertainties of the jet-flavor-tagging efficiency are expressed in terms of jet-$p_T$, jet-$\eta$, and flavor of the fragmenting parton. They are discussed in detail in section 3.3.6. The jet-flavor-tagging efficiencies are known up to $\approx 5\%$ for jets originating from fragmenting b quarks, $\approx 10\%$ for jets from c-quark fragmentation, and about $\approx 15\%$ for jets from remaining partons (cf. fig. 3.14). In order to study the impact of the jet-flavor-tagging-efficiency uncertainties in this analysis, the data-simulation-scale factors are varied by their corresponding uncertainties, applied for each jet according to the recipe discussed in section 3.5.4, and the full analysis is re-done subsequently.

Figure 6.1 shows the impact of jet-flavor-tagging uncertainty on the BDT-discriminator distribution. The shape dependence of the BDT discriminator on the jet-flavor-tagging efficiency is small, and only some residual shape effects due to the limited number of simulated events are obtained. The rate-only effects are of the order of a few percent.

Figure 6.2 shows the impact of the mistag uncertainty on the BDT-discriminator distribution. A “mistag” refers to a jet that stems from fragmentation of u, d, s, and g partons and that is tagged as a b-jet. Events with muons are exemplarily shown, while the effect is the same for events with electrons. The mistag efficiency has a negligible impact for $t$-channel events since the spectator jet mostly is out of the tracker acceptance (fig. 6.2, left plot). However, the uncertainty of the mistagging efficiency has at least a small impact on the shape and rate of the BDT-discriminator distribution for W-boson-plus-jets events (fig. 6.2, right plot). W-boson production in association with light-flavored jets has a large cross section. About 10% of the W-boson-plus-jets events in the signal region have mistagged jets (cf. table 4.3). The remaining events are W bosons that are produced in association with heavy-flavored jets. Uncertainties of the mistagging efficiency are assumed to be uncorrelated to uncertainties of the tagging efficiency for jets from fragmentation of b or c quarks.

The change in event yield, which is obtained when varying the jet-flavor-tagging efficiency for jets from fragmentation of b or c quarks, is summarized in the upper part of table 6.1. The jet-flavor-tagging efficiency is varied within its uncertainties at 68% CL. The main processes are listed in this table. All values (in [\%]) are calculated relative to the event yield from the nominal sample. The values obtained for $t$-channel top-quark and top-antiquark production are well compatible with each other. Event-yield changes are a bit smaller for $t\bar{t}$ events due to combinatorics, since $t\bar{t}$ events have two jets from b-quark fragmentation. Larger variations are observed for W ($\rightarrow l\nu$) + cX events since uncertainties are doubled for jets from c-quark fragmentation (w.r.t. jets from b-quark fragmentation). W-boson-plus-light-jets events only contain mistags. Hence, no variation of the event yield is expected. The obtained numbers are consistent between events with electrons and muons, which gives additional confidence in the modeling of the b-tagged-jet requirement in the trigger for events with electrons. The statistical uncertainty is of order $O(1\%)$ (depending on the category).

The lower part in table 6.1 summarizes the changes of the event yield when varying the jet-flavor-tagging efficiency for jets from fragmentation of u, d, s, g partons (mistag). The mistag uncertainty has a negligible impact for events that contain final-state-b quarks, but sizable effects for W-boson-plus-light-jets events (15-30%).
Figure 6.1: Impact of the uncertainty due to the b-tag-algorithm efficiency for jets resulting from b- or c-parton fragmentation for $t$-channel events (top row) and $t\bar{t}$ events (bottom row).

Figure 6.2: Impact of the uncertainty due to the b-tag-algorithm efficiency uncertainty for jets from fragmentation of $u$, $d$, $s$, or $g$ partons (mistag) for $t$-channel events (left) and W-boson-plus-jets events (right). BDT-discriminator distributions are shown for events with muons.
### 6.1 Experimental Uncertainties

<table>
<thead>
<tr>
<th></th>
<th>2 jets, 1 btag</th>
<th>2 jets, 2 btags</th>
<th>3 jets, 1 btag</th>
<th>3 jets, ≥ 2 btags</th>
<th>≥ 4 jets, 1 btag</th>
<th>≥ 4 jets, ≥ 2 btags</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>b and c quarks</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$e$ t-Channel (top quark)</td>
<td>-4</td>
<td>4</td>
<td>-8</td>
<td>9</td>
<td>-2</td>
<td>2</td>
</tr>
<tr>
<td>$e$ t-Channel (top antiquark)</td>
<td>-4</td>
<td>4</td>
<td>-9</td>
<td>9</td>
<td>-2</td>
<td>2</td>
</tr>
<tr>
<td>$\tau\tau$</td>
<td>-3</td>
<td>3</td>
<td>-8</td>
<td>8</td>
<td>-2</td>
<td>2</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) +$ light jets</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) + bX$</td>
<td>-4</td>
<td>4</td>
<td>-8</td>
<td>9</td>
<td>-3</td>
<td>3</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) + cX$</td>
<td>-8</td>
<td>8</td>
<td>-17</td>
<td>19</td>
<td>-8</td>
<td>8</td>
</tr>
<tr>
<td><strong>$\mu$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau\tau$</td>
<td>-3</td>
<td>3</td>
<td>-8</td>
<td>8</td>
<td>-2</td>
<td>2</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) +$ light jets</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) + bX$</td>
<td>-4</td>
<td>4</td>
<td>-9</td>
<td>9</td>
<td>-3</td>
<td>3</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) + cX$</td>
<td>-8</td>
<td>8</td>
<td>-12</td>
<td>12</td>
<td>-8</td>
<td>8</td>
</tr>
<tr>
<td><strong>u, d, s, and g partons</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$e$ t-Channel (top quark)</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$e$ t-Channel (top antiquark)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\tau\tau$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) +$ light jets</td>
<td>-15</td>
<td>15</td>
<td>-27</td>
<td>32</td>
<td>-14</td>
<td>15</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) + bX$</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) + cX$</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$\mu$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau\tau$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) +$ light jets</td>
<td>-14</td>
<td>14</td>
<td>-27</td>
<td>32</td>
<td>-14</td>
<td>16</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) + bX$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$W (\rightarrow \ell \nu) + cX$</td>
<td>0</td>
<td>0</td>
<td>-10</td>
<td>10</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6.1: Change of the event yield (in [%]) when varying the jet-flavor-tagging efficiency for jets from fragmentation of b or c quarks (upper part) as well as for jets originating from u, d, s, or g partons (lower part). The jet-flavor-tagging efficiency is varied within its uncertainties at $68\%$ CL.
6.1.3 Jet-Transverse-Momentum Resolution

The residual correction of the simulated jet-transverse-momentum resolution and its corresponding uncertainties are discussed in section 3.3.5. The correction factors are jet-\(\eta\) dependent. The uncertainties of the jet-\(p_T\) resolution are about 5-10\% for central jets and 15\% for jets with \(|\eta| > 2.3\). Figure 6.3 shows the resulting BDT-discriminator distribution if the jet-\(p_T\) resolution is varied at a level that corresponds to one standard deviation. Here, events with electrons in the “2 jets, 1 btag” category for \(t\)-channel events (fig. 6.3, left) and for the sum of all simulated background events (fig. 6.3, right) are shown. The uncertainty of the jet-\(p_T\) resolution has negligible effects on the shape of the BDT-discriminator distributions. Furthermore, it does not significantly change the rate of signal or background events. The impact of this uncertainty is of the same order for events with muons and electrons.

![Figure 6.3: Impact of the jet-energy-resolution uncertainty for \(t\)-channel events (left) and for the sum of all simulated background events (right). Events with electrons in the “2 jets, 1 btag” category are shown.](image)

6.1.4 Jet-Energy Scale

The jet-energy-scale uncertainties are expressed as a function of jet-\(p_T\) and jet-\(\eta\) and are about 1-3\% in the central-detector region. They are larger (up to \(\approx 5\%\)) for jets at the barrel-endcap-transition region and for jets close to the beamline. The jet-energy scale and their uncertainties are discussed in detail in section 3.3.5. \(E_{\text{miss}}\) is accordingly re-calculated when applying variations to the jet-energy scale (cf. sec. 6.1.6). Figure 6.4 shows the resulting BDT-discriminator distributions for \(t\)-channel events (top row) and the sum of all simulated background events (bottom row) when varying the jet-energy scale by its uncertainties. The rate-changing effect is small on average, but can be as large as \(\approx 10\%\) in single bins.

Table 6.2 summarizes the impact of the jet-energy-scale variations on the event yield for the main processes. A variation of the jet-energy scale at 68\% CL changes the amount of predicted events by about \(\pm 5\%\) for events with up to three jets, and about \(\pm 10\%\) for events with at least four jets. The rate changes are small for \(t\)-channel events, but larger for \(t\bar{t}\) events. Events migrate among jet-multiplicity bins according to the jet-\(p_T\) spectrum and jet-multiplicity distribution of the individual process. In particular, variations of the jet-energy scale have opposite effects on the event yield for \(t\)-channel and \(t\bar{t}\) events. The reasons is that jet-multiplicity distribution peaks at four jets for \(t\bar{t}\) events. Instead, most of the \(t\)-channel events have two jets...
within the acceptance. The obtained rate changes are compatible for events with electrons and muons. Residual differences can be explained due to the different $p_T$ thresholds for electrons and muons, as well as the jet-$p_T$ requirement within the electron-cross trigger.

Figure 6.4: Impact of the jet-energy-scale uncertainty for $t$-channel events (top row) and on the sum of all simulated background events (bottom row). Shown are the BDT-discriminator distributions for events with electron (left) and muon final states (right).
### 6 Sources of Systematic Uncertainties and Expected Sensitivity

<table>
<thead>
<tr>
<th></th>
<th>2 jets, 1 btag</th>
<th>2 jets, 2 btags</th>
<th>3 jets, 1 btag</th>
<th>≥ 2 jets, ≥ 4 jets, 1 btag</th>
<th>≥ 2 jets, ≥ 4 jets, ≥ 2 btags</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Channel</th>
<th>2 jets, 1 btag</th>
<th>2 jets, 2 btags</th>
<th>3 jets, 1 btag</th>
<th>≥ 2 jets, ≥ 4 jets, 1 btag</th>
<th>≥ 2 jets, ≥ 4 jets, ≥ 2 btags</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>-2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>-6</td>
</tr>
<tr>
<td>t-channel (top quark)</td>
<td>-2</td>
<td>1</td>
<td>4</td>
<td>-7</td>
<td>-6</td>
</tr>
<tr>
<td>t-channel (top antiquark)</td>
<td>5</td>
<td>-6</td>
<td>4</td>
<td>-6</td>
<td>2</td>
</tr>
<tr>
<td>t tubing</td>
<td>2</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>-1</td>
</tr>
<tr>
<td>W (→ lν) + light jets</td>
<td>-2</td>
<td>2</td>
<td>1</td>
<td>-2</td>
<td>-3</td>
</tr>
<tr>
<td>W (→ lν) + bX</td>
<td>-4</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>-5</td>
</tr>
<tr>
<td>W (→ lν) + cX</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>-9</td>
<td>-5</td>
</tr>
<tr>
<td>W (→ lν) + light jets</td>
<td>-2</td>
<td>1</td>
<td>-3</td>
<td>1</td>
<td>-8</td>
</tr>
<tr>
<td>W (→ lν) + bX</td>
<td>-3</td>
<td>2</td>
<td>0</td>
<td>-1</td>
<td>-8</td>
</tr>
<tr>
<td>W (→ lν) + cX</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Channel</th>
<th>2 jets, 1 btag</th>
<th>2 jets, 2 btags</th>
<th>3 jets, 1 btag</th>
<th>≥ 2 jets, ≥ 4 jets, 1 btag</th>
<th>≥ 2 jets, ≥ 4 jets, ≥ 2 btags</th>
</tr>
</thead>
<tbody>
<tr>
<td>μ</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>-9</td>
<td>-5</td>
</tr>
<tr>
<td>t-channel (top quark)</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>-2</td>
<td>-6</td>
</tr>
<tr>
<td>t-channel (top antiquark)</td>
<td>7</td>
<td>-7</td>
<td>5</td>
<td>-7</td>
<td>3</td>
</tr>
<tr>
<td>t tubing</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-8</td>
</tr>
<tr>
<td>W (→ lν) + light jets</td>
<td>-2</td>
<td>1</td>
<td>-3</td>
<td>1</td>
<td>-8</td>
</tr>
<tr>
<td>W (→ lν) + bX</td>
<td>-3</td>
<td>2</td>
<td>0</td>
<td>-1</td>
<td>-8</td>
</tr>
<tr>
<td>W (→ lν) + cX</td>
<td>-3</td>
<td>2</td>
<td>0</td>
<td>-1</td>
<td>-8</td>
</tr>
</tbody>
</table>

Table 6.2: Change of the event yield (in [%]) when varying the jet-energy scale within its uncertainties at 68% CL. Entries are marked with a “-” if the statistical precision of the simulation is too low to be conclusive.
6.1 Experimental Uncertainties

6.1.5 Luminosity Determination

The uncertainty of the luminosity determination is 2.2\% [183] (cf. sec. 3.4). The integrated luminosity of the analyzed data is fully correlated between events with electrons and muons, and fully correlated among all simulated processes.

6.1.6 Missing Transverse Energy

Uncertainties of the $\vec{p}_T^{\text{miss}}$ reconstruction in simulated events are considered as follows. The $\vec{p}_T^{\text{miss}}$ is factorized into its basic constituents, namely jets, unclustered energy, electrons, and muons. Tau leptons and photons are clustered into jets as discussed in section 3.3.5. Each constituent is independently varied by its respective response and resolution uncertainties, and $\vec{p}_T^{\text{miss}}$ is re-calculated accordingly. While the underlying assumption is that the energy-scale uncertainties are uncorrelated among different physics-object types, they are taken fully correlated among the same type of physics objects.

Jet-energy-scale uncertainties are simultaneously applied to jets and $\vec{p}_T^{\text{miss}}$. Uncertainties of the transverse-momentum response and resolution of electrons [169, 236] and muons [164] between simulation and data are neglected as they are much smaller than for jets or unclustered energy.

Unclustered energy refers to the part of $\vec{p}_T^{\text{miss}}$ that remains after vectorially subtracting all jets with uncorrected transverse momentum $p_T > 10\text{GeV}/c$ within $|\eta| < 5$, as well as muons and electrons with $p_T > 10\text{GeV}/c$ within $|\eta| < 2.5$ if they are not clustered in jets. The uncertainty of the amount of unclustered energy is estimated to be $\pm 10\%$ [68, 237]. The impact of unclustered-energy variations on the BDT-discriminator shape is small (cf. fig. 6.5). Furthermore, the variations of the unclustered energy only marginally affect the predicted event yield for all simulated processes in each category.

6.1.7 Pile-Up

The number of additional interactions in data is estimated within an uncertainty of $\pm 5\%$ (cf. the discussion in section 3.5.4). This uncertainty covers systematic shifts on the total inelastic proton-proton cross section as well as uncertainties due to the reweighting procedure. Furthermore, this uncertainty is expected to cover uncertainties due to the modeling of soft interactions with the PYTHIA 6 generator, e.g. particle energy spectra, multiplicities, or their angular distributions. Additional interactions potentially affect the jet kinematics and charged-lepton isolation. Both effects are damped by dedicated algorithms. The CHS technique of the PF algorithm subtracts charged constituents from jets that are incompatible with the primary vertex (cf. sec. 3.3.5). Charged particles are further excluded from the isolation calculation for muons (cf. 3.3.3).

