ETP-KA/2018-10

Single Top Quark Production at 13 TeV with the CMS Experiment: from Rediscovery to Search for Rare Channels and Determination of Higgs Boson Couplings

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
von der Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

Dipl.-Phys. Nils Faltermann
aus Kandel

Mündliche Prüfung: 20. Juli 2018

Referent: Prof. Dr. Th. Müller
Institut für Experimentelle Teilchenphysik

Korreferent: Prof. Dr. U. Husemann
Institut für Experimentelle Teilchenphysik

Introduction

With the observation of the Higgs boson in 2012 [1, 2], the last missing particle of the standard model of particle physics was discovered. The standard model has successfully described the subatomic world for half a century, but also predicted new particles, which have been discovered afterwards at particle colliders. Although the standard model is a complete and self-consistent theory, there are unexplained phenomena in nature, whose existence let one come to the conclusion that the standard model is not the final answer to all fundamental questions in particle physics. Many extensions of the standard model have been predicted, but only experimental evidence will finally decide if one of these theories is realized in nature.
The purpose of the Large Hadron Collider (LHC), the biggest machine ever built by mankind, is to probe the standard model up to the highest energies ever achieved under laboratory conditions. This is accomplished by accelerating two reverse proton beams up to beam energies of 6.5 TeV and cross their paths at certain interaction points to induce particle collisions. Sophisticated multi-purpose particle detectors, such as the Compact Muon Solenoid (CMS) experiment, are located at crossing points of the beams to record the signatures of each proton-proton collision.
With a mass of roughly the same as a gold atom, the top quark is the heaviest fundamental particle in the standard model. Although mainly produced in top quark-antiquark pairs at the LHC through the strong interaction, the production of single top quarks allows to probe the electroweak sector of the standard model. It is not only possible to measure observables of the electroweak theory with single top quark processes at high precision, but also to search for deviations in predictions and data that could be a hint for physics beyond the standard model. Due to its high mass, the top quark is also an excellent candidate to search for associated production with a Higgs boson, in which the coupling of the Higgs boson to the top quark and other fundamental particles can be determined.
After the first data-taking period from 2010 to 2012 at center-of-mass energies of 7 and 8 TeV and a two-year shutdown for upgrades of the accelerator, the LHC started operations again in 2015 with Run II at 13 TeV center-of-mass energy. Since then, the LHC has surpassed its design goal in terms of collision intensity and provided more than a quadrillion proton-proton collisions, from which only a tiny fraction is of interest for single top quark research. This thesis follows the journey of the LHC Run II and the data recorded by the CMS detector from the first collisions in 2015 up to the beginning of the precision era at the end of 2017.

In the first chapter, the theoretical foundation of the standard model is provided. The chapter starts with a general introduction, with the focus shifted afterwards on the physics of the top quark and the Higgs boson, as well as their interplay.
The second chapter introduces the statistical methods employed for the different analyses in this thesis. In the first part, sophisticated multivariate analysis techniques are outlined that help

to identify signal candidates out of background-dominated data. The second part explains the underlying statistical reasoning behind the results.
The description of the experimental apparatus, namely the particle collider LHC and the CMS detector, are subjects of interest in the third chapter. The journey of the proton is followed from a simple hydrogen bottle to the collision point and from the decay products of the collisions to the signals in the readout electronics.
The fourth chapter first reveals how simulated proton-proton collisions are generated to describe the measured data. In the second part of the chapter, the reconstruction procedure is introduced, which is used to reassemble physics objects from detector signals for simulation and data alike.
The purpose of the fifth chapter is to describe a common physics object definition, shared between the different analyses presented in this thesis. In addition, certain quality criteria are defined that ensure proper modeling of the measured data by the predictions.
In the sixth chapter, a measurement of the cross section of single top quark production in the t channel is presented. The measurement is based on the first year of proton-proton collision data at a center-of-mass energy of 13 TeV in 2015. The t channel is the most dominant production mode of single top quarks at the LHC and is therefore well suited for a first single top quark measurement at Run II of the LHC.
With the increased amount of data recorded during Run II, the focus shifts to more rare single top quark production channels. One of these channels is the s-channel single top quark production, the only major single top quark production mode yet unobserved at the LHC. The seventh chapter is dedicated to the search for this channel. To increase the sensitivity of this search, the combined data sets of 2016 and 2017 are used.
In the eighth chapter of this thesis, a search for the associated production of a Higgs boson with a single top quark is presented. This associated production mode is highly sensitive to the coupling of the Higgs boson to the top quark and to vector bosons.
The final chapter summarizes the three different single top quark analyses, presents a conclusion and provides an outlook for the promising field of single top quark research.

Contents

1 Theoretical Foundation 1
- 1.1 The Standard Model of Particle Physics . . . 1
  - 1.1.1 Fermions . . . 1
  - 1.1.2 Bosons . . . 2
- 1.2 Theoretical Description . . . 3
  - 1.2.1 Electromagnetic Interaction . . . 4
  - 1.2.2 Weak Interaction . . . 5
  - 1.2.3 Electroweak Symmetry Breaking and the Higgs Mechanism . . . 5
  - 1.2.4 Strong Interaction . . . 8
- 1.3 From Partons to Particles . . . 10
- 1.4 The Top Quark . . . 10
  - 1.4.1 Production . . . 11
  - 1.4.2 Decay . . . 14
- 1.5 The Higgs Boson . . . 15
  - 1.5.1 Production . . . 15
  - 1.5.2 Decay . . . 16
  - 1.5.3 Couplings . . . 17
- 1.6 Associated Production of a Higgs Boson with Top Quarks . . . 18
- 1.7 Indications for Physics beyond the Standard Model . . . 21