A negligible effect on the BDT-discriminator distribution is obtained due to variations of the number of additional interactions (fig. 6.6). Large effects are noted in a few bins in the tails of the discriminator distributions. However, only few events remain in these bins, and the effects can be attributed to the statistical precision of the simulated template. The rate-changing effect is small in each category.
Figure 6.5: Impact of the $E_{T}^{miss}$-scale uncertainty for $t$-channel events (top row) and the sum of all simulated background events (bottom row). Shown are the BDT-discriminator distributions for events with electron (left) and muon final states (right).
Figure 6.6: Impact of the number of additional interactions (Pile-Up) on the BDT-discriminator distribution for $t$-channel events (top row) and for the sum of all simulated background events (bottom row).
6 Sources of Systematic Uncertainties and Expected Sensitivity

6.1.8 Identification, Reconstruction, and Triggering of Charged Leptons

The uncertainties of charged-lepton identification, reconstruction, and triggering are discussed in detail for electrons in section 3.3.4 and for muons in section 3.3.3. Identification, reconstruction, and isolation efficiency have an uncertainty of 3% for each electrons and muons. The uncertainty of the efficiency estimation for triggering charged leptons is below 1%.

6.1.9 Hadronic-Trigger Efficiencies

The uncertainties of the trigger efficiency of b-tagged jets within the electron-cross trigger are discussed in section 3.5.4. They are dominated by the uncertainty of the additional scale-factor, which is ±5%. The uncertainties due to the parametrization of the Gompertz curve are much smaller.

6.1.10 Charge Ratio for W-Boson-plus-Jets Events

The observed muon-charge ratio is smaller than predicted by simulation as discussed in section 5.3.1. Figure 6.7 shows the ratio of $W^+$-boson production with positively-charged muons and $W^-$-boson production with negatively-charged muons (“muon-charge ratio”) as a function of the absolute pseudo-rapidity of the muon. The left plot shows the absolute muon-charge-ratio distributions for data and simulation. The right plot shows the ratio of observed and predicted muon-charge-production rates. The difference is stable over the full pseudo-rapidity range up to ≈ 2%.

The relatively small $\eta$ dependence of the muon-charge ratio can become a sizable effect only if the discriminator distributions significantly differ between events with negatively-charged muons and positively-charged muons. Figure 6.8 shows a comparison of the BDT discriminator-output distributions for $W^+$-boson production with positively-charged muons ($\mu^+$) and $W^-$-boson production with negatively-charged muons ($\mu^-$). The distributions are
shown for data events in the “2 jets, 0 btags” category, which is dominated by W-boson-plus-jets production. The discriminator distributions are shown for central muons ($|\eta(\mu)| < 0.9$) (left) and non-central muons ($|\eta(\mu)| \geq 0.9$). The pull plot shows the differences w.r.t. the “nominal” event selection, i.e. a combination of $W^+$-boson production to $W^-$-boson production according to the ratio as obtained from simulation (top row) or data (bottom row).

While the $\eta$-dependent charge ratio is estimated up to an uncertainty of $\approx 2\%$, figure 6.8 shows the two extreme variations “only $W^+$ bosons” or “only $W^-$ bosons”. The resulting differences between the discriminator distributions for events with negatively-charged muons and positively-charged muons are rather small. They are even smaller for central muons. Furthermore, it is obtained that the observed discriminator distribution for an inclusive muon charge is well described by the simulation (cf. fig. 5.27).

For the charge-ratio measurement, events with muons in the “2 jets, 0 btags” category are included in the statistical inference. These events constrain the charge-production ratio of W-boson-plus-jets events in situ. The W-boson-plus-jets yield for events with $W^+$ or $W^-$ bosons is allowed to independently float with an a-priori uncertainty of 30%.

It is assumed that the production rates of events with $W^+$ and $W^-$ bosons are fully correlated for events with electrons and muons, as well as independent of the number and flavor of additional jets. Moreover, the charge-production ratios for the processes $W (\rightarrow l\nu) + bX$, $W (\rightarrow l\nu) + cX$, and $W (\rightarrow l\nu) + \text{light jets}$ are allowed to differ among each other.

Furthermore, $s$-channel single-top-quark processes are lepton-charge asymmetric, too. $s$-channel production is expected to give negligible contributions due to its small production cross section. $s$-channel top-quark production and top-antiquark production are completely decoupled in the statistical model. QCD-multijet events are expected to be produced lepton-charge symmetric.
6 Sources of Systematic Uncertainties and Expected Sensitivity

6.11 Uncertainties of Background Normalization

All signal and background processes are normalized to the SM prediction (cf. table 3.4), except for QCD-multijet events whose normalization is estimated with observed data. In this section, the uncertainty of the background normalization is discussed.

Theoretical calculations usually include uncertainties due to the choice of the renormalization and factorization scales and parametrizations of the PDFs (cf. table 3.4). The uncertainty of the PDF parametrization usually is estimated by varying the parametrization within a set that corresponds to one particular fitting group. However, these variations might underestimate the variations that are obtained when comparing PDF sets of different fitting groups. Even if those effects can be small for calculations of the inclusive production cross section of a particular process, such variations can have a sizable effect on the acceptance in the dedicated phase space in which this analysis is performed [55, 238]. Therefore, the uncertainties of the background normalization are increased compared to the uncertainties of theoretical calculations.

Table 6.3 summarizes the assumed uncertainties of the background normalization. These uncertainties are included into the statistical model by adding nuisance parameters with log-normal prior-probability distributions. Log-normal distributions are preferred over Gaussian distributions, since Gaussian distributions possibly lead to a bias in the measurement for large uncertainties as they become truncated at zero. For $t$-channel signal events, a flat prior is used. The uncertainty of other single-top-quark processes, $s$-channel and $tW$-channel, is estimated to be ±30%, which is comparable to the experimental resolution of the $tW$-channel-cross-section measurement [94, 95].

The uncertainty of the $tt$ production is estimated to be ±15% (based on the NLO calculation with MCFM [76, 77], cf. table 3.4). This uncertainty also covers calculations in approximate NNLO accuracy (cf. the overview in ref. [43]) and NNLO accuracy [74]. Moreover, this uncertainty covers the measurements of the inclusive $tt$-production-cross-section measurements obtained by the ATLAS and CMS experiments [86, 87, 239–241].

Event generation of $W$ bosons that are produced in association with jets from b or c quarks involves large uncertainties. A typical source of systematic uncertainties in the production of $W$-boson-plus-jets events is the amount of gluon radiation and the fraction of gluon splitting into pairs of b or c quarks. Therefore, $W$-boson-plus-jets processes are grouped into three categories, depending on the flavor of the jets (b, c, or light) that are produced in association with the $W$ boson (cf. sec. 3.5.2). Furthermore, the production of $W$-boson-plus-jets events is sensitive to variations of the $Q^2$ scale with yield changes up to 100% (as discussed in section 6.2.3). Here, the yield of $W$-boson-plus-jets events is independently varied in categories of two, three, or at least four jets.

Thus, several nuisance parameters are introduced for the $W$-boson-plus-jets processes. In total, nine nuisance parameters ($3_{\text{flavor}} \otimes 3_{Q^2-\text{scale}}$) are used to describe the uncertainties of $W$-boson-plus-jets-event normalization. $W$-boson production in association with light-flavored jets is assumed to be known with an uncertainty of ±50%. $W$-boson production in association with jets from b or c quark fragmentation is assumed to be known with an uncertainty of ±100%.

The uncertainty of the predicted event yield of $Z$-boson-plus-jets events is correlated to the uncertainty of the predicted number of $W$-boson-plus-jets events in a certain category. Here, an uncertainty of 30% for $Z$-boson-plus-jets events is used in addition to the uncertainty of $W$-boson-plus-jets prediction.
The uncertainty of diboson production in association with jets is estimated to be known within 30%. Diboson events are a minor contribution among all background processes. The uncertainties of the estimated yield for QCD-multijet events are discussed in section 5.1.

<table>
<thead>
<tr>
<th>Process</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single top quark</td>
<td></td>
</tr>
<tr>
<td>s-channel</td>
<td>30%</td>
</tr>
<tr>
<td>tW-channel</td>
<td>30%</td>
</tr>
<tr>
<td>t(\bar{t})</td>
<td>15%</td>
</tr>
<tr>
<td>W boson plus jets</td>
<td></td>
</tr>
<tr>
<td>2 jets</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + light jets</td>
<td>50%</td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + bX</td>
<td>100%</td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + cX</td>
<td>100%</td>
</tr>
<tr>
<td>3 jets</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + light jets</td>
<td>50%</td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + bX</td>
<td>100%</td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + cX</td>
<td>100%</td>
</tr>
<tr>
<td>(\geq 4) jets</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + light jets</td>
<td>50%</td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + bX</td>
<td>100%</td>
</tr>
<tr>
<td>(W \rightarrow l\nu) + cX</td>
<td>100%</td>
</tr>
<tr>
<td>(Z/\gamma^* \rightarrow l^+l^-) + jets</td>
<td>30% (\oplus) unc. on W-boson-plus-jets prediction</td>
</tr>
<tr>
<td>Diboson</td>
<td></td>
</tr>
<tr>
<td>WW</td>
<td>30%</td>
</tr>
<tr>
<td>WZ</td>
<td>30%</td>
</tr>
<tr>
<td>ZZ</td>
<td>30%</td>
</tr>
<tr>
<td>QCD multijet</td>
<td>39-120%, category dependent</td>
</tr>
</tbody>
</table>

Table 6.3: Uncertainties of the normalization of background processes in the dedicated phase space in which this analysis is performed.
6.2 Generator-related Uncertainties

In the following section, systematic uncertainties that are related to event generation are discussed. These include the modeling of additional partons for $t\bar{t}$ and W-boson-plus-jets events, the uncertainty in the parametrization of the PDFs, the choice of the renormalization and factorization scales ($Q^2$ scale), and $t$-channel-event modeling. Variations to the $Q^2$ scale and the signal modeling are the most important among all generator-related uncertainties.

6.2.1 Modeling of Additional Partons

Simulated $t\bar{t}$ and W-boson-plus-jets events are generated with MadGraph +Pythia. Additional jets are obtained either from the matrix-element calculation or from the parton shower, and the corresponding events are combined according to the MLM-matching prescription with $k_T$ jets [66]. Two main parameters, matching threshold and matching scale, define the matching procedure. The matching threshold is varied by factors 0.5 and 2, and the matching scale is re-calculated accordingly (cf. also sec. 3.5.2). Both parameters are independently calculated for $t\bar{t}$ and W-boson-plus-jets process.

The obtained variations of the BDT-discriminator distribution are shown in figure 6.9 for $t\bar{t}$ (top row) and W-boson-plus-jets events (bottom row). Fluctuations in single bins can be large due to the statistical precision of the simulated samples, e.g. only a few events with large weights are expected in bins of the tails of the discriminator distributions. The overall variations are not significant.

Figure 6.10 shows the obtained jet-multiplicity distributions when varying the matching parameters for both $t\bar{t}$ (left plot) and W-boson-plus-jets events (right plot). Events with muons are exemplarily shown.

The event yield of $t\bar{t}$ events is stable against variations of the matching scale. The event yield only changes by few percent for $t\bar{t}$ events with up to four jets. The impact of the matching scale is more significant for $t\bar{t}$ events with more than five jets, with up to 10% for events with seven jets. However, only a small fraction of all produced $t\bar{t}$ events have more than five jets.

The impact of the matching variations are also relatively small for W-boson-plus-jets events. The event-yield changes by about 10% for events with up to four jets, and by about 20% for events with more than four jets. In particular, these changes are small compared to (and fully covered by) the assumed uncertainties of the event yield (50-100%) as discussed in one of the following paragraphs (sec 6.1.11). Therefore, the matching uncertainty is not explicitly modeled for W-boson-plus-jets events in the following.
6.2 Generator-related Uncertainties

Figure 6.9: Impact of varied matching parameters on the BDT-discriminator distribution for $t\bar{t}$ events (top row) and W-boson-plus-jets events (bottom row).

Figure 6.10: Impact of varied matching parameters on the jet-multiplicity distribution for $t\bar{t}$ events (left) and W-boson-plus-jets events (right) with muons.
6 Sources of Systematic Uncertainties and Expected Sensitivity

6.2.2 Parton-Distribution Functions

Simulated events are generated with the “central” (best-fit) PDF set of the CTEQ6M NLO PDF set [52]. In order to account for uncertainties in the parametrizations of the best-fit PDF set, the simulated signal sample is reweighted to each of the 40 eigenvector sets of the CTEQ6M PDF set. Events are reweighted using the LHAPDF package [242].

The POWHEG BOX event generators sets the scale to the transverse momentum of the hardest emission such that the PYTHIA showering is only used for subsequent radiations with lower transverse momenta. For the PDF reweighting, the event scale is re-calculated from generator information such that it equals the top-quark mass. The simulated events are properly re-normalized such that the overall normalization of the (inclusive) production cross section is constant for \( t \)-channel events. Instead, the acceptance changes due to altered event kinematics. Furthermore, the shape of the BDT discriminator distribution can be distorted.

The signal-extraction procedure is re-iterated 40 times by generating pseudo-signal events according to each of the altered PDFs but leaving the background contribution unchanged. The \( t \)-channel-cross-section measurement, instead, is evaluated with the nominal templates. The signal cross section is determined for each pseudo-experiment. The shift between the nominal scenario and a certain PDF parametrization is obtained for each of the pseudo-experiments. Half of the 40 altered PDF sets represent variations in the “+” direction or “−” direction respectively. For each direction “+” and “−” separately, the obtained shifts are added quadratically. The square roots of the obtained values are taken as a contribution to the systematic uncertainty due to the PDF set parametrization in either “+” or “−” directions. Since the CTEQ collaboration provides PDF variations at 90% CL, the obtained shifts are divided by 1.64485 such that the resulting uncertainty corresponds to a variation at 68% CL [55, 243].

However, these variations are expected not to cover deviations that are observed among PDF sets of different fitting groups [55]. Therefore, this analysis follows the recommendations given in reference [55]. A second contribution to the systematic uncertainty due to the parametrization of the PDFs is calculated by reweighting the signal events from the CTEQ6M PDF set to the mean values of all 100 replica of the NNPDF2.1 NLO PDF set [244], as well as to the best fit of the MSTW2008 NLO PDF set [75]. Moreover, a newer PDF set of the CTEQ collaboration (CTEQ10 NLO PDF set [51]) is checked.

Reweighted events are properly normalized such that the inclusive production cross section remains constant. Pseudo-signal events are generated according to the altered PDFs and the signal cross section is measured in altered pseudo events using the nominal templates derived with the CTEQ6M PDF set. The resulting mean shifts of the measured \( t \)-channel cross section are +0.8% for NNPDF2.1 and −0.6% for MSTW2008 PDF sets. Changing the PDF set to the best-fit CTEQ10 NLO PDF set results in a negligible shift of the cross-section estimate.

Table 6.4 summarizes the estimated impact on the signal-cross-section measurement due to the choice of the PDF set. A small dependence is observed for the inclusive-cross-section measurement.

1Technically, “+” and “−” variations are stored in alternating order and identified by using the index of the PDF member.

2These recommendations are adapted to this analysis by using a newer version of the NNPDF set.
6.2 Generator-related Uncertainties

<table>
<thead>
<tr>
<th>PDF set</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 Eigenvectors of CTEQ6M NLO [52]</td>
<td>±1.3%</td>
</tr>
<tr>
<td>Mean PDF set of NNPDF2.1 NLO [244]</td>
<td>+0.8%</td>
</tr>
<tr>
<td>Best-fit PDF set of MSTW2008 NLO [75]</td>
<td>−0.6%</td>
</tr>
<tr>
<td>Best-fit PDF set of CTEQ10 NLO [51]</td>
<td>+0.01%</td>
</tr>
<tr>
<td>Total</td>
<td>±1.5%</td>
</tr>
</tbody>
</table>

Table 6.4: Estimated impact on the signal-cross-section measurement due to the choice of the PDFs.

6.2.3 Renormalization and Factorization Scales

Factorization scale $\mu_F$ and renormalization scale $\mu_R$ are simultaneously varied at a common scale $Q^2 = \mu_F^2 = \mu_R^2$. The definition of $Q^2$-scale variations is discussed in more detail in section 3.5.2. In particular, the scales are changed in both the matrix-element calculation and parton-shower modeling. The $Q^2$ scale is varied by factors $(0.5)^2$ and $(2)^2$ w.r.t. the nominal values. Matching scale (sec. 6.2.1) and $Q^2$ scale are varied independently. Furthermore, the $Q^2$ scale is assumed to be uncorrelated among $t\bar{t}$, W-boson-plus-jets, and single-top-quark processes. Thus, three independent nuisance parameters are incorporated in the statistical model.

The $Q^2$ scale has an impact on the radiation of additional jets. Therefore, variations to the $Q^2$ scale shift the jet-multiplicity distribution and vary the kinematics of physics objects in an event, e.g. the jet-$p_T$ spectra among others. However, they do not affect the number of b-tagged jets. In the following paragraphs, the impact of the $Q^2$-scale variations on the jet-multiplicity distribution and BDT-discriminator distributions are discussed.

Jet-Multiplicity Distribution Figure 6.11 shows the $Q^2$-scale variations for $t$-channel events with an inclusive number of b-tagged jets. Halving the $Q^2$ scale decreases the expected number of events with two jets by a few percent. The number of events with three jets remains stable for events with muons, but decreases for events with electrons, which is attributed to the limited statistical precision of the simulation. $t$-channel events migrate to higher jet-multicities with four jets when decreasing the $Q^2$ scale. Here, the effect can be sizable with up to 10%. A $Q^2$ scale that is twice the nominal value enhances the number of events with two jets, but decreases the number of events with three or four jets. The jet multiplicity remains relatively stable against variations of the $Q^2$ scale, and the obtained shifts are consistent between events with electrons and muons. The $Q^2$-scale variations are approximately symmetric around events with three jets.

The jet-multiplicity distributions of $t\bar{t}$ and W-boson-plus-jets events are much more sensitive to the choice of the $Q^2$ scale as shown in figure 6.12. Halving the $Q^2$ scale results in a significant shift of $t\bar{t}$ events with less than four jets to higher jet-multicities. Instead, events with more than four jets migrate to lower jet-multiplicity bins when doubling the $Q^2$ scale. The turning point at events with four jets is expected from the $t\bar{t}$-event signature with semi-leptonically decaying W bosons.

The number of jets which are produced in association with W bosons is even more sensitive to the $Q^2$ scale (fig. 6.12, bottom row). Halving the $Q^2$ scale increases the number of W-boson-plus-jets events with at least two jets by more than 100%. Doubling the $Q^2$ scale decreases the number of events by about 30-50%. The $Q^2$-scale variations have a reverse effect on the
6 Sources of Systematic Uncertainties and Expected Sensitivity

jet-multiplicity distribution for W-boson-plus-jets production compared to $t\bar{t}$ production. The reason is that the jet-multiplicity distribution is steeply falling for W-boson-plus-jets production, while the $t\bar{t}$-event signature involves at least four jets.

The amount of variations that are applied to the $Q^2$ scale, i.e. doubling or halving the scale, are kind of an “educated guess” that is used to estimate the importance of the choice of the renormalization and factorization scales. As such huge effects on the predicted event yield are obtained, it is questionable whether the event generator reasonably describes the variations of W-boson-plus-jets events w.r.t. the $Q^2$ scale. While the $Q^2$ scale has a small impact on the kinematics of W-boson-plus-jets events, the shift of the jet-multiplicity distribution for W-boson-plus-jets events is questioned in particular.

Figure 6.13 shows the obtained jet-multiplicity distribution in comparison to the simulated prediction with nominal $Q^2$ scale. The pull plot shows the yield changes when varying the renormalization and factorization scales for events with muons w.r.t. the nominal $Q^2$ scale. The considered $Q^2$-scale variations are significantly disfavored by the obtained data.

Therefore, the $Q^2$-scale uncertainty is factorized into two components in order to disentangle the overall yield normalization from shape effects on jet-multiplicity or BDT-discriminator distributions. A shape-changing component alters the jet-multiplicity distributions, as well as the kinematics of the W-boson-plus-jets events, but preserves the normalization of W-boson-plus-jets events. The rate-changing component alters the event yield in each jet-multiplicity bin separately. The yield of W-boson-plus-jets events is independently varied in categories of two, three, or at least four jets by 50-100% (cf. sec. 6.1.11).
6.2 Generator-related Uncertainties

Figure 6.12: Effect on the jet-multiplicity distribution of $t\bar{t}$ events (top row) and W-boson-plus-jets events (bottom row) due to variations of the renormalization and factorization scales.

Figure 6.13: Comparison of predicted and observed jet-multiplicity distributions for the categories enriched in W-boson-plus-jets production. The pull plot shows the variations of the renormalization and factorization scales w.r.t. the nominal $Q^2$ scale.
The impact of the $Q^2$-scale variations on the BDT-discriminator distribution is exemplarily shown for $t$-channel events in the “2 jets, 1 btag” category in figure 6.14 (top row). The effect on the BDT-discriminator shape can be as large as 10% in isolated bins. However, the observed fluctuations can be explained by the finite statistical precision of the simulation. Especially the left tail of the distribution ($bdt < -0.7$) is dominated by single events.

Figure 6.14: Influence of variations to the renormalization and factorization scales on the BDT-discriminator distribution for $t$-channel events (top row), $t\bar{t}$ events (second row), and W-boson-plus-jets events (third row) in the “2 jets, 1 btag” category.
Furthermore, figure 6.14 shows the BDT-discriminator distributions with varied $Q^2$ scale for $t\bar{t}$ (second row) and W-boson-plus-jets production (third row). Shape effects on the BDT-discriminator distribution are larger for these processes than for signal events. However, also the number of generated events corresponds to a lower integrated luminosity. The templates for the W-boson-plus-jets events show significant fluctuations due to large event weights.

### 6.2.4 Signal Modeling

The nominal $t$-channel-signal events are generated in NLO accuracy with the POWHEG BOX-event generator in the 5-flavor scheme. Here, massless initial-state- $b$-quarks are used in the matrix-element calculation. In order to estimate the effect that the event generator has on the signal modeling, the nominal BDT-discriminator distributions are compared with the distributions obtained with the COMPHEP generator. The COMPHEP generator implements a matching of events which are generated according to $2 \rightarrow 2$ or $2 \rightarrow 3$ diagrams. Therefore, COMPHEP generates events in an “effective-NLO accuracy”. Furthermore, COMPHEP generates events with initial-state- $b$-quarks that are massive. However, the kinematics of all particles are described in LO+PS accuracy.

In particular, the comparison with COMPHEP is expected to give a reasonable estimate of the impact that the modeling of the spectator- $b$-quark has on the $t$-channel-cross-section measurement, since POWHEG BOX generates events with massless initial-state- $b$-quarks, while COMPHEP takes the b-quark mass into account. Both generators model the kinematics of the spectator- $b$-quark in LO+PS accuracy at the highest possible order. An overview of all available event generators for $t$-channel processes is given in section 3.5.1.

In figure 6.15, the BDT-discriminator distributions as obtained with the POWHEG BOX generator are compared with those from the COMPHEP generator. All distributions are shown for events in the “2 jets, 1 btag” category. Two effects are obtained. First, the predicted BDT-discriminator distributions that are obtained with the POWHEG BOX event generator are much more signal-like than the distributions obtained with the COMPHEP generator. Here, “signal-like” refers to high discriminator values. Differences in the predicted kinematics between both generators are expected, since POWHEG BOX generates events in NLO while COMPHEP generates event kinematics in LO accuracy for all particles. Therefore, differences (to some extent) are expected in the shape of the BDT-discriminator distribution.

Second, the predicted event yield depends on the event generator. COMPHEP generator predicts 1.6% more events with muons in the “2 jets, 1 btag” category than the POWHEG BOX generator, and 5.3% more events with electrons. Larger discrepancies are observed for events with more than two jets, but are of minor importance as these categories are dominated by $t\bar{t}$ events. The obtained differences can be only partially attributed to the modeling of the spectator-$b$ quark. Furthermore, the obtained differences are smaller for events with muons than for events with electrons, which is not expected a priori. Table 6.5 summarizes the predicted event yields for each generator in each category.
Figure 6.15: Comparison of the $t$-channel signal modeling with both POWHEG BOX and COMPHEP generators. Shown are the BDT-discriminator distributions for events with electron (left) and muon final states (right).

<table>
<thead>
<tr>
<th></th>
<th>2 jets, 1 btag</th>
<th>2 jets, 2 btags</th>
<th>3 jets, 1 btag</th>
<th>≥ 3 jets, ≥ 2 btags</th>
<th>≥ 4 jets, 1 btag</th>
<th>≥ 4 jets, ≥ 2 btags</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$e$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t$-channel POWHEG BOX</td>
<td>414.8</td>
<td>16.0</td>
<td>172.1</td>
<td>37.7</td>
<td>60.5</td>
<td>18.7</td>
</tr>
<tr>
<td>$t$-channel COMPHEP</td>
<td>436.8</td>
<td>14.4</td>
<td>193.7</td>
<td>41.1</td>
<td>58.6</td>
<td>16.6</td>
</tr>
<tr>
<td>Ratio ($\frac{N_{\text{COMPHEP}} - N_{\text{POWHEG BOX}}}{N_{\text{POWHEG BOX}}}$)</td>
<td>+5.3%</td>
<td>-9.7%</td>
<td>+12.5%</td>
<td>+9.0%</td>
<td>-3.1%</td>
<td>-11.1%</td>
</tr>
<tr>
<td><strong>$\mu$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t$-channel POWHEG BOX</td>
<td>690.2</td>
<td>20.2</td>
<td>258.6</td>
<td>53.3</td>
<td>82.7</td>
<td>22.9</td>
</tr>
<tr>
<td>$t$-channel COMPHEP</td>
<td>701.3</td>
<td>19.4</td>
<td>280.9</td>
<td>52.8</td>
<td>80.2</td>
<td>22.0</td>
</tr>
<tr>
<td>Ratio ($\frac{N_{\text{COMPHEP}} - N_{\text{POWHEG BOX}}}{N_{\text{POWHEG BOX}}}$)</td>
<td>+1.6%</td>
<td>-3.9%</td>
<td>+8.6%</td>
<td>-0.9%</td>
<td>-3.0%</td>
<td>-3.9%</td>
</tr>
</tbody>
</table>

Table 6.5: Comparison of expected event yields obtained with POWHEG BOX and COMPHEP event generators after the full event selection. Expected event yields are scaled to integrated luminosities of 1560.7 pb$^{-1}$ for events with electrons and 1165.6 pb$^{-1}$ for events with muons.
6.3 Impact of Systematic Uncertainties on the Measurement

In the first part of this section, the impact of experimental uncertainties on the cross-section measurement is discussed. Experimental uncertainties include uncertainties related to the description of the detector performance, background normalization, luminosity determination, and statistical precision of the simulated and QCD-multijet templates. These uncertainties are incorporated as nuisance parameters in the statistical model, and they are eliminated by marginalization (cf. sec. 4.6). Definitions of the total experimental uncertainty, the impact that each systematic uncertainty has on the cross-section measurement, the statistical uncertainty, and the uncertainty due to the limited statistical precision of the simulated and QCD-multijet distributions are given.

In the second part of this section, the impact of generator-related uncertainties on the cross-section measurement is described. Generator-related uncertainties are not incorporated as nuisance parameters in the statistical model, but their impact on the cross-section measurement is estimated with pseudo-experiments, while using the nominal statistical model to extract the signal cross section.

All estimates of the expected impact of each systematic uncertainty, as well as the statistical uncertainty, are evaluated at the SM-$t$-channel-cross-section value. A cross check of the linearity of the signal-cross-section measurement is given in the next section (“Expected Sensitivity”, 6.4). The prior-probability-density distribution for signal events is a uniform distribution in the interval $[0, \infty]$.

Table 6.6 presents a summary of the statistical model that is used to measure the $t$-channel-production cross section. Rate-changing uncertainties have an effect on the normalization of processes, but do not alter the BDT-discriminator distribution of a particular process. However, the sum of BDT-discriminator distributions of all processes is altered due to the varied composition of the contributing processes. Rate-and-shape-changing uncertainties affect the normalization and they change the shape of the BDT-discriminator distribution of a particular process. Here, both variations are fully correlated. Variations to the renormalization and factorization scales for W-boson-plus-jets events affect the BDT-discriminator distribution only, since the impact on the rate is fully covered by the normalization uncertainties (cf. sec. 6.2.3).
<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>Type</th>
<th>Nuisance parameter</th>
<th>Externalized</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Background prediction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single top quark, s-channel</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Single top quark, tW-channel</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(\bar{t}t)</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu + \text{light jets, 2 jets})</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu + bX, 2) jets</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu + cX, 2) jets</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu + \text{light jets, 3 jets})</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu + bX, 3) jets</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu + cX, 3) jets</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu + \text{light jets, } \geq 4) jets</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu + bX, \geq 4) jets</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(W \rightarrow l\nu + cX, \geq 4) jets</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(Z/\gamma^* \rightarrow l^+l^- + \text{jets})</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(WW)</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(WZ)</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(ZZ)</td>
<td>rate</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>QCD multijet, (e)</td>
<td>rate</td>
<td>6 (1 per category)</td>
<td></td>
</tr>
<tr>
<td>QCD multijet, (\mu)</td>
<td>rate</td>
<td>6 (1 per category)</td>
<td></td>
</tr>
<tr>
<td><strong>Experimental</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jet-flavor-tagging efficiency (b, c)</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Jet-flavor-tagging efficiency (u, d, s, g)</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Jet-energy scale</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Jet-transverse-momentum resolution (E_T^{\text{miss}})</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Pile-Up</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Electron id., rec., and trigger</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Muon id., rec., and trigger</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Hadronic part of the electron trigger</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parametrization 1</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Parametrization 2</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Parametrization 3</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Additional scale factor</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Limited stat. precision of templates</td>
<td>rate &amp; shape</td>
<td>1 per bin</td>
<td></td>
</tr>
<tr>
<td><strong>Generator-related</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renormalization and factorization scales</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(t\bar{t})</td>
<td>rate &amp; shape</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>W boson plus jets</td>
<td>shape</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>(t-, s)-channel</td>
<td>rate &amp; shape</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Parton-Distribution Functions</td>
<td>rate &amp; shape</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Modeling of add. partons, (t\bar{t})</td>
<td>rate &amp; shape</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>(t)-channel modeling</td>
<td>rate &amp; shape</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.6: Summary of systematic uncertainties and corresponding number of nuisance parameters in the statistical model. Generator-related uncertainties are not incorporated as nuisance parameters in the statistical model.
6.3 Impact of Systematic Uncertainties on the Measurement

6.3.1 Experimental and Normalization Uncertainties

Experimental uncertainties and uncertainties of the normalization of the background prediction are incorporated as additional nuisance parameters in the statistical model. The total experimental uncertainty \( \Delta_{\text{total exp.}} \) includes the statistical uncertainty, the uncertainty due to the limited statistical precision of the simulated and QCD-multijet templates, experimental uncertainties, and uncertainties of background normalization. The total experimental uncertainty is defined as the central-68% credible interval of the obtained signal-strength-posterior-probability distribution, and adding, in quadrature, the uncertainty of the integrated luminosity. The posterior-probability distribution is evaluated using “Asimov data” (cf. [245]).

Asimov data are pseudo-data, in which all nuisance parameters as well as the signal strength are set to their nominal values (i.e. Poisson-mean values, Gaussian-mean values, and median values for log-normal distributions), and which is generated without statistical fluctuations.