2 Statistical Analysis 25
- 2.1 Multivariate Analysis . . . 25
  - 2.1.1 Boosted Decision Trees . . . 25
  - 2.1.2 Artificial Neural Networks . . . 27
  - 2.1.3 Deep Neural Networks . . . 29
  - 2.1.4 Training and Validation . . . 30
- 2.2 Statistical Methods . . . 31
  - 2.2.1 Maximum Likelihood Estimation . . . 31
  - 2.2.2 Systematic Uncertainties . . . 32
  - 2.2.3 Hypothesis Testing and Exclusion Limits . . . 33
  - 2.2.4 Statistical Significance . . . 35

3 The Large Hadron Collider and the Compact Muon Solenoid Experiment 37
- 3.1 The Large Hadron Collider . . . 37

3.2 The Compact Muon Solenoid . . . 40
3.2.1 Tracking System . . . 42
3.2.2 Electromagnetic Calorimeter . . . 43
3.2.3 Hadron Calorimeter . . . 45
3.2.4 Superconducting Solenoid . . . 45
3.2.5 Muon System . . . 46
3.2.6 Trigger System and Data Acquisition . . . 48
3.2.7 Computational Infrastructure . . . 49

4 Event Simulation and Reconstruction 51
4.1 Simulation . . . 51
4.1.1 Hard Scattering . . . 51
4.1.2 Parton Shower . . . 54
4.1.3 Processes with Bottom Quarks in the Initial State . . . 54
4.1.4 Hadronization . . . 55
4.1.5 Underlying Event and Pileup . . . 55
4.1.6 Monte Carlo Event Generators . . . 55
4.1.7 Detector Simulation . . . 57
4.2 Reconstruction with the Particle-Flow Algorithm . . . 57
4.2.1 Tracks . . . 57
4.2.2 Vertices . . . 58
4.2.3 Muons . . . 59
4.2.4 Electrons . . . 59
4.2.5 Photons, Charged Hadrons and Neutral Hadrons . . . 60
4.2.6 Jets . . . 60
4.2.7 Missing Transverse Momentum . . . 66

5 Preselection of Physics Objects and Additional Reconstruction 69
5.1 Physics Objects . . . 69
5.1.1 Primary Vertex . . . 69
5.1.2 Muons . . . 69
5.1.3 Electrons . . . 72
5.1.4 Jets . . . 74
5.1.5 b Tagging . . . 77
5.2 W Boson Reconstruction . . . 78

6 Measurement of the *t*-Channel Single Top Quark Production Cross Section at $\sqrt{s}$ = 13 TeV 81
6.1 Analysis Strategy . . . 82
6.2 Event Topology . . . 84
6.2.1 Signal Process . . . 84
6.2.2 Background Processes . . . 84
6.3 Event Selection . . . 87
6.4 Simulation of Events and Corrections . . . 89

6.4.1 Number of Pileup Interactions . . . 89
6.4.2 Muon Efficiencies . . . 89
6.4.3 b Tagging Efficiencies . . . 90
6.5 Background Estimation for QCD Multijet Events . . . 91
6.6 Evidence for Single Top Quark Production at $\sqrt{s} = 13$ TeV with the First Data in 2015 . . . 93
6.7 Analysis with the Full Data Set of 2015 . . . 100
6.7.1 QCD Multijet Background Estimation . . . 100
6.8 Multivariate Analysis . . . 103
6.9 Signal Extraction . . . 108
6.10 Systematic Uncertainties . . . 112
6.10.1 Experimental Uncertainties . . . 112
6.10.2 Theoretical Uncertainties . . . 113
6.10.3 Impact of Systematic Uncertainties . . . 114
6.11 Results . . . 115
6.12 Summary and Outlook . . . 116

7 Search for *s*-Channel Single Top Quark Production at $\sqrt{s} = 13$ TeV 121
7.1 Analysis Strategy . . . 122
7.2 Event Topology . . . 123
7.2.1 Signal Process . . . 123
7.2.2 Background Processes . . . 124
7.3 Event Selection . . . 126
7.3.1 Electron Cross Trigger . . . 128
7.3.2 b Tagging Improvements . . . 128
7.4 Simulations of Events and Corrections . . . 129
7.4.1 Number of Pileup Interactions . . . 132
7.4.2 Lepton Efficiencies . . . 132
7.4.3 b Tagging Efficiencies . . . 132
7.5 Background Estimation for QCD Multijet Events . . . 133
7.5.1 Muon Channel . . . 133
7.5.2 Electron Channel . . . 134
7.6 Top Quark Reconstruction . . . 135
7.7 Event Classification . . . 136
7.8 Systematic Uncertainties . . . 144
7.8.1 Experimental Uncertainties . . . 144
7.8.2 Theoretical Uncertainties . . . 145
7.9 Results . . . 147
7.9.1 2016 Data . . . 147
7.9.2 2017 Data . . . 147
7.9.3 Combination of 2016 and 2017 Data . . . 147
7.10 Summary and Outlook . . . 152

8 Search for tH Production with H → $b\bar{b}$ at $\sqrt{s}$ = 13 TeV and Study of Higgs Boson Couplings 153
8.1 Analysis Strategy . . . 154
8.2 Event Topology . . . 156
8.2.1 Signal Processes . . . 157
8.2.2 Background Processes . . . 158
8.3 Event Selection . . . 160
8.3.1 Single Lepton Region . . . 161
8.3.2 Dileptonic Region . . . 164
8.4 Simulation of Events and Corrections . . . 165
8.4.1 Number of Pileup Interactions . . . 167
8.4.2 Lepton Efficiencies . . . 167
8.4.3 b Tagging Efficiencies . . . 168
8.5 Event Reconstruction . . . 171
8.5.1 Reconstruction under the tHq Hypothesis . . . 173
8.5.2 Reconstruction under the tHW Hypothesis . . . 176
8.5.3 Reconstruction under the Top Quark Pair Hypothesis . . . 179
8.6 Event Classification . . . 181
8.6.1 Single Lepton Region . . . 181
8.6.2 Dileptonic Region . . . 186
8.7 Systematic Uncertainties . . . 191
8.7.1 Experimental Uncertainties . . . 191
8.7.2 Theoretical Uncertainties . . . 192
8.8 Results and Interpretation . . . 193
8.8.1 tH Production . . . 193
8.8.2 κ_t/κ_V Coupling Ratios . . . 198
8.8.3 CP-Mixing . . . 198
8.9 Combination of the Results with Different Higgs Boson Decay Modes . . . 202
8.10 Summary and Outlook . . . 203