In table 6.7, the estimated total experimental uncertainty is summarized for measurements with both electron and muon categories, as well as for individual measurements in either electron or muon categories. The estimated total experimental uncertainty of the \( t \)-channel-cross-section measurement is \(-8.0/ + 8.1\%\). It is dominated by events with muons. Events with muons have a larger acceptance \((p_T > 20\text{ GeV}/c)\) compared to events with electrons \((E_T > 30\text{ GeV}/c^2)\), as well as higher reconstruction and identification efficiencies.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Expected experimental sensitivity (stat. ( \oplus ) sys.)</th>
<th>( \Delta_{\text{total exp.}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e \oplus \mu )</td>
<td></td>
<td>(-8.0/ + 8.1%)</td>
</tr>
<tr>
<td>( e ) only</td>
<td></td>
<td>(-12.3/ + 13.1%)</td>
</tr>
<tr>
<td>( \mu ) only</td>
<td></td>
<td>(-9.4/ + 9.8%)</td>
</tr>
</tbody>
</table>

Table 6.7: Expected sensitivity of the \( t \)-channel-cross-section measurement.

In the following, a break down of the total experimental uncertainty into several sub-components is given. The impact of each individual systematic uncertainty, as well as the statistical uncertainty, and the uncertainty due to the limited statistical precision of the templates are quantified.

The impact of each individual systematic uncertainty on the cross-section measurement is estimated as follows. A pseudo-experiment is generated, in which a particular systematic uncertainty is set to either the ↑ or ↓ variation, but all other parameters are set to their nominal values. Here, the ↑ and ↓ variations correspond to variations at 68% CL. The nominal, unchanged signal-extraction procedure is used to obtain a signal-cross-section estimate \( \sigma_{\text{exp.}, i} \). The median of the signal-strength-posterior-probability distribution is used as the signal-cross-section estimate. In particular, all nuisance parameters are eliminated by marginalization. The obtained difference in the signal-cross-section estimate,

\[
\Delta_{\text{exp.}, i} = \frac{(\sigma_{\text{exp.}, i} - \sigma_{\text{nominal}})}{\sigma_{\text{nominal}}}, \tag{6.1}
\]

is calculated for each systematic uncertainty \( i \) and variation (↑ and ↓). Several pseudo-experiments are generated for each variation (↑ or ↓) of a particular systematic uncertainty. This is done in order account for the finite length of the Markov Chain that is used for integration. The median difference is an estimate of the impact that a particular systematic uncertainty has on the signal-cross-section measurement.

\footnote{This reference discusses Asimov data in terms of frequentistic statistical tests with profiled nuisance parameters.}
6 Sources of Systematic Uncertainties and Expected Sensitivity

Figure 6.16 summarizes the impact that each experimental systematic uncertainty has on the cross-section measurement. The dominant experimental uncertainty is due to the normalization of W-boson production in association with heavy-flavored jets ($-3.5/ + 2.5\%$). The second-largest experimental uncertainty of the cross-section measurement is the uncertainty due to the limited statistical precision of the simulated and QCD-multijet templates ($\pm 3.1\%$). All other uncertainties have a smaller impact than the luminosity uncertainty ($\pm 2.2\%$).

The statistical uncertainty of the cross-section measurement $\Delta_{\text{stat}}$ is determined from the total uncertainty by subtracting, in quadrature, the contributions of all experimental uncertainties, as well as the contributions from uncertainties of background normalization (eq. [6.2]).

$$\Delta_{\text{stat}} = \sqrt{\Delta_{\text{total exp.}}^2 - \sum_i \Delta_{\text{exp.}, i}^2} \quad (6.2)$$

The statistical uncertainty for the measurement with electron and muon categories is $-4.7/ + 5.4\%$. In table 6.8 the statistical uncertainty is summarized also for the individual measurements with either electron or muon categories.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Statistical uncertainty $\Delta_{\text{stat}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e \oplus \mu$</td>
<td>$-4.7/ + 5.4%$</td>
</tr>
<tr>
<td>$e$ only</td>
<td>$-8.1/ + 10.5%$</td>
</tr>
<tr>
<td>$\mu$ only</td>
<td>$-7.3/ + 7.9%$</td>
</tr>
</tbody>
</table>

Table 6.8: Statistical uncertainty of the $t$-channel-cross-section measurement.

The uncertainty due to the limited statistical precision of the templates is referred to as $\Delta_{\text{BB}}$. Its implementation within the statistical model is discussed in section 6.1.1. In order to estimate the uncertainty due to the limited statistical precision of the simulated and QCD-multijet templates, first, the total experimental uncertainty is determined using a statistical model that does not include uncertainties due to the limited statistical precision of the templates. The width of the obtained signal-strength-posterior-probability density is referred to as $\Delta_{\text{total exp.}, \text{without BB}}$. Subtracting, in quadrature, $\Delta_{\text{total exp.}}$ and $\Delta_{\text{total exp.}, \text{without BB}}$ yields an estimate of the uncertainty due to the limited statistical precision of the templates (eq. [6.3]).

$$\Delta_{\text{BB}} = \sqrt{\Delta_{\text{total exp.}}^2 - \Delta_{\text{total exp.}, \text{without BB}}^2} \quad (6.3)$$

Table 6.9 summarizes the uncertainty due to the limited statistical precision of the simulated events for the combined measurement with electron and muon categories, as well as for individual measurements for events with either electrons or muons.

6.3.2 Generator-Related Uncertainties

The term “generator-related uncertainties” refers to uncertainties of assumptions that are made for event generation, i.e. they affect the theoretical description of a particular process. Generator-related uncertainties are not incorporated as nuisance parameters in the statistical model. The reason is that generator-related uncertainties cannot be sufficiently parametrized

---

4“BB” refers to Barlow and Beeston, who are the authors of the original approach to include uncertainties due to the limited statistical precision of a simulation [234].
6.3 Impact of Systematic Uncertainties on the Measurement

Figure 6.16: Expected impact of experimental uncertainties for the cross-section measurement with electron and muon events.
6 Sources of Systematic Uncertainties and Expected Sensitivity

<table>
<thead>
<tr>
<th>Categories</th>
<th>Uncertainty ∆BB</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e \oplus \mu$</td>
<td>± 3.1%</td>
</tr>
<tr>
<td>$e$ only</td>
<td>± 5.0%</td>
</tr>
<tr>
<td>$\mu$ only</td>
<td>± 3.4%</td>
</tr>
</tbody>
</table>

Table 6.9: Uncertainty due to limited statistical precision of the templates that are used for the $t$-channel-cross-section measurement.

since their evaluation requires too much computing time, in particular if they include the detector simulation. Therefore, only two points of the parametrization are generated (↑ and ↓ variations). The generated event samples cover the generator-related uncertainties at 68% CL.

As an example, in order to study the impact of the $Q^2$ scale variations for $t\bar{t}$ events, dedicated event samples are generated with varied $Q^2$ scales. $Q^2$-scale variations by factors of $(0.5)^2$ and $(2)^2$ are interpreted as ↓ and ↑ variations, i.e. the assumption is made that the variations cover changes to the $Q^2$ scale at 68% CL. Variations that cover the $Q^2$ uncertainty of smaller or larger confidence levels are not simulated due to limited computing resources. However, the extrapolation of the impact of the $Q^2$-scale variations on the $t\bar{t}$ yield and kinematics, which is known at 68% CL to smaller or larger variations is nontrivial.

The impact of generator-related uncertainties on the cross-section measurement is estimated with pseudo-experiments. Pseudo-experiments are generated according to ↑ and ↓ variations of a particular systematic uncertainty, but the nominal statistical model is used to estimate the signal cross section. The relative uncertainty $\Delta_{\text{gen., } i}$

$$\Delta_{\text{gen., } i} = \frac{\sigma_{\text{gen., } i} - \sigma_{\text{nominal}}}{\sigma_{\text{nominal}}}$$

is a measure of the impact of a particular systematic uncertainty $i$ on the cross-section measurement. Here, $\sigma_{\text{gen., } i}$ refers to the mean cross-section estimate that results for a particular generator-related systematic variation $i$. The cross-section estimate that is obtained in the nominal scenario is referred to as $\sigma_{\text{nominal}}$. The differences are separately calculated for variations in both directions (↑ and ↓) of each systematic uncertainty. An uncertainty that does not have a corresponding nuisance parameter in the statistical model is also referred to as “externalized” uncertainty in the following.

Figure 6.17 summarizes the expected impact of generator-related uncertainties of the cross-section measurement. The dominant contribution among the generator-related uncertainties is the signal modeling with an impact of ±4.6% on the cross-section measurement. The signal-modeling uncertainty is symmetrized in a sense that half of the total uncertainty is taken in both “+” and “−” directions (cf. [24]), which also takes into account that the nominal simulation in NLO+PS accuracy is compared to a simulation in effective NLO+PS accuracy. The second-largest uncertainty results from variations to the renormalization and factorization scales of W-boson-plus-jets events. Variations to the PDFs are the third largest generator-related uncertainty, though their impact is small with only ±1.5%.

The total generator-related uncertainty $\Delta_{\text{total gen.}}$ is defined as

$$\Delta_{\text{total gen.}} = \sqrt{\sum_i \Delta_{\text{gen., } i}^2},$$

in which the summation runs over all generator-related uncertainties. Table 6.10 summarizes the impact of all generator-related uncertainties for the combined measurement with electron
Figure 6.17: Expected impact of generator-related uncertainties for the cross-section measurement with electron and muon events.

and muon events, as well as for individual measurements in either electron or muon categories.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Uncertainty $\Delta_{\text{total_gen.}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e \oplus \mu$</td>
<td>$-4.9 / +6.0%$</td>
</tr>
<tr>
<td>$e$ only</td>
<td>$\pm 9.1%$</td>
</tr>
<tr>
<td>$\mu$ only</td>
<td>$\pm 7.3%$</td>
</tr>
</tbody>
</table>

Table 6.10: Generator-related uncertainty for the $t$-channel-cross-section measurement.
6 Sources of Systematic Uncertainties and Expected Sensitivity

6.4 Expected Sensitivity

The total expected uncertainty of the \( t \)-channel-cross-section measurement (\( \Delta_{\text{total}} \)) consists of two components. The first contribution is the total experimental uncertainty (\( \Delta_{\text{total exp.}} \)), which is obtained from signal-strength-posterior-probability distribution. The second contribution is the impact of all generator-related, externalized uncertainties, which are summed in quadrature (\( \Delta_{\text{total gen.}} \)). Both contributions \( \Delta_{\text{total exp.}} \) and \( \Delta_{\text{total gen.}} \) are added in quadrature to obtain the total expected uncertainty (eq. 6.6).

\[
\Delta_{\text{total}} = \sqrt{\Delta_{\text{total exp.}}^2 + \Delta_{\text{total gen.}}^2}
\]

The prior-probability-density distribution for signal events is a uniform distribution in the interval \([0, \infty]\). The expected sensitivity is estimated for the SM-\( t \)-channel-cross-section value. A linearity check of the statistical inference is presented in the last paragraph of this section.

Table 6.11 summarizes the expected sensitivity of the \( t \)-channel-cross-section measurements with electron and muon events, events with electrons only, and events with muons only. The total uncertainty of the cross-section measurement is estimated to be \(-9.4/ + 10.1\%\). The expected sensitivity in the individual charged-lepton-flavor categories is \(-15.3/ + 16.0\%\) for a measurement that only takes electron events into account, and \(-11.9/ + 12.2\%\) for a measurement with muon events. The simultaneous measurement of events with both charged-lepton flavors enhances the sensitivity compared to the individual measurements. The muon channel is more sensitive than the electron channel, since it has a larger acceptance and higher lepton reconstruction and identification efficiencies. Therefore, not only the statistical precision is larger in the muon channel, but also systematic uncertainties can be more constrained.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Expected sensitivity ( \Delta_{\text{total}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e \oplus \mu )</td>
<td>(-9.4/ + 10.1%)</td>
</tr>
<tr>
<td>( e ) only</td>
<td>(-15.3/ + 16.0%)</td>
</tr>
<tr>
<td>( \mu ) only</td>
<td>(-11.9/ + 12.2%)</td>
</tr>
</tbody>
</table>

Table 6.11: Expected sensitivity of the \( t \)-channel-cross-section measurement.

The uncertainty of the cross-section measurement due to the generator-related uncertainties is estimated to be \(-4.9/ + 6.0\%\). The statistical uncertainty is estimated to be \(-4.7/ + 5.4\%\) (for the SM-\( t \)-channel-cross-section prediction). The statistical uncertainty is of the same order as the impact of the generator-related uncertainties, but smaller than the impact from experimental uncertainties. The uncertainty due to the experimental-systematic uncertainties is estimated to be \(-6.5/ + 6.0\%\).

The impact that each uncertainty has on the cross-section measurement is summarized in table 6.12. Individual uncertainties are grouped together assuming Gaussian-error propagation for illustration purposes.

The dominant single sources of systematic uncertainties are the \( t \)-channel modeling, which has an impact of \( \pm 4.6\% \) on the measured signal cross section, and the modeling of the W-boson-plus-jets events, both in shape and rate. The total uncertainty of the \( t \)-channel cross-section estimate due to the W-boson-plus-jets modeling is \(-3.5/ + 4.2\%\), including variations of the renormalization and factorization scales. Here, the shape effects on the discriminator output
due to variations of the renormalization and factorization scales result in an uncertainty of $-0.0/ + 3.4\%$, which is of about the same order as the uncertainty due to the limited statistical precision of the simulated distributions. The uncertainty of the normalization of W-boson-plus-jets events, which are dominated by the uncertainty of W-boson production in association with heavy-flavored jets, have an impact of $-3.5/ + 2.5\%$ on the measurement. The uncertainty due to the limited statistical precision of the simulated samples and QCD-multijet events is the third most important contribution to the systematic uncertainty ($\pm 3.1\%$). Other sources of systematic uncertainties are of minor importance.

Figure 6.18 shows the obtained cross section as a function of the input-signal strength. This dependency is obtained with pseudo-experiments. The obtained cross-section estimate ($\sigma_{\text{measured}}$) linearly depends on the input cross section ($\sigma_{\text{input}}$), and its dependence can be described as

$$\sigma_{\text{measured}} = (0.58 \pm 0.17) + (0.998 \pm 0.002) \times \sigma_{\text{input}}.$$  

(6.7)

The interpretation is that the median of the signal-strength posterior is an unbiased estimator of the true $t$-channel cross section. For the final measurement, also the full posterior-probability distribution is presented.

![Graph showing the obtained cross section as a function of the input cross section](image)

Figure 6.18: Cross check of the linearity of the obtained $t$-channel cross-section estimate w.r.t. the input cross section. This dependency is estimated with pseudo-experiments. The filled area corresponds to the central-68% credible interval of the signal-strength-posterior-probability distribution.
### 6 Sources of Systematic Uncertainties and Expected Sensitivity

<table>
<thead>
<tr>
<th>Source of Uncertainty</th>
<th>Impact on cross-section measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistical</strong></td>
<td>−4.7/ + 5.4 %</td>
</tr>
<tr>
<td><strong>Experimental</strong></td>
<td>−6.5/ + 6.0 %</td>
</tr>
<tr>
<td>Statistical precision of the templates</td>
<td>±3.1%</td>
</tr>
<tr>
<td>Luminosity determination</td>
<td>±2.2%</td>
</tr>
<tr>
<td><strong>Detector related</strong></td>
<td></td>
</tr>
<tr>
<td>Jet-flavor-tagging efficiency</td>
<td>±1.6%</td>
</tr>
<tr>
<td>Jet-energy scale</td>
<td>±0.6%</td>
</tr>
<tr>
<td>Jet-transverse-momentum resolution</td>
<td>±0.1%</td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$</td>
<td>±0.2%</td>
</tr>
<tr>
<td>Pile-Up</td>
<td>±0.4%</td>
</tr>
<tr>
<td>Electron id., rec., and trigger</td>
<td>±1.2%</td>
</tr>
<tr>
<td>Muon id., rec., and trigger</td>
<td>±1.9%</td>
</tr>
<tr>
<td>Hadronic part of the electron trigger</td>
<td>±1.5%</td>
</tr>
<tr>
<td><strong>Background normalization</strong></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>±1.0%</td>
</tr>
<tr>
<td>$W$ boson plus jets</td>
<td>−3.5/ + 2.5 %</td>
</tr>
<tr>
<td>light-flavored jets (u, d, s, g)</td>
<td>±0.4%</td>
</tr>
<tr>
<td>heavy-flavored jets (c, b)</td>
<td>−3.5/ + 2.5 %</td>
</tr>
<tr>
<td>Single top quark ($s\rightarrow tW$-channel),</td>
<td>±0.6%</td>
</tr>
<tr>
<td>$Z/\gamma^*$ ($\rightarrow l^+l^-$) + jets, Diboson</td>
<td>±0.6%</td>
</tr>
<tr>
<td>QCD multijet, $\mu$</td>
<td>±1.7%</td>
</tr>
<tr>
<td>QCD multijet, $e$</td>
<td>±0.8%</td>
</tr>
<tr>
<td><strong>Generator related</strong></td>
<td>−4.9/ + 6.0 %</td>
</tr>
<tr>
<td>Renormalization and factorization scales</td>
<td></td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>±0.9%</td>
</tr>
<tr>
<td>$W$ boson plus jets</td>
<td>−0.0/ + 3.4 %</td>
</tr>
<tr>
<td>$t\rightarrow s\rightarrow tW$-channel</td>
<td>±0.2%</td>
</tr>
<tr>
<td>Parton-distribution functions</td>
<td>±1.5%</td>
</tr>
<tr>
<td>Modeling of add. partons, $t\bar{t}$</td>
<td>±0.4%</td>
</tr>
<tr>
<td>$t$-channel modeling</td>
<td>±4.6%</td>
</tr>
<tr>
<td><strong>Expected experimental uncertainty, stat. + sys.</strong></td>
<td>−8.0/ + 8.1 %</td>
</tr>
<tr>
<td><strong>Expected total uncertainty</strong></td>
<td>−9.4/ + 10.1 %</td>
</tr>
</tbody>
</table>

Table 6.12: Impact of systematic uncertainties on the inclusive $t$-channel-cross-section measurement.
6.5 Charge-Ratio Measurement

In this section, the statistical model and expected sensitivities of the cross-section measurements of single-top-quark production ($\sigma_t$) and single-top-antiquark production ($\sigma_{\bar{t}}$), as well as their ratio ($R_{\sigma_t/\sigma_{\bar{t}}}$) are discussed. The lepton charge $Q_l$ is used as an identifier for single top quark ($Q_l = +e$) and antiquark production ($Q_l = +e$), as well as for $W^\pm$ production.

The first part of this section discusses changes to the statistical model w.r.t. the statistical model that is used for the measurement of the inclusive production cross section. The second part refers to the expected sensitivity of all three measurements $\sigma_t$, $\sigma_{\bar{t}}$, and $R_{\sigma_t/\sigma_{\bar{t}}}$.

6.5.1 Statistical Model

The statistical model is extended for the charge-ratio measurement, as well as for the measurements of the individual cross-sections $\sigma_t$ and $\sigma_{\bar{t}}$. The changes to the statistical model include new analysis categories, additional nuisance parameters, and a second parameter of interest. These changes are discussed in the following paragraphs.

New analysis categories are introduced. All analysis categories, which are used for the measurement of the inclusive cross section (cf. sec. 4.1) are splitted by lepton charge ($\mu^+$, $\mu^-$, $e^+$, $e^-$), i.e. the number of used categories is doubled. Furthermore, the “2 jets, 0 btags” category is introduced for events with muons. The data then constrains the charge-production ratio for $W$-boson-plus-jets events in situ. Only one bin is used for the “2 jets, 0 btags” categories. In total, 26 categories ($6+1\ \mu^+$, $6+1\ \mu^-$, $6\ e^+$, $6\ e^-$) are used to infer $R_{\sigma_t/\sigma_{\bar{t}}}$ from data.

Furthermore, a new parameter of interest ($\mu_{R_{\sigma_t/\sigma_{\bar{t}}}}$) is introduced for the charge-ratio measurement. $\mu_{R_{\sigma_t/\sigma_{\bar{t}}}}$ varies the amount of $t$-channel events with positively-charged leptons (eq. 6.8). $\mu_{R_{\sigma_t/\sigma_{\bar{t}}}}$ effectively is a scale factor to the ratio of top-quark production to top-antiquark production ($R_{\sigma_t/\sigma_{\bar{t}}}$) as predicted by the SM. Furthermore, the signal-strength parameter $\mu$ is used to modify the inclusive $t$-channel cross section (cf. eq. 4.27). Both $\mu$ and $\mu_{R_{\sigma_t/\sigma_{\bar{t}}}}$ have flat prior-probability distributions (cf. sec. 4.6).

The joint-likelihood function (cf. eq. 4.37) for the charge-ratio measurement ($R_{\sigma_t/\sigma_{\bar{t}}}$) is given by

$$L(\mu, \tilde{\theta}, \tilde{x}) = \prod_{i=1}^{N_{\text{categories}}} \prod_{j=1}^{N_{\text{bins}(i)}} \left[ \text{Poisson}(x_{i,j}^+, \mu \times \mu_{R_{\sigma_t/\sigma_{\bar{t}}}} \times s_{i,j}^+ (\tilde{\theta}) + b_{i,j}^+ (\tilde{\theta})) + \text{Poisson}(x_{i,j}^-, \mu \times s_{i,j}^- (\tilde{\theta}) + b_{i,j}^- (\tilde{\theta})) \right].$$

Here, the indices “+” and “−” mean that only positively-charged or negatively-charged leptons are taken into account when counting the event yields for signal events ($s_{i,j} (\tilde{\theta})$) or background events ($b_{i,j} (\tilde{\theta})$) (cf. sec. 4.6.2).

For separate measurements of top-quark and top-antiquark production, the nuisance parameter $\mu$ is omitted for $t$-Channel events with top-quarks (+), i.e. one parameter varies the strength of top-quark production and one parameter varies the strength of top-antiquark production. The joint-likelihood function for the separate measurements of top-quark production...
and top-antiquark production is given by

\[ L(\mu, \vec{\theta}; \vec{x}) = \prod_{i=1}^{N_{\text{categories}}} \prod_{j=1}^{N_{\text{bins}(i)}} \left[ \text{Poisson}(x_{i,j}^+ | \mu^+ \times s_{i,j}^+ (\vec{\theta}) + b_{i,j}^+ (\vec{\theta})) + \text{Poisson}(x_{i,j}^- | \mu^- \times s_{i,j}^- (\vec{\theta}) + b_{i,j}^- (\vec{\theta})) \right], \tag{6.9} \]

in which \( \mu^+ \) refers to the signal strength of top-quark production and \( \mu^- \) refers to the signal strength of top-antiquark production. Both \( \mu^+ \) and \( \mu^- \) have flat prior-probability distributions.

**Additional nuisance parameters** are introduced to vary the production rates of charge-asymmetric processes. One additional nuisance is introduced for \( s \)-channel-top-quark production, since it is charge-asymmetric, too. Furthermore, it is assumed that QCD-multijet production is charge symmetric, and only one nuisance parameter is introduced to account for the QCD-multijet yield in the “2 jets, 0 btags” category.

For \( W \)-boson-plus-jets processes, one additional nuisance parameter is included for events with \( W^+ \) bosons, and one additional nuisance parameter is included for events with \( W^- \) bosons. Both nuisance parameters are independent of the number and flavor of associated jets. Hence, the statistical model is able to infer the \( W^\pm \)-production rates from data by using the “2 jets, 0 btags” category. The nuisance parameter that scales the event yield for “\( W (\to l\nu) + \) light jets events with 2 jets” is omitted (cf. table 6.6).

Moreover, two additional nuisance parameters are introduced that independently vary the amount of events with \( W^+ \) bosons for \( W (\to l\nu) + cX \) production and \( W (\to l\nu) + bX \) production. Both nuisance parameters are independent of the number and flavor of associated jets. These variations effectively introduce an extrapolation uncertainty, which accounts for the assumption that the charge-asymmetry, which is derived in \( W \)-boson-plus-light-jets events in data, can be extrapolated to events with heavy-flavored jets.

Table 6.13 summarizes the additional nuisance parameters, which are used for the extended statistical model, w.r.t. the statistical model that is used for the inclusive cross-section measurement (cf. table 6.6).

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Additional nuisance parameter</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Charge-ratio modifier</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t )-channel (top quark)</td>
<td>rate</td>
<td>1 (“( \mu R_{\sigma_t/\sigma_\ell} )” or “( \mu^+ )”</td>
<td>flat prior</td>
</tr>
<tr>
<td>( s )-channel (top quark)</td>
<td>rate</td>
<td>1</td>
<td>30%</td>
</tr>
<tr>
<td>( W^+ )-boson production, any jets</td>
<td>rate</td>
<td>1</td>
<td>30%</td>
</tr>
<tr>
<td>( W^- )-boson production, any jets</td>
<td>rate</td>
<td>1</td>
<td>30%</td>
</tr>
<tr>
<td>( W^+ )-boson production + ( cX )</td>
<td>rate</td>
<td>1</td>
<td>10%</td>
</tr>
<tr>
<td>( W^+ )-boson production + ( bX )</td>
<td>rate</td>
<td>1</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Background prediction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QCD multijet, “2 jets, 0 btags”</td>
<td>rate</td>
<td>1</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 6.13: Additional nuisance parameters as implemented in the “extended” statistical model, which is used for the measurement of \( R_{\sigma_t/\sigma_\ell} \), \( \sigma_t \), and \( \sigma_\ell \).
### 6.5.2 Expected Sensitivity

The total expected experimental uncertainty is obtained with Asimov data from the posterior-probability distribution (similar as discussed for the inclusive cross-section measurement in section 6.3.1). The uncertainty of the integrated luminosity is added in quadrature for the \( \sigma_t \) and \( \bar{\sigma}_t \) measurements.

The statistical uncertainty is estimated by constraining each nuisance parameters to its nominal value, generating Asimov data (cf. sec. 6.3.1), and determining the 68\%-credible interval of the posterior-probability distribution. In particular, this definition of the statistical uncertainty differs from the definition given in section 6.3.1. The experimental systematic uncertainty is calculated from the total expected experimental uncertainty by subtracting in quadrature the statistical uncertainty.

In table 6.14, the expected sensitivity with its individual components (statistical, experimental, and generator-related) is summarized for the measurements of \( \sigma_t \), \( \bar{\sigma}_t \), and \( R_{\sigma_t/\bar{\sigma}_t} \). The absolute numerical precision of each uncertainty value is ±0.5\%, except for the statistical uncertainties, which are known up to an absolute precision of ±0.05\%.

<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>Impact on measurement</th>
<th>( \sigma_t )</th>
<th>( \bar{\sigma}_t )</th>
<th>( R_{\sigma_t/\bar{\sigma}_t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical</td>
<td></td>
<td>±5.8%</td>
<td>±9.0%</td>
<td>−10.0/ + 11.5%</td>
</tr>
<tr>
<td>Experimental</td>
<td></td>
<td>±5.2%</td>
<td>±7.3%</td>
<td>−4.8/ + 6.7%</td>
</tr>
<tr>
<td>Generator related</td>
<td></td>
<td>−7.5/ + 6.4%</td>
<td>−17.8/ + 5.4%</td>
<td>−3.8/ + 14.5%</td>
</tr>
<tr>
<td>Renormalization and factorization scales</td>
<td></td>
<td>±3.1%</td>
<td>±3.1%</td>
<td>±0.5%</td>
</tr>
<tr>
<td>t( \bar{t} )</td>
<td></td>
<td>−1.3/ + 5.2%</td>
<td>−1.3/ + 0.0%</td>
<td>−0.0/ + 5.8%</td>
</tr>
<tr>
<td>W boson plus jets</td>
<td></td>
<td>±0.7%</td>
<td>−4.9/ + 0.0%</td>
<td>−1.9/ + 4.8%</td>
</tr>
<tr>
<td>t( \bar{t} )-s-channel</td>
<td></td>
<td>±0.9%</td>
<td>±2.2%</td>
<td>±2.0%</td>
</tr>
<tr>
<td>Parton-distribution functions</td>
<td></td>
<td>±0.8%</td>
<td>±1.8%</td>
<td>±1.7%</td>
</tr>
<tr>
<td>40 EVs of CTEQ6M NLO [52]</td>
<td></td>
<td>−0.4%</td>
<td>−1.3%</td>
<td>+0.4%</td>
</tr>
<tr>
<td>Best-fit of MSTW2008 NLO [75]</td>
<td></td>
<td>+0.5%</td>
<td>+1.3%</td>
<td>−1.1%</td>
</tr>
<tr>
<td>Mean of NNPDF2.1 NLO [244]</td>
<td></td>
<td>±1.6%</td>
<td>±3.8%</td>
<td>±2.5%</td>
</tr>
<tr>
<td>Modeling of add. partons, tt</td>
<td></td>
<td>−6.4%</td>
<td>−16.2%</td>
<td>+12.0%</td>
</tr>
<tr>
<td>t-channel modeling</td>
<td></td>
<td>±7.8%</td>
<td>±11.6%</td>
<td>−11.1/ + 13.3%</td>
</tr>
<tr>
<td>Expected exp. uncertainty, stat. &amp; sys.</td>
<td></td>
<td>−10.8/ + 10.1%</td>
<td>−21.3/ + 12.8%</td>
<td>−11.7/ + 19.7%</td>
</tr>
<tr>
<td>Expected total uncertainty</td>
<td></td>
<td>−10.8/ + 10.1%</td>
<td>−21.3/ + 12.8%</td>
<td>−11.7/ + 19.7%</td>
</tr>
</tbody>
</table>

Table 6.14: Expected sensitivity of separate measurements of top-quark production (\( \sigma_t \)) and top-antiquark production (\( \bar{\sigma}_t \)), as well as their ratio (\( R_{\sigma_t/\bar{\sigma}_t} \)). The abbreviation “exp.” refers to “experimental”, and “EVs” refers to “Eigenvectors”.

**Generator-related uncertainties** The dominant sources of systematic uncertainties are generator-related uncertainties.

The dominant generator-related uncertainty is the uncertainty of the \( t \)-channel modeling as obtained by a comparison of the POWHEG BOX and COMPHEP event generators. Both event generators sample events with the CTEQ6M NLO PDF set [52].

---

5Technically, \( \delta \) distributions are used as prior-probability distributions to constrain the nuisance parameters.
If the $t$-channel signal is generated with the COMPHEP generator and the POWHEG BOX distributions are used to determine the signal strengths (as well as their ratio), the top-antiquark-production cross section is estimated to be 16.2\% lower. The top-quark-production cross section is determined to be 6.4\% lower in this case. Since the relative uncertainties of the top-quark-production cross section ($\Delta_{\sigma_t}$) and top-antiquark-production cross section ($\Delta_{\bar{\sigma}_t}$) propagate to the charge-production ratio ($R_{\sigma_t/\sigma_{\bar{t}}}$) like $R' \approx R \times (1 + \Delta_{\sigma_t})/(1 + \Delta_{\bar{\sigma}_t})$, the uncertainty of $R_{\sigma_t/\sigma_{\bar{t}}}$ becomes +12.0\%.

The obtained shift of the inclusive production cross section is $\approx 11\%$, which is compatible with the modeling uncertainty that is quoted in table 6.12. For the inclusive cross-section measurement, half of the shift was used as an uncertainty in both $\uparrow$ and $\downarrow$ directions. For the charge-production-ratio measurement, the uncertainty is not symmetrized.

The $Q^2$-scale variations for $W$-boson-plus-jets events are estimated by using dedicated samples for all categories, except for the “2 jets, 1 btag” categories. In order to enrich the statistical precision of the $Q^2$-scale varied templates, the $Q^2$-scale variations are extrapolated from the muon-“2 jets, 0 btags” category to the “2 jets, 1 btag” categories. Only shape effects are taken into account, since the background normalization is allowed to float within each jet-multiplicity bin individually (similar to the inclusive cross-section measurement). Hence, the normalization effect is already covered by the nuisance parameters for background normalization. Varying the $Q^2$-scale up yields a simultaneous change in $\sigma_t$ and $\sigma_{\bar{t}}$ of $-1.3\%$, which cancels for the charge-production-ratio measurement. Varying down the $Q^2$ scale has a significant impact ($+5.2\%$) on the obtained value of $\sigma_t$, while the impact on $\sigma_{\bar{t}}$ is negligible. This uncertainty directly translates into a large uncertainty of the charge-ratio measurement ($+5.8\%$).