9 Conclusion 205

A Appendix: Measurement of the *t*-Channel Single Top Quark Production Cross Section at $\sqrt{s}$ = 13 TeV 209
A.1 Data Sets and MC Simulation Samples . . . 209
A.2 Input Variables in the 2-Jets-1-Tag Region . . . 211
A.3 Input Variables in the 3-Jets-1-Tag Region . . . 214
A.4 Input Variables in the 3-Jets-2-Tags Region . . . 218

B Appendix: Search for *s*-Channel Single Top Quark Production at $\sqrt{s}$ = 13 TeV 223
B.1 Data Sets and MC Simulation Samples . . . 223
B.2 Corrections to Simulated Events . . . 226
B.2.1 Number of Pileup Interactions . . . 226
B.2.2 Lepton Efficiencies . . . 227

B.2.3 b Tagging Efficiencies . . . 229
B.3 QCD Estimation in the 2-jets–1-tag Control Region . . . 230
B.4 QCD Estimation in the 3-jets–1-tag Control Region . . . 231
B.5 Input Variables for the DNN . . . 232
B.6 Fit Result with 2016 Data . . . 238
B.7 Fit Result with 2017 Data . . . 239
B.8 Pulls in the Combined Fit . . . 240

C Appendix: Search for tH Production with H → $b\bar{b}$ at $\sqrt{s}$ = 13 TeV and Study of Higgs Boson Couplings 241
C.1 Data Sets and MC Samples . . . 241
C.2 Reweighting of Final-State Object Kinematics . . . 243
C.3 Two-dimensional Distributions of Final-State Object Kinematics . . . 245
C.4 Settings for the Training of BDTs . . . 247
C.5 Input Variables for the tHq Hypothesis . . . 248
C.6 Input Variables for the tHW Hypothesis . . . 251
C.7 Input Variables for the $t\bar{t}$ Hypothesis . . . 254
C.8 Input Variables for the Event Classification in the Single-Lepton Region . . . 256
C.9 Input Variables for the Event Classification in the Dileptonic Region . . . 262

List of Figures 265

List of Tables 271

Bibliography 273

1. Theoretical Foundation

The concept of science is a cycle of observation, conclusion, theory and prediction. No prediction can be made without a proper theory, which not only describes the already known, but can also predict yet undiscovered phenomena. In case of elementary particles, this theory is the standard model (SM) of particle physics. This chapter gives an overview of the different phenomena described by the SM of particle physics and also provides the current experimental status for comparison. For the sake of simplicity, natural units are used throughout this thesis, where $\hbar = c = 1$.

1.1. The Standard Model of Particle Physics

The SM was developed in the first part of the second half of the 20th century and was since then very successful in describing elementary particles and their interactions. Furthermore, all predicted particles have been experimentally observed at particle colliders afterwards, with the Higgs boson in 2012 as last missing particle [1, 2]. The particles in the SM are classified into two groups depending on their spin. All particles with half-integer spin are called fermions and all particles with integer spin are called bosons.

1.1.1. Fermions

The SM predicts twelve fermions in total, which are equally divided into two groups: quarks and leptons. Furthermore, all fermions can also be arranged based on their mass into three generations, each containing two quarks (up-type and down-type) and two leptons (electrically charged and uncharged). Particles inside the same generation are divided by the value of their weak isospin. The quarks and leptons of the SM are shown in Table 1.1 and Table 1.2, respectively, along with some of their properties.
The quarks of the first generation are the up quark (u) and down quark (d), the second generation includes the charm quark (c) and the strange quark (s) and the third generation is built up from the top quark (t) and the bottom quark (b). Quarks are not observed as free particles in nature, instead most of them form bound states with each other, called hadrons. Since fermions of higher generations are unstable, the matter of everyday life, such as the electron, the proton (uud) and the neutron (udd), consists of fermions of the first generation. For every fermion in the SM there exists a corresponding antifermion with the same properties, but opposite electric' charge. Quarks are the only particles in the SM that couple to all known forces. The group of leptons consists of the electron (e), the muon (μ), the tau (τ) and for each of them a corresponding neutrino, the electron neutrino (ν_e), muon neutrino (ν_μ) and tau neutrino (ν_τ). Neutrinos are massless in the SM by construction, but the mass of the other fermions range

Table 1.1.: The six different quarks of the SM, listed with their mass, electric charge and weak isospin, make up half of the fermions. Values are taken from Ref. [3]. The quarks are grouped into three different generations with increasing mass, each with an up- and down-type quark.

		Generation	Mass	Electric charge (e)	Weak isospin
up	(u)	1	$2.2^{+0.6}_{-0.4}$ MeV	$+^2/_3$	$+^1/_2$
down	(d)	1	$4.7^{+0.5}_{-0.4}$ MeV	$-^1/_3$	$-^1/_2$
charm	(c)	2	1.28 ± 0.03 GeV	$+^2/_3$	$+^1/_2$
strange	(s)	2	96^{+8}_{-4} MeV	$-^1/_3$	$-^1/_2$
top	(t)	3	173.1 ± 0.6 GeV	$+^2/_3$	$+^1/_2$
bottom	(b)	3	$4.18^{+0.04}_{-0.03}$ GeV	$-^1/_3$	$-^1/_2$

Table 1.2.: The second group of fermions in the SM are the six leptons. Similar to Table 1.1, their mass, electric charge and weak isospin are shown [3]. Similar to the quarks, the leptons are arranged into three different generations.