The uncertainties due to the parametrization of the PDFs (sec. 6.2.2) have a minor impact on the $R_{\sigma_t/\sigma_{\bar{t}}}$ measurement. Similar to the inclusive cross-section measurement, PDF reweighting is used to alter the acceptance of $t$-channel events while keeping the normalization constant.

**Statistical and experimental uncertainties** The statistical uncertainties become important if $\sigma_t$ and $\sigma_{\bar{t}}$ are measured separately. Top-antiquark production has the lowest event yield and the largest statistical uncertainty. Since the statistical uncertainties of the individual contributions $\sigma_t$ ($\pm 5.8\%$) and $\sigma_{\bar{t}}$ ($\pm 9.0\%$) quadratically add up for the charge-ratio measurement, the statistical uncertainty of $R_{\sigma_t/\sigma_{\bar{t}}}$ is driven by the event yield of top-antiquark production. In particular, the charge-ratio measurement $R_{\sigma_t/\sigma_{\bar{t}}}$ is limited rather by the statistical precision ($-10.0/ + 11.5\%$) than by experimental uncertainties.

The experimental uncertainties of the $R_{\sigma_t/\sigma_{\bar{t}}}$ measurement are $-4.8/ + 6.7\%$, which is about half of the statistical uncertainties. The experimental uncertainties for the measurement of $\sigma_{\bar{t}}$ are $\pm 7.3\%$, and they are $\pm 5.2\%$ for the measurement of $\sigma_t$.

---

6Technically, the shape variations of the BDT discriminator distributions due to the $Q^2$-scale variations are calculated in the muon-“2 jets, 0 btags” category. The obtained shape variations are applied to the (nominal) templates for $W$-boson-plus-jets events in the “2 jets, 1 btag” categories.
7 Measurement Results

This chapter is divided into three parts. The first section (7.1) discusses the results of the inclusive $t$-channel-production-cross-section measurement $\sigma_{t\text{-channel}}$. The second part (7.2) focuses on the measurement of the CKM matrix element $|V_{tb}|$, which is derived from the inclusive cross-section measurement. Finally, the individual cross sections of top-quark-production ($\sigma_t$) and top-antiquark-production ($\sigma_{\bar{t}}$), as well as their ratio $R_{\sigma_t/\sigma_{\bar{t}}}$, are determined (sec. 7.3).

7.1 Measurement of the Inclusive $t$-Channel Cross Section

The single-top-quark $t$-channel production cross section is measured to be

$$\sigma_{t\text{-channel}} = 66.6^{+5.4}_{-5.3} \text{ (exp.) } +4.0_{-3.3} \text{ (gen.) pb}$$

(7.1)

The individual measurement for events with electrons results in

$$\sigma_{t\text{-channel}} = 66.4^{+8.4}_{-7.9} \text{ (exp.) } +6.0_{-6.0} \text{ (gen.) pb},$$

(7.2)

and the measurement for events with muons results in

$$\sigma_{t\text{-channel}} = 66.6^{+7.0}_{-6.6} \text{ (exp.) } +4.8_{-4.8} \text{ (gen.) pb}.$$  

(7.3)

The measured $t$-channel cross section is consistent with the SM prediction, which is $64.6^{+2.1}_{-0.7}+1.0$ pb in approximate NNLO accuracy for a top-quark mass of 172.5 GeV/$c^2$ [8]. The individual cross-section measurements in either electron or muon categories are consistent with each other and the combined measurement. Figure 7.1 shows a comparison of this measurement to the SM prediction in approximate NNLO QCD accuracy [8], as well as to other measurements [19, 20, 22, 100].

The prior-probability-density distribution for signal events is chosen as uniformly distributed in the interval $[0, \infty]$. The 50% quantile of the posterior-probability distributions is used as the cross-section estimate. The obtained posterior-probability distribution for the signal-production cross section is given in figure 7.2. Since generator-related uncertainties are externalized, the posterior-probability distribution only includes experimental systematic uncertainties. The median of the posterior-probability distribution is represented by the straight line, and the central-68% interval by the filled area. The SM cross-section prediction in approximate NNLO [8] is represented by the dashed line. The hatched area represents uncertainties of the theoretical prediction. Uncertainties due to the choice of the renormalization and factorization scales are estimated by simultaneously varying both scales by factors two and half. Uncertainties of the parametrization of the MSTW2008 NNLO PDF set [75] are given at 90% CL. Both measurement and SM prediction are in good agreement with each other.
Figure 7.1: Comparison of this measurement (blue dot) to the SM prediction in approximate NNLO QCD accuracy [8], as well as to other measurements [19, 20, 22, 100].

Figure 7.2: Observed signal-strength-posterior-probability distribution for the combined cross-section measurement with electron and muon categories. The central-68% interval covers experimental uncertainties only, since generator-related uncertainties are externalized.
7.1 Measurement of the Inclusive $t$-Channel Cross Section

**Observed Nuisance parameters**  Figure 7.3 shows the preferred nuisance-parameter values (dot-shaped markers) and their central-68% intervals (error markers) that are obtained from the posterior-probability distributions. The posterior-probability distribution for each nuisance parameter is obtained by marginalizing all nuisance parameters, except the nuisance parameter under study. The preferred value is defined as the median of the posterior-probability distribution of the corresponding nuisance parameter. The nuisance parameters are shown for the combined measurement with both electron and muon events.

The median values of the prior-probability distributions, i.e. the nominal values of each corresponding systematic uncertainty, are represented by the straight vertical lines in figure 7.3. The central-68% and central-95% intervals of the prior-probability distributions are shown as filled areas for each nuisance parameter.

A comparison between the central-68% intervals of the prior-probability and posterior-probability distributions reveals the amount of in-situ constraint of the corresponding nuisance parameter by observed data. A smaller width of the posterior-probability distribution w.r.t. the prior-probability distribution indicates an in-situ constraint, and, therefore, a reduced impact on the cross-section measurement of the corresponding systematic uncertainty.

The central-68% and central-95% intervals of the prior-probability distributions are scaled by 5.0 for the $t\bar{t}$-normalization-nuisance parameter and by 10.0 for the lepton-efficiency-nuisance parameters. The median values itself, which are represented with dot-shaped markers, are not scaled.

The posterior-probability distributions of all nuisance parameters are compatible with the a-priori expectation. All preferred nuisance parameters are within the central-68% intervals of the prior-probability distributions, except for the nuisance parameters for QCD-multijet events with muons, three jets, and two b-tagged jets, and the trigger parametrization for events with electrons. The preferred values of the latter two nuisance parameters are covered at 95% CL w.r.t. the prior-probability distributions.

The nuisance parameters for muon and electron reconstruction efficiencies slightly constrain each other. The preferred muon-efficiency is $1.023 \pm 2.5\%$, the preferred electron-efficiency is $0.975 \pm 2.5\%$. Diboson processes, Z-boson-plus-jets production, and single-top-quark $s$-Channel and $tW$-Channel processes are minor backgrounds that cannot be sufficiently constrained by data.

The uncertainty of the top-quark-pair normalization is constrained by data to approximately one third of the input value. The preferred $t\bar{t}$ normalization is $(1.003 \pm 4.8\%) \times 157.5$ pb, and, therefore, close to the input value. The $t\bar{t}$ normalization can be interpreted as a $t\bar{t}$-cross-section measurement with experimental uncertainties only.

The data excess in the “2 jets, 1 btag” category is partially explained by an excess of W-boson production in association with b-flavored jets. The event yield of W-boson production in association with two jets, in which at least one jet is originating from b-quark fragmentation, is $(62 \pm 15)\%$ larger in data than predicted by simulation. Instead, the amount of W-boson events, which have two jets and at least one jet originating from c-quark fragmentation, is found to be lower in data compared to simulation ($-51\% +82\% -51\%$). Both process, “W-boson +bX” and “W-boson +cX” are anti-correlated. “W-boson +bX” production is underestimated in categories with three or four jets as well. The a-posteriori normalization of W-boson production in association with light-flavored jets is compatible with the a-priori prediction.

---

1The uncertainties include the statistical precision and systematic uncertainties, except for the luminosity uncertainty of 2.2%, parametrization of PDFs, and generator-related uncertainties.
7 Measurement Results

The a-priori normalization of QCD-multijet events with electrons is dominated by the definition of the data-sideband region. Here, QCD-multijet events are normalized to the average of two yield estimations, which result from different data-sideband regions. These data-sideband regions are either enriched in electrons from photon conversions or inverted electron-identification criteria, except conversion rejection (cf. sec. 5.1). The uncertainty of the QCD-multijet normalization for events with electrons is significantly constrained by data. The preferred normalizations are lower than the input values among all electron categories, and the obtained values are consistent with the a-priori normalization that is obtained with isolated electrons from photon conversions. The observed normalization of QCD-multijet events with muons is compatible with the prior expectation.

The impact on the cross-section measurement due to variations of the jet-flavor-tagging efficiency and jet-energy scale is significantly constrained by data. The impact due to uncertainties of the jet-energy scale is reduced by a factor of 4. The impact due to uncertainties of jet-flavor-tagging efficiencies are constrained by a factor 2. The mistag efficiency is not constrained by data, since the event yield of W-boson-plus-light-jets events is relatively small compared to other background processes. Variations to the jet-$p_T$ resolution, to the number of additional Pile-Up interactions, $E_T^{\text{miss}}$, or different parametrizations of the hadronic part of the electron-cross-trigger efficiency have a negligible impact on the measured $t$-channel cross section.

The median of the posterior for the nuisance parameter “Trigger parametrization: SF” is shifted by $-1.6\sigma$ w.r.t. the median of the prior distribution ($0 \pm 1$). This nuisance parameter corresponds to an additional scale factor to the electron-cross-trigger efficiency. It is used to correct the electron-cross-trigger efficiency for the total number of b-flavored jets in an event (cf. sec. 3.5.4). The additional scale factor of the electron-cross-trigger efficiency is derived by comparing two different versions of the CMS reconstruction software (cf. sec. 3.5.4). Residual differences in the physics-objects reconstruction, e.g. in the charged-lepton isolation as required by the electron-cross trigger, can explain the observed shift. However, the observed shift of this nuisance parameter by $-1.6\sigma$ can be also explained by statistics when bearing in mind that the statistical model has in total 40 nuisance parameters, whereof 39 parameters have posterior distributions that are compatible with the prior distributions at 1$s$ level. Moreover, the posterior distribution nearly has the same width as the prior distribution, which means that the statistical model only has a small sensitivity to this nuisance parameter.
7.1 Measurement of the Inclusive $t$-Channel Cross Section

Figure 7.3: Preferred nuisance-parameter values and central-68% intervals of the posterior-probability distributions for the cross-section measurement with electron and muon events. The central-68% and central-95% intervals of the prior-probability distributions are shown as underlying boxes for each nuisance parameter.
7 Measurement Results

Best-Fit BDT-Discriminator Distributions In the following paragraphs, the BDT-discriminator distributions are shown for observed data and the best-fit model. In order to obtain the best-fit BDT-discriminator distributions, simulated signal and background contributions, as well as QCD-multijet events, are scaled to the best-fit results of the combined measurement with electron and muon events. Here, the set of nuisance-parameter values that corresponds to the maximum-likelihood estimate is referred to as the “best-fit” result. The best fit is obtained using the nominal statistical model, identical to the model used for the inclusive cross-section measurement. In particular, all obtained nuisance-parameter estimates are compatible with the results that are obtained from the Bayesian inference.

The agreement between the best-fit model and the observed data is quantified with two statistical hypothesis tests (as implemented in ref. [206]). The results of these tests, the KS-test probabilities and $\chi^2$-test p-values, are shown in the upper right corner of each plot. Large KS-test probabilities and high $\chi^2$-test p-values are obtained, and both tests are passed in each individual category.

Figure 7.4 shows the comparison between observed and best-fit BDT-discriminator distributions in the signal-enriched categories “2 jets, 1 btag” (top-left plot) and “3 jets, 1 btag” (top-right plot), as well as for the sum of all signal-depleted categories (bottom plot) that are used in the statistical inference. Each distribution is shown for the sum of electron and muon events. A good agreement between the best-fit model and observed data is obtained. In the “2 jets, 1 btag” category (top-left plot), the right tail of the BDT-discriminator distribution consists of a almost pure signal region. The left tail of the distribution is signal depleted and well reproduced by the simulation. The comparison of the BDT-discriminator distributions for the sum of all signal-depleted categories (bottom plot) shows that the statistical model, which consists of the simulated prediction and corresponding experimental uncertainties, is able to describe the observed data. In figures 7.5 and 7.6, the best-fit BDT-discriminator distributions are individually compared to the observed data for events with electrons (left columns) and events with muons (right columns). Again, simulated contributions are scaled to the best-fit result of the combined measurement with electron and muon events.

Figure 7.5 refers to the signal-enriched categories “2 jets, 1 btag” (top row) and “3 jets, 1 btag” (bottom row). The observed BDT-discriminator distributions are well described by the best-fit distributions in each individual category. While the background contributions are constrained by the left tails of the BDT-discriminator distributions, the right tails, which are enriched in signal events, allow for a precise determination of the signal contribution.

Figure 7.6 shows the BDT-discriminator distributions in the signal-depleted categories (“2 jets, 2 btags”, “3 jets, $\geq 2$ btags”, “$\geq 4$ jets, 1 btag”, and “$\geq 4$ jets, $\geq 2$ btags”), which are simultaneously used with the signal-enriched categories within the statistical inference. W-boson-plus-jets production has a significant contribution in the “2 jets, 2 btags” categories, however, the overall event yield is low in this category. Signal events populate the “2 jets, 2 btags” and “3 jets, $\geq 2$ btags” categories if the b-tagged jet from spectator-b-quark fragmentation is within the acceptance of this analysis. Best-fit and observed BDT-discriminator distributions in the “2 jets, 2 btags” and “3 jets, $\geq 2$ btags” categories (first and second rows) are in good agreement with each other within the statistical precision of observed and simulated distributions. The dominant contribution in each category is $t\bar{t}$ production. The good modeling of $t\bar{t}$ production is confirmed in the “$\geq 4$ jets, 1 btag” and “$\geq 4$ jets, $\geq 2$ btags” categories (third and fourth row), which have a high statistical precision.
7.1 Measurement of the Inclusive $t$-Channel Cross Section

Figure 7.4: Observed BDT-discriminator distributions in the signal-enriched categories “2 jets, 1 btag” (top-left plot) and “3 jets, 1 btag” (top-right plot), as well as for the sum of all signal-depleted categories (bottom plot). Simulated signal and background contributions, as well as QCD-multijet events, are scaled to the best-fit results of the combined measurement with electron and muon events.
Figure 7.5: Observed BDT-discriminator distributions in the signal-enriched categories “2 jets, 1 btag” (top row) and “3 jets, 1 btag” (bottom row). Events with electrons are shown in the left column, and events with muons are shown in the right column. Simulated signal and background contributions, as well as QCD-multijet events, are scaled to the best-fit results of the combined measurement with electron and muon events.
7.1 Measurement of the Inclusive $t$-Channel Cross Section

Figure 7.6: Observed $\text{BDT}$-discriminator distributions in the signal-depleted categories. Events with electrons (muons) are shown in the left (right) column. Simulated signal and background contributions, as well as QCD-multijet events, are scaled to the best-fit results of the combined measurement with electron and muon events.
7 Measurement Results

7.2 Measurement of the CKM-Matrix Element $|V_{tb}|$

The measurement of the inclusive $t$-channel-production cross section offers a unique access to the CKM-matrix element $|V_{tb}|$ without assuming CKM-matrix unitarity or a particular number of quark generations. The measurement principle is well known in the literature (cf. e.g. [102, 225, 226]).