		Generation	Mass	Electric charge (e)	Weak isospin
electron neutrino	(ν_e)	1	< 225 eV	0	$+^1/_2$
electron	(e)	1	510.999 keV	-1	$-^1/_2$
muon neutrino	(ν_μ)	2	< 0.19 MeV	0	$+^1/_2$
muon	(μ)	2	105.658 MeV	-1	$-^1/_2$
tau neutrino	(ν_τ)	3	< 18.2 MeV	0	$+^1/_2$
tau	(τ)	3	1.776 GeV	-1	$-^1/_2$

over several orders of magnitude, starting from the electron with 511 keV up to the top quark with about 173 GeV.

1.1.2. Bosons

All bosons of the SM have spin 1 (vector bosons), with the exception of the already mentioned Higgs boson (spin 0). The spin-1 bosons are also called gauge bosons, as they mediate the different interactions of the SM between elementary particles. These are the photon (γ), the two W bosons (W) and the Z boson (Z), as well as the eight gluons (g), each of the three boson groups being the mediator of a different force. The photon transmits the electromagnetic interaction, which has an infinite range due to the photon being massless. In contrast to this, the weak interaction mediated by the W and Z bosons has only a short range because of their relatively high mass of about 80 GeV and 91 GeV, respectively. Although the gluons are massless, the interaction also has only a short range because of the self-interaction of gluons. Thus, the only fundamental force described by the SM that plays a role on a macroscopic scale is the electromagnetic interaction. The Higgs boson is not a mediator of a fundamental interaction,

Table 1.3.: Five different bosons are described by the SM. They are either massless or their mass is in the 100 GeV range. Except for the scalar Higgs boson (spin 0), all other bosons are vector bosons (spin 1) [3]. The vector bosons are a result of the possible interactions between elementary particles, while the scalar Higgs boson is evidence for the mass-generating mechanism described by the SM.

		Mass	Electric charge (e)	Spin
photon	(γ)	0	0	1
W bosons	($W^{\pm}$)	80.385 ± 0.015 GeV	± 1	1
Z boson	(Z)	91.188 ± 0.002 GeV	0	1
gluons	(g)	0	0	1
Higgs boson	(H)	125.09 ± 0.24 GeV	0	0

but instead the excitation of the Higgs field, which gives all particles of the SM their masses. An overview of all bosons in the SM is provided in Table 1.3.

1.2. Theoretical Description

As the SM is a quantum field theory (QFT), all particles are described as excitations of quantum fields ϕ. The dynamics of fields are defined by the principle of least action. The action of a system is defined as

$$S = \int \mathcal{L}(\phi, \partial_\mu \phi)\, \mathrm{d}^4 x, \tag{1.1}$$

where $\mathcal{L}$ is the Lagrangian density of the system (in the following simply called Lagrangian) and x is a four-vector in spacetime with the relation $\partial_\mu = \frac{\partial}{\partial x^\mu}$ ($\mu = 0, 1, 2, 3$). The principle of least action then states that variations of the action vanish, i. e., $\delta S = 0$. This condition is satisfied by the Euler-Lagrange equation

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0, \tag{1.2}$$

from which the equations of motion of the system can be derived.
As an example, the Lagrangian of a system of spin-less particles without interaction is defined by the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu - \frac{1}{2} m^2 \phi^2. \tag{1.3}$$

Applying the Euler-Lagrange equation to this particular Lagrangian yields

$$(\partial_\mu \partial^\mu - m^2)\phi = 0, \tag{1.4}$$

which is known as the Klein-Gordon equation. Similarly, from the Lagrangian of non-interacting spin-$^1/_2$ fermions

$$\mathcal{L}_{\text{Dirac}} = \overline{\psi}(\mathrm{i}\gamma^\mu \partial_\mu - m)\psi, \tag{1.5}$$

with ψ and $\overline{\psi} = \psi^\dagger \gamma^0$ being the spinor and adjoint spinor, respectively, and γ^μ the gamma matrices, the Dirac equation can be derived:

$$(\mathrm{i}\gamma^\mu \partial_\mu - m)\psi = 0. \tag{1.6}$$

The local invariance of the Lagrangian under certain gauge transformations results in interaction terms of particles.

1.2.1. Electromagnetic Interaction

The theory of quantum electrodynamics (QED) [4–6] describes the interaction between light, matter and all other electromagnetic phenomena. It is commonly referred to as the most precisely tested theory in the history of science, giving accurate predictions over several orders of magnitude [7, 8]. This huge success has made its approaches a baseline for any other QFT in the SM.
The theory of QED states that the Lagrangian should be invariant under a gauge transformation of the U(1) symmetry group, i. e., a phase of the type

$$\psi(x) \mapsto \psi'(x) = \mathrm{e}^{\mathrm{i}q\varphi(x)}\psi(x), \tag{1.7}$$

with the electric charge q and an arbitrary, but spacetime-dependent phase φ. Applying this transformation to Eq. 1.5 results in an additional term of the transformed Lagrangian:

$$\mathcal{L}(\psi', \partial_\mu \psi') = \mathcal{L}(\psi, \partial_\mu \psi) + q\overline{\psi}\gamma^\mu \psi \varphi(x). \tag{1.8}$$

To recover local gauge invariance, an additional term, representing a gauge field, has to be added to the Lagrangian. This is achieved by replacing the normal derivative with the covariant derivative

$$\partial_\mu \mapsto D_\mu = \partial_\mu - \mathrm{i}qA_\mu(x), \tag{1.9}$$

with A_μ as the four-potential and the gauge transformation

$$A_\mu(x) \mapsto A'_\mu(x) = A_\mu(x) + \partial_\mu \varphi(x). \tag{1.10}$$

Along with the Lagrangian of this so-called gauge field $\mathcal{L}_A$, the Lagrangian of QED can be written as

$$\begin{aligned}
\mathcal{L}_{\mathrm{QED}} &= \mathcal{L}_\psi + \mathcal{L}_A \\
&= \overline{\psi}(\mathrm{i}\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\
&= \overline{\psi}(\mathrm{i}\gamma^\mu \partial_\mu - m)\psi + q\overline{\psi}\gamma^\mu \psi A_\mu - \frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2.
\end{aligned} \tag{1.11}$$

The first term in Eq. 1.11 represents the fermion field propagator, as introduced by the Dirac Lagrangian (1.5), and the last term is the propagator of the gauge field, which corresponds to the electromagnetic field tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. The second term describes the interaction between both fields. Excitation of the gauge field are the photons with spin 1. They transmit the electromagnetic force between fermions, which is proportional to the electric charge q. In each vertex of electromagnetic interaction, the electric charge is conserved. The photon itself is massless by construction, as a mass term of any gauge field would have the non-gauge invariant form $m_A^2 A_\mu A^\mu$.

1.2.2. Weak Interaction

Similar to the theory of QED, the same approach can be used to describe the weak interaction between particles. The corresponding gauge group is the SU(2), as the weak interaction mediates between different particles changing their weak charge, called weak isospin. Therefore, the fermion field consists of two components

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}. \tag{1.12}$$

A covariant derivative is defined by

$$\begin{aligned} D_\mu &= \partial_\mu - \frac{1}{2} \mathrm{i} g W_\mu^j \sigma^j \\ &= \partial_\mu - \frac{1}{2} \mathrm{i} g \begin{pmatrix} W_\mu^3 & W_\mu^1 - \mathrm{i} W_\mu^2 \\ W_\mu^1 + \mathrm{i} W_\mu^2 & -W_\mu^3 \end{pmatrix}, \end{aligned} \tag{1.13}$$

with the weak charge g, $W_\mu^{1,2,3}$ the three gauge fields and σ^j the Pauli matrices. In contrast to the electromagnetic interaction, parity is not conserved [9]. Instead, parity is maximally violated, meaning that the weak interaction can only couple to one chiral eigenstate of the fermion field

$$\psi_{\mathrm{L,R}} = \frac{1}{2}(1 \mp \gamma^5)\psi, \tag{1.14}$$

where $\psi_{\mathrm{L,R}}$ are the left- and right-handed components of the field ψ. It turns out that the weak interaction only couples to left-handed fermions and right-handed antifermions. Mathematically this is realized by grouping left-handed fermions into isospin doublets and right-handed fermions into isospin singlets. An exception are the neutrinos, which do not exist as right-handed particles in the SM. Exemplary for the first generation of fermions this means:

$$\begin{pmatrix} \mathrm{e} \\ \nu_{\mathrm{e},} \end{pmatrix}_{\mathrm{L}}, \begin{pmatrix} \mathrm{u} \\ \mathrm{d} \end{pmatrix}_{\mathrm{L}}, \left(\mathrm{e}\right)_{\mathrm{R}}, \left(\mathrm{u}\right)_{\mathrm{R}}, \left(\mathrm{d}\right)_{\mathrm{R}}. \tag{1.15}$$

Combinations of W_1 and W_2 can be associated to the W bosons from the weak interaction:

$$W^\pm = \frac{1}{\sqrt{2}}(W_1 \mp \mathrm{i} W_2). \tag{1.16}$$

On the other hand, W_3 cannot be associated directly with the Z boson as there is no coupling to right-handed fermions allowed. Due to parity violation the theory is only gauge invariant if all fermions are massless. This circumstance is resolved in the following section.

1.2.3. Electroweak Symmetry Breaking and the Higgs Mechanism

One of the major accomplishments of the SM is the unification of the electromagnetic interaction and the weak interaction to the electroweak interaction [10]. The weak interaction adds two additional gauge bosons, namely the W bosons. However, in contrast to the massless photons,

the W bosons have non-zero masses, which is not allowed by gauge invariance. Furthermore, all fermions should be massless for the same reason. This is resolved by the Higgs mechanism, developed by Peter Higgs among others in 1964 [11–13], and the concept of electroweak symmetry breaking [14, 15].
The Higgs mechanism introduces a new complex scalar field ϕ with SU(2) symmetry:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{1.17}$$

The corresponding Lagrangian of this field can be written as

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \tag{1.18}$$

with the Higgs potential

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \frac{1}{2} \lambda (\phi^\dagger \phi)^2 \tag{1.19}$$

and the covariant derivative

$$D_\mu = \partial_\mu - \frac{1}{2} \mathrm{i} g' B_\mu - \frac{1}{2} \mathrm{i} g W_\mu^j \sigma^j. \tag{1.20}$$

Similar to QED, the field B_μ is the gauge field of a U(1) symmetry. The charge associated to this field is the weak hypercharge

$$Y = 2(Q - T_3), \tag{1.21}$$

where T_3 is the third component of the weak isospin. Depending on the value of the μ^2 term, the Higgs potential has either one global minimum at 0 ($\mu^2 \leq 0$) or a degenerate minimum in case of $\mu^2 > 0$ at

$$v = \sqrt{\frac{\mu^2}{\lambda}} \tag{1.22}$$

on a circle around the origin of the plane spanned by the real and imaginary components of the Higgs field. In this case, the minimum is at the so-called vacuum expectation value (VEV) $v = 246\,\text{GeV}$, which can be derived from measurements of the Fermi coupling constant G_F [16]. The form of the potential in case of $\mu^2 > 0$ is shown in Fig. 1.1. The Higgs mechanism allows the combination of the electromagnetic interaction from the theory of QED and the weak interaction. The symmetry group of this electroweak interaction is the combined $\text{SU(2)}_\text{L} \times \text{U(1)}_Y$ group with the gauge fields $W_{1,2,3}$ and B. This symmetry is spontaneously broken by the Higgs mechanism such that only a U(1)_Q symmetry remains. This is called electroweak symmetry breaking. This procedure introduces a mixture of the W_3 and B fields

$$\begin{pmatrix} A \\ Z \end{pmatrix} = \begin{pmatrix} \cos\theta_\text{W} & \sin\theta_\text{W} \\ -\sin\theta_\text{W} & \cos\theta_\text{W} \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix} \tag{1.23}$$