The $t$-channel production cross section is proportional to the squared CKM-matrix element $|V_{tb}|$

$$\sigma_{t\text{-channel}} \propto |f_{V_L} \times V_{tb}|^2. \quad (7.4)$$

$f_{V_L}$ is a form factor that can effectively modify the strength of the electroweak interactions, e.g. in terms of new physics phenomena (cf. sec. 2.2.3).

The measurement of the $t$-channel cross section allows for two direct measurements of the CKM-matrix element $|V_{tb}|$. The first measurement corresponds to the strength of the left-handed vector coupling at the $Wtb$ vertex without any assumption on $f_{V_L}$, i.e. the measurement of $|f_{V_L} \times V_{tb}|$ with $|f_{V_L} \times V_{tb}| \geq 0$. The second measurement refers to the interpretation of the measured $t$-channel cross section in terms of SM couplings, i.e. $f_{V_L} \equiv 1$. Here, $|V_{tb}|$ is constrained to be in the interval $[0, 1]$.

For the unconstrained $|f_{V_L} \times V_{tb}|$ measurement, a prior probability that is uniformly distributed in $|f_{V_L} \times V_{tb}|$ in the interval $[0, \infty]$ is used. For the constrained measurement, the prior probability for $|f_{V_L} \times V_{tb}|^2$ is uniformly distributed in the interval $[0, 1]$, and zero outside this interval. A flat prior-probability distribution in $|f_{V_L} \times V_{tb}|^2$ means that the prior-probability distribution is flat in the measured $t$-channel cross section, and, therefore, flat for the Poisson means of $t$-channel events.

**Unconstrained $|f_{V_L} \times V_{tb}|$ Measurement**

The CKM-matrix element $V_{tb}$ is measured to be

$$|f_{V_L} \times V_{tb}| = \sqrt{\frac{\sigma_{t\text{-channel}}^{\text{meas.}}}{\sigma_{t\text{-channel}}^{\text{SM}}}} = 1.016 \pm 0.048 \text{ (meas.)} \pm 0.018 \text{ (theor.)}$$

$$= 1.016 \pm 0.051,$$

assuming a $t$-channel SM cross-section prediction of $64.6 \pm 3.4\% \text{ pb}$ [8], and assuming that $|V_{td}|, |V_{ts}| \ll |V_{tb}|$, i.e. assuming that the top quark predominantly decays into a $W$ boson and a $b$ quark. The measured value is well compatible with the expected value for three quark generations as inferred from unitarity of the CKM matrix ($V_{tb, \text{SM fit}} = 0.999146^{+0.000021}_{-0.000046}$) [11].

The first uncertainty in eq. (7.5) is the (symmetrized) uncertainty due to the cross-section measurement (cf. eq. (7.1)), which includes experimental and generator-related uncertainties. The second uncertainty is due to the uncertainty of the theoretical prediction in NNLO accuracy, which includes coherent variations of the renormalization and factorization scales by factors 0.5 and 2.0, and variations to the MSTW2008 NNLO PDF parametrization [75] at 90% CL.

It is assumed that the top quark predominantly decays into a $W$ boson and a $b$ quark, which is confirmed by measuring the top-quark-branching ratio $R = B(t \to Wb)/\sum_{q=d,s,b} B(t \to Wq)$ in data (cf. discussion in sec. 2.2.4). $|V_{tb}|$ can be inferred from the $R$ measurement in $t\bar{t}$ events.
with $R = \frac{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2}$, which, however, implies unitarity of the CKM matrix and three quark generations ($|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2 = 1$). $|V_{tb}| = 0.911^{+0.018}_{-0.017}$ is measured with $t\bar{t}$ events at $\sqrt{s} = 8$ TeV (CMS experiment [126]). $|V_{tb}| = 0.95 \pm 0.02$ (D0 experiment [124]) and $|V_{tb}| = 0.97 \pm 0.05$ (CDF experiment [122]) are obtained in $t\bar{t}$ events with $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. Here, the obtained result from the D0 collaboration is compatible with the SM prediction from the CKM matrix fit within 2.5 standard deviations.

In figure 7.7 the $|f_{V_L} \times V_{tb}|$ measurement of this analysis is compared with results from other experiments, center-of-mass energies, or production channels. Only measurements obtained in single-top-quark production, i.e. without assuming CKM-matrix unitarity and three generations are shown in fig. 7.7. This measurement is in agreement with previous measurements [19, 20, 22, 23, 94–99]. This analysis constitutes the most precise measurement of the strength of the left-handed vector coupling at the $Wtb$ vertex ($|f_{V_L} \times V_{tb}|$) without assuming CKM-matrix unitarity or three generations.

Figure 7.7: Direct measurements of the CKM-matrix element $|V_{tb}|$ without assuming CKM-matrix unitarity [19, 20, 22, 23, 94–99].
Constrained $|V_{tb}|$ Measurement

When assuming SM couplings, i.e. $f_{V_L} \equiv 1$ and $0 \leq |V_{tb}| \leq 1$, a lower bound on $|V_{tb}|$ of

$$|V_{tb}| > 0.910$$

(7.6)

is obtained at 95% CL from the posterior-probability distribution of $|V_{tb}|^2$ (fig. 7.8 left). In order to account for non-experimental uncertainties, two additional nuisance parameters are added to the statistical model. Each nuisance parameter has a prior probability according to a Gaussian distribution with mean 1 and standard deviation $\sigma_{sd}$. Both nuisance parameters have an effect on $t$-channel events only. One nuisance parameter accounts for the (symmetrized) generator-related uncertainties and has a standard deviation of $\sigma_{sd} = 5.5\%$. The other nuisance parameter accounts for the uncertainties of the SM-cross-section prediction ($\sigma_{sd} = 3.4\%$).

For this measurement, a prior probability is used that is uniformly distributed in $|V_{tb}|^2$ within the interval $[0, 1]$, and zero outside this interval (fig. 7.8, left plot, solid line). This prior-probability distribution is flat w.r.t. the number of $t$-channel events, and is the typical choice in direct $|V_{tb}|$ measurements from single-top-quark events, e.g. in ref. [7, 225].

However, one can argue that the prior believe should be rather flat in the fundamental SM parameter $|V_{tb}|$ instead of being flat in the observed number of $t$-channel events, which are proportional to $|V_{tb}|^2$. The dashed line in figure 7.8 (left plot) shows the posterior-probability distribution that is obtained with a prior-probability distribution that is flat in $|V_{tb}|$. It is obtained that the lower bound on $|V_{tb}|$ only slightly depends on the prior-probability distribution. The same conclusion is obtained for the unconstrained $|f_{V_L} \times V_{tb}|$ measurement.

As an alternative cross check to the Bayesian-interval estimation, $|V_{tb}|$ is estimated with the approach of Feldman and Cousins [208, 246]. Here, an ordering rule that is based on a likelihood ratio is used to construct confidence intervals (cf. [246]). $|V_{tb}|$ is obtained to be in the interval

$$0.914 < |V_{tb}| \leq 1$$

(7.7)

at 95% CL, which is compatible with the result from Bayesian inference. Figure 7.8 (right) shows the obtained 95% confidence interval as a function of the measured $|V_{tb}|$ value when assuming a Gaussian-distributed uncertainty of 5.06% (cf. eq. 7.5).
7.3 Measurement of the Ratio of Top-Quark-Production and Top-Antiquark-Production Cross Sections \( R = \sigma_t / \sigma_{\bar{t}} \)

The cross sections of \( t \)-channel top-quark production \( (\sigma_t) \) and top-antiquark production \( (\sigma_{\bar{t}}) \) are measured to be

\[
\sigma_t = 40.0 \pm 3.2 \text{ (exp.)} \pm^{+2.0}_{-3.0} \text{ (gen.)} \text{ pb} = 40.0^{+4.4}_{-4.4} \text{ pb} \\
\text{and} \sigma_{\bar{t}} = 23.6 \pm 2.8 \text{ (exp.)} \pm^{+1.3}_{-1.2} \text{ (gen.)} \text{ pb} = 23.6^{+3.1}_{-5.0} \text{ pb}. \tag{7.8}
\]

The first uncertainty corresponds to experimental uncertainties and the second uncertainty corresponds to generator-related uncertainties.

The measured values are compatible with the SM prediction of \( 41.92^{+1.59+0.83}_{-0.21-0.83} \text{ pb} \) for top-quark production and \( 22.65^{+0.59+0.68}_{-0.50-0.91} \text{ pb} \) for top-antiquark production \( \text{[8]} \). Here, the first uncertainty corresponds to the choice of renormalization and factorization scales, the second uncertainty corresponds to the parametrization of the PDFs. The SM prediction is calculated in approximate NNLO QCD accuracy \( \text{[8]} \).

The PDF sets include ABM11 in the 5-flavor scheme \( \text{[248]} \), CTEQ6M \( \text{[52]} \), CT10 \( \text{[51]} \), CT10w \( \text{[51]} \), HERAPDF15 \( \text{[247]} \), MSTW2008 \( \text{[75]} \), and NNPDF 2.3 \( \text{[249]} \). These PDF sets differ in their methodology, assumptions, and analyzed data\(^2\) to derive the PDFs (cf. sec. 2.2.1).

The ratio of \( t \)-channel top-quark and top-antiquark production cross sections is determined to be

\[
R_{\sigma_t/\sigma_{\bar{t}}} = 1.70^{+0.23}_{-0.20} \text{ (exp.)} ^{+0.25}_{-0.06} \text{ (gen.)} = 1.70^{+0.34}_{-0.21}. \tag{7.9}
\]

The measured value is in agreement with the SM prediction of \( 1.85^{+0.03+0.05}_{-0.03-0.04} \text{, which is calculated with the MSTW2008 NNLO PDFs \text{[75]} in approximate NNLO QCD accuracy \text{[8]} \). Here, the first uncertainty corresponds to the choice of renormalization and factorization scales, and the second uncertainty corresponds to the parametrization of the PDFs at \( 68\% \text{ CL} \)\(^3\).

In figure 7.9, the measured ratio of top-quark production to top-antiquark production for \( t \)-channel events is shown in comparison to theoretical predictions. The theoretical predictions are obtained for various PDF sets in NNLO QCD accuracy. The theoretical predictions agree with the measured charge-production ratio within uncertainties.

The PDF sets include ABM11 in the 5-flavor scheme \( \text{[248]} \), CTEQ6M \( \text{[52]} \), CT10 \( \text{[51]} \), CT10w \( \text{[51]} \), HERAPDF15 \( \text{[247]} \), MSTW2008 \( \text{[75]} \), and NNPDF 2.3 \( \text{[249]} \). These PDF sets differ in their methodology, assumptions, and analyzed data\(^2\) to derive the PDFs (cf. sec. 2.2.1).

The expected ratio \( R_{\sigma_t/\sigma_{\bar{t}}} \) is calculated with MCFM (v6.6) \( \text{[76]} \) in combination with the LHAPDF package (v5.9.1) \( \text{[242]} \). The top-quark mass is set to \( 172.5 \text{ GeV/c}^2 \). The 5-flavor scheme is used as the default.

\(^2\) The PDF uncertainties are quoted at \( 90\% \text{ CL} \) in ref. \( \text{[8]} \) and have been converted to \( 68\% \text{ CL} \).

\(^3\) In particular, the CT10w set \( \text{[51]} \) includes measurements of the lepton-charge asymmetry from W-boson production of the D0 experiment, which are not included in CT10. The lepton-charge asymmetry measurement mostly provides additional constraints to the up and down quark PDFs.
Figure 7.9: Measurement of the ratio of top-quark production to top-antiquark production in $t$-channel events in comparison to theoretical predictions. The predictions are calculated in NLO QCD accuracy with MCFM [76, 115] for various PDF sets [51, 52, 75, 247–249]. Uncertainties on these predictions include parametrizations of the PDFs at 68% CL, the choice of the renormalization and factorization scales, and differences between 4-flavor-scheme and 5-flavor-scheme calculations.

Uncertainties of the parametrizations of the PDFs are considered by applying the specific prescription of each fitting group (as implemented in LHAPDF). For NNPDF, the uncertainties are calculated according to eq. 158 of ref. [243], and the average of all 100 replica is used as the default value. For all other groups, the uncertainties are calculated according to eq. 43 in ref. [56]. Uncertainties due to the parametrization of the PDFs are calculated at 68% CL. Since the CTEQ collaboration provides parametrizations at 90% CL, the resulting uncertainties for the sets CTEQ6M, CT10, and CT10w are divided by 1.645, such that they correspond to 68% CL [55, 243].

Furthermore, the calculations are performed within the 4-flavor scheme with massive b quarks in the initial state ($m_b = 4.7\text{GeV}/c^2$, $m_c = 1.5\text{GeV}/c^2$). Dedicated 4-flavor-scheme PDF sets are used for CT10, CT10w, ABM11, and MSTW2008 PDF sets. Here, the scales are set according to ref. [115]. The obtained differences between calculations in the 4-flavor and 5-flavor scheme are rather small with up to 0.6%, except for the ABM11 set, in which the predicted ratio is 1.2% larger for the 4-flavor-scheme calculation than for the 5-flavor-scheme calculation. The differences between 4-flavor-scheme and 5-flavor-scheme calculations are taken as additional systematic uncertainties into account.
The statistical uncertainty due to the limited numerical precision of the calculation is 0.2%. Uncertainties due to the choice of the renormalization and factorization scales are included by simultaneously varying both scales ($Q = \mu_R = \mu_F$) by factors of $(2Q)^2$ and $(Q/2)^2$. The scale $Q$ is assumed to be fully correlated for top quark and top antiquark production. All mentioned uncertainties are summed up in quadrature in figure 7.9.
In this thesis, measurements of the inclusive $t$-channel single-top-quark-production cross section, the CKM-matrix element $V_{tb}$, and the ratio of $t$-channel top-quark-production and top-antiquark-production cross sections were presented. Proton-proton ($pp$) collisions with a center-of-mass energy of $\sqrt{s} = 7$ TeV were analyzed. These collisions were recorded with the CMS experiment at the particle-accelerator complex LHC, which is operated by CERN near Geneva, Switzerland. The analyzed data correspond to an integrated luminosity of $1.6 \text{ fb}^{-1}$.

This analysis used events with at least two jets and either an electron or muon. Signal and background processes were discriminated using Boosted Decision Trees (BDTs). Kinematic distributions of basic and composite physics objects, angular correlations among them, and observables that describe the event topology were studied as input for the BDT-classifier training. The most discriminating variable was the pseudo-rapidity $\eta$ of the "spectator" jet, which is produced in addition to the single top quark. This jet can be as close to the beamline as $|\eta| \approx 4.5$.

The kinematics of signal and background processes were modeled with event generators, except for QCD-multijet processes without top quarks. These processes were derived from data. A blinding procedure was used for the analysis of the inclusive cross section. In a first step, the modeling of the dominant background processes, top-quark-pair and W-boson-plus-jets production, was checked in dedicated data-control regions. Distributions of basic kinematic quantities, the top-quark reconstruction, the variables used for the BDT-classifier training, and the BDT-discriminator output were scrutinized. It was found that the simulated processes well describe the data in the signal-depleted phase space. In a second step, the signal region was unblinded. Similar checks as for the background processes were performed. A good agreement between simulation and data was found. In particular, no unphysical outliers were observed.

The signal cross section was simultaneously measured in twelve orthogonal categories. Each event was classified according to the flavor of the charged lepton ($e, \mu$), the number of jets, and the number of b-tagged jets. The charge of the electron or muon was used as an additional parametrization for the individual measurements of top-quark-production and top-antiquark-production cross sections, as well as their ratio. The categorization of events enhances the signal acceptance and improves the measurement by exploiting diversified signal-to-background ratios. Furthermore, signal-depleted categories were used to gain confidence in the background modeling and to constrain the background normalization "in-situ" in the statistical inference.

A Bayesian approach was used to infer the signal cross section from data. Particular emphasis was placed on the modeling of systematic uncertainties and the evaluation of their impact on the measurement. Systematic uncertainties were incorporated as additional nuisance parameters into the likelihood function. Marginalization was used to eliminate the nuisance parameters.

The dominant experimental uncertainties were the production cross sections of $W^\pm$ bosons, which are produced in association with b-flavored and c-flavored jets, and the statistical precision of simulated events and QCD-multijet events. Systematic uncertainties related to the
reconstruction of physics objects with the CMS detector were well understood and had a small impact on the measurements.

The dominant single systematic uncertainty was the modeling of the t-channel signal. The signal process was simulated with POWHEG BOX [60–62, 188] in NLO+PS QCD accuracy. The COMPHEP event generator [189, 190] was used for comparison.

The single-top-quark t-channel production cross section was measured to be

\[ \sigma_{t\text{-channel}} = 66.6^{+6.7}_{-6.2} \text{ pb.} \] (8.1)

The measured value is in agreement with the SM prediction of \( 64.6^{+2.6}_{-1.8} \text{ pb} \) in approximate NNLO-QCD accuracy [8] and in agreement with a measurement performed by the ATLAS experiment [22]. With a relative uncertainty of \( -9.3/ +10.1\% \), this measurement is significantly more precise than previous measurements in pp and pp collisions [19, 20, 22, 100]. The “strength” of the electroweak coupling at the Wtb vertex was measured to be

\[ |f_{VL} \times V_{tb}| = 1.016 \pm 0.051, \] (8.2)

in which \( f_{VL} \) is a general form factor. It was assumed that the top quark predominantly decays into a W boson and b quark, i.e. \( |V_{td}|, |V_{ts}| \ll |V_{tb}| \). In terms of SM electroweak couplings \( (f_{VL} \equiv 1) \), the absolute value of the CKM-matrix element \( V_{tb} \) was determined to be

\[ |V_{tb}| = 1.016 \pm 0.051. \] (8.3)

With a relative uncertainty of \( \pm 5.0\% \), this is the most precise single direct measurement of \( |V_{tb}| \) to date. In particular, no assumptions were made on the unitarity of the CKM matrix or the number of quark generations. If \( 0 \leq |V_{tb}| \leq 1 \) is assumed in addition,

\[ |V_{tb}| > 0.910 \] (8.4)

was obtained at 95% confidence level. The measured values are compatible with previous direct measurements [19, 20, 22, 23, 94–99] and with the predicted value of \( V_{tb, \text{SM fit}} = 0.999146^{+0.000021}_{-0.000021} \) which was indirectly inferred using unitarity of the CKM matrix with three quark generations [11].

The ratio of t-channel top-quark-production and top-antiquark-production cross sections was measured to be

\[ R_{\sigma_{t}/\sigma_{\bar{t}}} = 1.70^{+0.34}_{-0.21}. \] (8.5)

The measurement is in agreement with SM predictions, which depend on the parametrizations of the PDFs and are between 1.84 and 2.11 for calculations in NLO+QCD accuracy [76, 113] with recent PDFs [51, 52, 75, 247–249]. Furthermore, this measurement is compatible with a previous measurement of the ATLAS experiment [17]. This is the first measurement of \( R_{\sigma_{t}/\sigma_{\bar{t}}} \) in proton-proton collisions at \( \sqrt{s} = 7 \text{ TeV} \) that is performed by the CMS experiment.

The results of the inclusive t-channel cross-section measurement were pre-published in [24]. This measurement was combined with two other analyses [25, 27] using the method of the “best linear unbiased estimate” [24, 250]. Their combination yields \( \sigma_{t\text{-channel}} = 67.2 \pm 6.1 \text{ pb} \) and \( |f_{VL} \times V_{tb}| = 1.020 \pm 0.049 \) with a \( p \)-value of 0.90 [24]. All three individual measurements are in excellent agreement with each other.

The experimental resolution of the t-channel-cross-section measurement already challenges the theoretical predictions. Future measurements are expected to benefit from the analysis strategy, the statistical model, and the treatment of systematic uncertainties that were elaborated in this thesis. With increased data, measurements of the t-channel single-top-quark-production cross section with a precision well below 10% become feasible at the LHC.
Bibliography


http://cdsweb.cern.ch/record/1633190

arXiv:1205.5764 [hep-ex]

http://cdsweb.cern.ch/record/1600799

http://cdsweb.cern.ch/record/1563135

http://cdsweb.cern.ch/record/1478935

http://cdsweb.cern.ch/record/1478371

http://cdsweb.cern.ch/record/1601029


arXiv:1110.4592 [hep-ex]


http://cdsweb.cern.ch/record/1601800


Bibliography


Bibliography


[220] C. J. Clopper and E. S. Pearson, “The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial”, *Biometrika* 26 (1934) 404.


List of Acronyms

AdaBoost – Adaptive Boost: An algorithm used to improve machine-learning techniques.
ALICE – A Large Ion Collider Experiment: General-purpose heavy-ion detector at the LHC.
ATLAS: Multi-purpose detector at the LHC.
AUC – Area under the Receiver-Operating-Characteristic curve: In case of machine-learning techniques, the AUC is used as an indicator to evaluate the performance of a particular algorithm or setup.
BDT – Boosted Decision Tree: A machine-learning technique.
CASTOR – Centauro And Strange Object Research: Calorimeter that is part of the CMS detector.
CDF – Collider Detector at Fermilab: Multi-purpose detector at the Tevatron.
CERN – Conseil Européen pour la Recherche Nucléaire: in English “European Council for Nuclear Research”
CHS – Charged Hadron Subtraction: Algorithm that is used to subtract Pile-Up contributions from reconstructed physics objects, e.g. jets.
CKM – Cabibbo-Kobayashi-Maskawa: The CKM matrix is a (complex) $3 \times 3$ matrix that relates weak and mass eigenstates. The squared absolute matrix elements describe the transition probability of quark flavor $i$ to $j$ via weakcharged currents.
CL – Confidence level
CMS – Compact Muon Solenoid: Multi-purpose detector at the LHC.
CMSSW – CMS Software
CP – Charge conjugation parity
CSC – Cathode Strip Chamber: Sub-detector in the end-cap disks of the CMS detector muon system.
DT – Drift Tube: Sub-detector used in the barrel part of the CMS detector muon system.
ECAL – Electromagnetic Calorimeter: Sub-detector system of the CMS detector.
FCNC – Flavor-changing neutral current
GSF – Gaussian-Sum Filter: Generalization of the Kalman Filter that is able to cope with non-Gaussian probability-density functions.
GZK – Greisen-Szepin-Kusmin: The GZK limit refers to an upper energy limit of cosmic rays [141, 142].
HCAL – Hadron Calorimeter: Sub-detector system of the CMS detector.
HERA – Hadron-Elektron-Ring-Anlage: Former collider at Deutsches Elektronen-Synchrotron (DESY) that was primarily used to study the structure of protons and its constituents.
HF – Hadron-Forward Calorimeter: Sub-detector of the CMS Hadron Calorimeter in the forward region of the CMS detector.
HFS – Hadronic final state: The HFS is the sum of four-vectors of all final-state jets in an event.
### List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLT</td>
<td>High Level Trigger: The highest trigger level of the CMS trigger system.</td>
</tr>
<tr>
<td>ISR</td>
<td>Initial-State Radiation: Gluon emission in the initial state of a hard-scattering process.</td>
</tr>
<tr>
<td>KS</td>
<td>Kolmogorov–Smirnov: The KS test is a statistical test.</td>
</tr>
<tr>
<td>L1</td>
<td>Level-1 Trigger: The lowest trigger level of the CMS trigger system.</td>
</tr>
<tr>
<td>LEP</td>
<td>Large Electron-Positron Collider: Former collider at CERN with a center-of-mass energy of up to $\sqrt{s} = 209$ GeV.</td>
</tr>
<tr>
<td>LHC</td>
<td>Large Hadron Collider: A hadron collider at CERN that is located at the Franco-Swiss border near Geneva.</td>
</tr>
<tr>
<td>LHCb</td>
<td>Large Hadron Collider beauty: Experiment at the LHC.</td>
</tr>
<tr>
<td>LHCf</td>
<td>Large Hadron Collider forward: Experiment at the LHC.</td>
</tr>
<tr>
<td>LO</td>
<td>Leading order</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov-chain-Monte-Carlo</td>
</tr>
<tr>
<td>n.d.o.f.</td>
<td>Number of degrees of freedom</td>
</tr>
<tr>
<td>NLO</td>
<td>Next-To-Leading Order</td>
</tr>
<tr>
<td>NNLL</td>
<td>Next-To-Next-To-Leading-Logarithmic Order</td>
</tr>
<tr>
<td>NNLO</td>
<td>Next-To-Next-To-Leading Order</td>
</tr>
<tr>
<td>PDF</td>
<td>Parton-Distribution Function</td>
</tr>
<tr>
<td>PF</td>
<td>Particle Flow: Algorithm, which is implemented in CMSSW that provides a list of reconstructed particles in an event similar to the list obtained from simulation truth.</td>
</tr>
<tr>
<td>PS</td>
<td>Preshower: Sub-detector system of the CMS detector.</td>
</tr>
<tr>
<td>PS</td>
<td>Parton-Showers: A parton shower describes the evolution of highly energetic partons down to low energy scales. PS models include entire particle cascades by subsequent gluon emissions and splittings into quark-antiquark pairs.</td>
</tr>
<tr>
<td>PXL</td>
<td>Physics eXtension Library: Software library for high-energy-physics analyses that is part of the VISPA package.</td>
</tr>
<tr>
<td>QCD</td>
<td>Quantum Chromodynamics</td>
</tr>
<tr>
<td>QED</td>
<td>Quantum Electrodynamics</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver operating characteristic: In case of machine-learning techniques, a ROC is used to visualize and evaluate the performance of a particular algorithm or setup.</td>
</tr>
<tr>
<td>RPC</td>
<td>Resistive Plate Chamber: Sub-detector of the CMS detector muon system which is used in both the barrel end-cap regions.</td>
</tr>
<tr>
<td>SM</td>
<td>Standard Model: The Standard Model of particle physics is the theory in which the fundamental particles, the fundamental interactions (except gravity), and the mass-generation mechanism are described.</td>
</tr>
<tr>
<td>SUSY</td>
<td>Supersymmetry</td>
</tr>
<tr>
<td>TCHE</td>
<td>Track Counting High Efficiency: Algorithm to identify jets originating from b quarks. The algorithm uses the second highest impact parameter significance of all tracks of a jet.</td>
</tr>
<tr>
<td>TCHP</td>
<td>Track Counting High Purity: Algorithm to identify jets originating from b quarks. The algorithm uses the third highest impact parameter significance of all tracks of a jet.</td>
</tr>
<tr>
<td>TEC</td>
<td>Tracker EndCaps: Sub-detector component of the silicon-strip tracker.</td>
</tr>
<tr>
<td>TIB</td>
<td>Tracker Inner Barrel: Sub-detector component of the silicon-strip tracker.</td>
</tr>
<tr>
<td>TID</td>
<td>Tracker Inner Disk: Sub-detector component of the silicon-strip tracker.</td>
</tr>
</tbody>
</table>
List of Acronyms

TMVA – Toolkit for Multivariate Data Analysis: Toolkit with implementations of various machine-learning techniques.

TOB – Tracker Outer Barrel: Sub-detector component of the silicon-strip tracker.

TOTEM – Total Elastic and Diffractive Cross Section Measurement: Experiment at the LHC

VdM – Van der Meer

VEV – Vacuum-Expectation value


WIMP – Weakly Interacting Massive Particle
Contributions to Previous Publications
(Vorveröffentlichungen)

A previous iteration of the inclusive-cross-section measurement with data corresponding to an integrated luminosity of 36 pb$^{-1}$ was published in ref. [1] and [2]. These publications describe the evidence for $t$-channel single-top-quark production at the LHC using two distinct analysis methods, one of them a BDT analysis, which drives the precision of the result. The BDT analysis (incl. figures) was developed and performed by myself, I also contributed revisions to the text. This previous analysis is also described in an internal analysis note of the CMS experiment [3]. My contributions to ref. [3] include the text, figures, and data analysis, except for the text in the subsections about “lepton trigger, identification and isolation efficiencies from Tag and Probe” and the description of the QCD-multijet estimation.

The analysis of the $t$-channel cross-section was significantly improved afterwards. Among others, more data was analyzed (1.6 fb$^{-1}$), the categorisation of events was introduced, the BDT-classifier training was improved, the statistical model extended, all systematic uncertainties were revisited and updated, a new QCD-multijet modeling was introduced. A summary of analysis strategy and results was published in ref. [4]. I contributed the BDT analysis, figures related to the BDT analysis, and significant revisions to the text. In particular, this publication contains modified versions of fig. 7.5 (top row) and 7.4 (bottom row).

I presented the analysis strategy and results of the inclusive cross-section measurement at the conference “TOP 2012, 5th International Workshop on Top Quark Physics” in Winchester, United Kingdom. I summarized my contribution in ref. [5]. This publication also contains a modified version of figure 7.5 (top right).

A previous iteration of the analysis presented in this thesis is also documented in an internal analysis note of the CMS experiment [6]. I contributed the text, figures, and data analysis to ref. [6], except for the text in the sections “Overview” and “Determination of the Cross Section”, which have been revised by me. In particular, ref. [6] contains modified versions of figures presented in chapter 5, fig. 7.5 and fig. 7.6.

Danksagung (Acknowledgments)

Mein besonderer Dank gilt Herrn Prof. Dr. Martin Erdmann für die exzellente Betreuung meiner Arbeit. Ich habe die wöchentlichen Arbeitsgruppentreffen immer sehr genossen.

Herrn Prof. Dr. Thomas Hebbeker danke ich dafür, dass er sich bereit erklärt hat, das Zweitgutachten zu übernehmen.


Ich danke insbesondere Prof. Dr. Martin Erdmann, Matthias Komm, Dr. Jan Steggemann und David Walz für das Korrekturlesen dieser Arbeit.

Special thanks go to the single-top-quark group of the CMS experiment. I would like to thank the group and sub-group convenors for their support, supervision, and helpful advices: Dr. Roberto Chierici, Prof. Dr. Andrea Giammanco, Dr. Luca Lista, Dr. Frank-Peter Schilling, Prof. Dr. Roberto Tenchini, and Dr. Jeannine Wagner-Kuhr. Moreover, I would like to thank all group members I closely collaborated with: Dr. Rebeca Gonzalez Suarez, Dr. Orso Iorio, Dr. Jyothsna Komaragiri, Dr. Daniel Martschei, Dr. Mario Merola, Dr. Thomas Peiffer, Andrey Popov, Dr. Jochen Ott, and Steffen Röcker. I highly appreciated our weekly meetings and I enjoyed the friendly welcome each time I was at CERN.

Prof. Dr. Wolfgang Wagner danke ich für die interessanten Diskussionen über Single-Top-Quark Physik auf diversen Konferenzen. Diese haben meinen Blick für das Thema über das CMS Experiment hinaus geschärft.

Many thanks to Dr. Peter Kreuzer and his team. I always enjoyed the CMS-computing shifts, which provided interesting insights into the computing infrastructure of the CMS experiment.

Ich danke dem gesamten Team des III. Physikalischen Instituts, insbesondere den Teams Grid-Computing und EDV (stellvertretend erwähnt seien Michael Bontenackels, Dr. Thomas Kreß, Dr. Andreas Nowack), Adriana Del Piero, Dr. Markus Merschmeyer, Melanie Roder und Iris Rosewick. Eure Unterstützung war immer tatkräftig und zuvorkommend.

Meiner Lebensgefährtin Linda Buchmüller, meinen Freunden und meiner Familie danke ich für ihre großartige Unterstützung und ihr unbegrenztes Vertrauen.