to a new field Z and the already known field A from QED. The angle is called the weak mixing angle or Weinberg angle $\sin^2\theta_\text{W} = 0.231$ [3]. The newly introduced Z boson associated to the Z field receives a mass of

$$m_\text{Z} = \frac{m_\text{W}}{\cos\theta_\text{W}}, \tag{1.24}$$

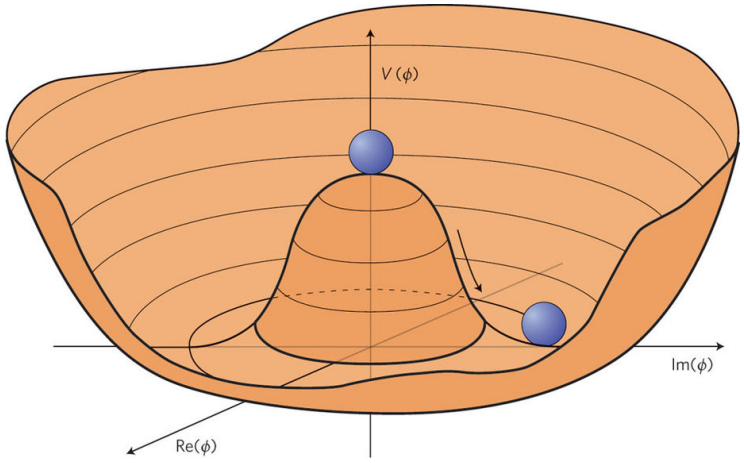


Figure 1.1.: The Higgs potential for the case of $\mu^2 > 0$, in which the potential is also referred to as *mexican hat* potential. The minimum of the potential is not at 0 (as it would be for $\mu^2 < 0$), but instead the potential has an infinite number of minima at $|\phi| = v$, the so-called vacuum expectation value. Figure taken from [17].

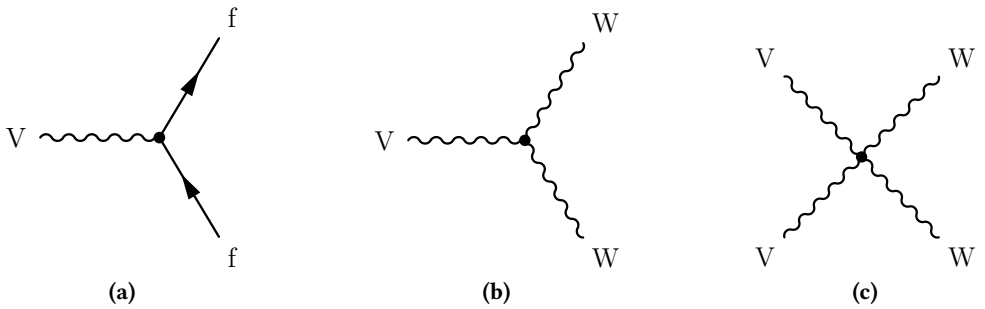


Figure 1.2.: The three basic electroweak interactions in the SM. A vector boson can either couple to a fermion-antifermion pair (a) or to two other vector bosons with triple vertex (b), or four vector bosons can interact through a quartic vertex (c).

while the photon remains massless. The basic interactions of all four gauge bosons are shown in Fig. 1.2 with Feynman diagrams.
Through this mechanism the weak gauge bosons receive their masses, but in case of quarks and leptons, the Higgs field couples through a Yukawa interaction with the corresponding fermion fields. In case of an electron, where no coupling to up-type fermions is allowed, the Lagrangian can simply be written as

$$\mathcal{L}_{\text{Yukawa}} = -y_e(\overline{\psi}_L \phi \psi_R + \overline{\psi}_R \phi \psi_L), \tag{1.25}$$

with the Higgs potential ϕ, the electron field $\psi_{L/R}$ for a left- and right-handed state, respectively, and the Yukawa coupling constant y_e for the electron. This results in a simple mass term for the electron:

$$m_e = \frac{y_e v}{\sqrt{2}}. \tag{1.26}$$

A similar term can be derived for quarks, while the situation is slightly different due to the isospin doublet partner. Here, the flavor eigenstates (d', s' and b'), which take part in the weak interaction, are a mixture of mass eigenstates of the quarks (d, s and b). This is parametrized by the Cabibbo–Kobayashi–Maskawa (CKM) matrix [18–20]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.27}$$

The values of the individual CKM matrix elements can be obtained from different measurements. Along with unitary conditions, such as $|V_{\mathrm{ud}}|^2 + |V_{\mathrm{us}}|^2 + |V_{\mathrm{ub}}|^2 = 1$, the values can be determined precisely [3]:

$$\begin{pmatrix} |V_{\mathrm{ud}}| & |V_{\mathrm{us}}| & |V_{\mathrm{ub}}| \\ |V_{\mathrm{cd}}| & |V_{\mathrm{cs}}| & |V_{\mathrm{cb}}| \\ |V_{\mathrm{td}}| & |V_{\mathrm{ts}}| & |V_{\mathrm{tb}}| \end{pmatrix} = \begin{pmatrix} 0.97434\ ^{+0.00011}_{-0.00012} & 0.22506 \pm 0.0005 & 0.00357 \pm 0.00015 \\ 0.22492 \pm 0.00050 & 0.97351 \pm 0.0001 & 0.0411 \pm 0.0013 \\ 0.00875\ ^{+0.00032}_{-0.00033} & 0.0403 \pm 0.001 & 0.99915 \pm 0.00005 \end{pmatrix}. \tag{1.28}$$

The diagonal elements of the CKM matrix are close to unity and the off-diagonal elements close to zero, meaning that the coupling to two quarks of the same generation is more likely than to quarks of different generations.

1.2.4. Strong Interaction

The remaining force in the SM, the strong interaction, is described by the theory of quantum chromodynamics (QCD) [21–23]. Compared to the interactions described above, the gauge group of QCD is more complex, following an SU(3) symmetry. The corresponding gauge-invariant Lagrangian is defined by:

$$\mathcal{L}_{\mathrm{QCD}} = \overline{\psi}(\mathrm{i}\gamma^\mu D_\mu - m)\psi - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a, \tag{1.29}$$

with the gluon field strength tensor $G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu$ and the covariant derivative $D_\mu = \partial_\mu - \mathrm{i}g A^a_\mu \lambda_a$, where the indices a, b and c range from 1 to 8. Here, A^a_μ represents the gluon field, g is the coupling strength and f^{abc} is the SU(3) structure constant. The matrices λ_a are called Gell-Mann matrices and are the generators of the SU(3) group. The force-mediating boson of the strong interaction is the gluon g, which exists in eight different variants, depending on the charge.
To better understand the symmetry rules of QCD and the resulting properties of this interaction, an analogy from the macroscopic world is often used, the principle of color charge. In the theory of QCD there exist three different types of charge, called red, green and blue, while for antiparticles the charges are antired, antigreen and antiblue. Quarks are the only fermions that carry color charge, while gluons are the only bosons to do so. Like for any other interaction, the color charge in the strong interaction is conserved in any vertex. The analogy to color is used to represent that the combination of all three different (anti)colors or of a color and its corresponding anticolor results in a net color charge of zero, which is equal to white. The

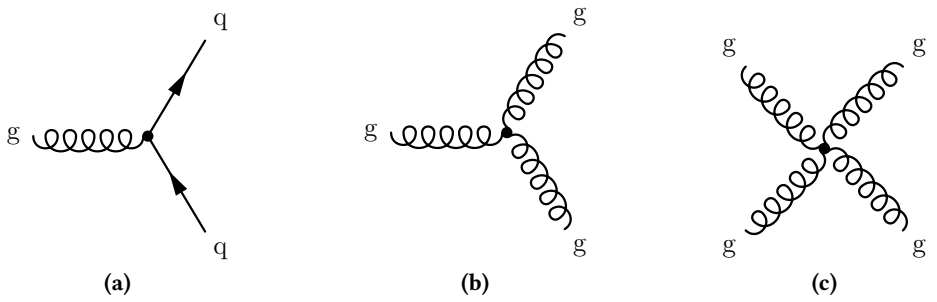


Figure 1.3.: The three basic interactions of the strong force. The gluons as color-charged boson can either interact with a quark-antiquark pair (a), with two other gluons (b) or four gluons can interact with a quadric vertex (c).

potential of the strong interaction can be parametrized as a combination of a Coulomb potential and linear term:

$$V(r) = -\frac{4\alpha_s}{3r} + kr, \tag{1.30}$$

with the coupling constant $\alpha_s = g^2/4\pi$ and a constant factor k of the linear term. Since the potential energy of QCD rises with longer distances, free quarks cannot be observed, as the energy at a certain distance between two color-charged objects will be released in the creation of a new quark-antiquark pair from a gluon of the field. This mechanism is called color-confinement. Thus, only bound states of three quarks or antiquarks, called baryons, or of a quark and antiquark, called mesons, can be observed at larger scales. In contrast to the confinement at longer distances, the force between color-charged objects at small distances becomes weaker, resulting in an effectively free movement, which is called asymptotic freedom. This effect is a result of the gluons carrying color charge themselves, allowing not only quark-gluon interactions, but also strong interactions among gluons. Interactions described by QCD are illustrated in Fig. 1.3. Similar to QED, virtual quark-antiquark pairs created from gluons effectively weaken the color charge of a single quark (screening). Because of the color charge of gluons the opposite effect occurs for virtual gluon loops (antiscreening). The strength of the combined effect depends on the distance to the object, or equivalent the energy scale. This behavior is absorbed in the coupling constant α_s, making it a scale-depended coupling:

$$\alpha_s(\mu_R^2) = \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2)\beta \ln \frac{\mu_R^2}{\mu^2}}. \tag{1.31}$$

Using this formula, the coupling strength of the strong interaction can be determined for any scale μ_R^2, called renormalization scale, from a measurement of a fixed scale μ^2. The value of β depends on the number of color charges n_c and the number of quark flavor n_f:

$$\beta = \frac{1}{12\pi}(11n_c - 2n_f). \tag{1.32}$$

As long as the energy scale is high enough, the coupling is smaller than one and QCD can be described with a perturbation series. This is possible until the so-called QCD scale Λ_{QCD}, at which the coupling is too large for the series to provide useful results.

1.3. From Partons to Particles

Elementary particles of interest for this thesis are only an intermediate state due to their limited lifetime between the initial protons in the collider and the decay products which are measured in the detector. The interactions described in the previous sections are the subject to study and occur multiple times in each collision, but usually only the hardest interaction, i. e., the interaction with the highest momentum transfer, is of interest. The initial particles that took part in the interaction are not the protons themselves, but the so-called partons of the proton, which each carry a fraction x of the total momentum of the proton. Because of QCD, the partons can either be the valence quarks of the proton (u and d), gluons or sea quarks, which are created via pair production through gluons. The probability to find a particular parton with momentum fraction x is described by the parton distribution function (PDF) $f(x, x_\mathrm{F})$, where the factorization scale x_F denotes the energy scale at which the PDF was evaluated. Measured at a given scale x_F, the PDF can be determined for other scales using the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equations [24–26].
A measure of the likelihood of the hard interaction, the cross section $\hat{\sigma}$, with the involved partons can then be calculated using the scattering matrix, which is given by the Dyson series of the interaction term. For proton-proton collisions, the total cross section to produce a particle X can be calculated as

$$\sigma(\mathrm{pp} \to \mathrm{X}) = \sum_{i,j} \iint \mathrm{d}x_1 \, \mathrm{d}x_2 f_i(x_1, x_\mathrm{F}) f_j(x_2, x_\mathrm{F}) \hat{\sigma}(\mathrm{ij} \to \mathrm{X}), \tag{1.33}$$

where $\hat{\sigma}$ is the cross section of the hard process to produce particle X from the partons i and j inside the proton and f_i, j are the PDFs of the two partons.
Given the high momentum of the final-state particles created in the hard scattering, additional gluons are radiated subsequently ,which carries away momentum. This process is called parton shower, as the emitted gluons can create additional partons themselves. This leads to a large increase of partons in the final state. In addition, color-charged objects cannot be separated for longer distances (see Eq. 1.30). Thus, all partons, i. e., the partons from the hard process, the shower, but also the remnants of the two protons need to form color-neutral bound states, a process which is called hadronization. This leads to various mesons and baryons, some of which decay further before they can reach the detector, creating even more additional particles. At this stage the particles are measured inside the detector. The directions of these hadrons are aligned and in the experiment they are not treated individually, but combined as bunches, called jets [27]. The whole chain is illustrated in Fig. 1.4.

1.4. The Top Quark

The top quark plays a special role in the SM as the heaviest elementary particle with a mass of $173.1 \pm 0.6\,\mathrm{GeV}$ [3]. Because of this circumstance, the observation of the top quark was not possible for a long time, although the existence of a sixth quark was already expected for symmetry reasons since the first observation of the bottom quark in 1977 [29] and even earlier to explain CP violation in Kaon decays [20]. The top quark was finally observed at the Tevatron

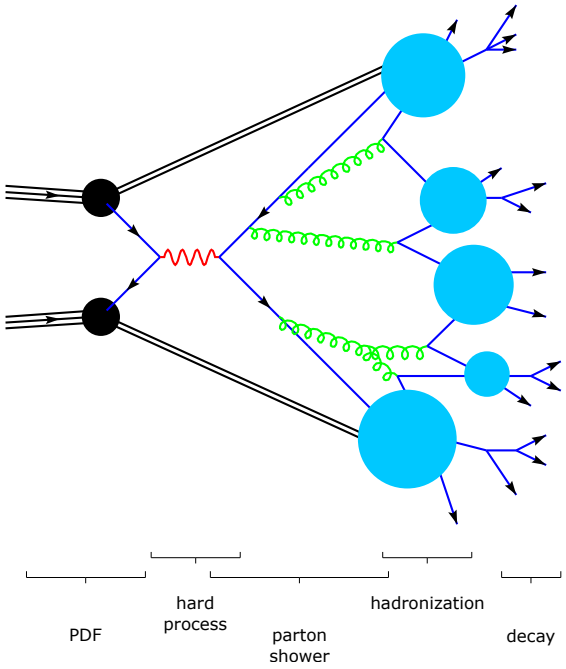


Figure 1.4.: Illustration of the transition from colliding protons to particles measured in a particle detector. Only one parton of each proton takes part in the hard scattering, carrying a large fraction of the proton's momentum. The partons in the final state of the hard scattering undergo a subsequent parton shower until the energy is low enough for color-neutral hadrons to form (hadronization), which may decay before they reach the detector. Adapted from [28].

collider (Chicago, USA) by the two experiments CDF and DØ in 1995 [30, 31]. Interestingly, the top-Yukawa coupling is

$$y_t = \frac{\sqrt{2}m_t}{v} \approx 1, \tag{1.34}$$

and it is unclear if this is just a coincidence or the result of an undiscovered theory.

1.4.1. Production

Top quarks can be produced via the strong and electroweak force, since quarks couple to all interactions described in the SM. Due to the high mass of the top quark, only two particle colliders in history were able to produce top quarks, namely the Tevatron, which was shut down in 2011, and the Large Hadron Collider (LHC), both hadron colliders. There are plans for future lepton colliders with sufficient energy for top quark production [32, 33].

Pair Production

The dominant production mode of top quarks is pair production via strong interaction. Exemplary leading-order Feynman diagrams are shown in Fig. 1.5. This involves the creation of a gluon through a quark-antiquark pair or two gluons in the initial state (Fig. 1.5(a) and 1.5(b)) or

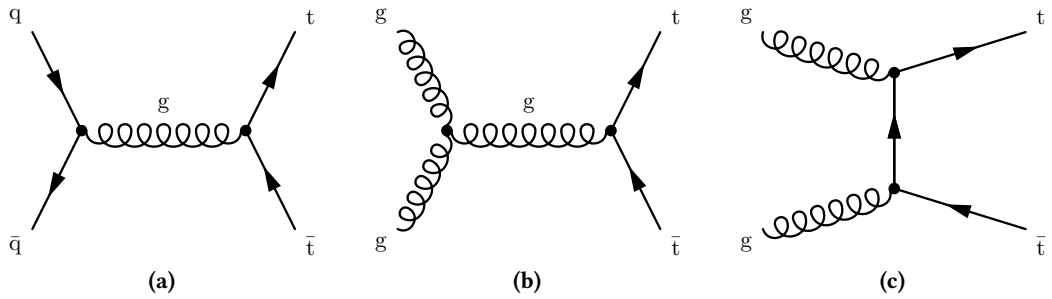


Figure 1.5.: The three main possibilities to create top quark-antiquark pairs at the LHC. Despite the annihilation of a quark-antiquark pair in the initial state into a gluon is still possible (a), it was relatively more likely at the proton-antiproton Tevatron than the LHC. Instead, the production with gluons in the initial state through time-like (b) and space-like diagrams (c) happens more frequently.

in the t channel (Fig. 1.5(c)). The production through quarks in the initial state was dominant at the proton-antiproton collider Tevatron with a center-of-mass-energy of $\sqrt{s} = 1.96\,\text{TeV}$, as anti-quarks were present as valence quarks there. In contrast, at the proton-proton collider LHC, the gluon-induced process is dominant. In general, the cross section of top quark pair production is higher at proton-antiproton colliders compared to proton-proton colliders. However at higher center-of-mass energies, the gluon-gluon initial state is also dominant for proton-antiproton colliders, thus making this difference smaller. The predicted cross section at the LHC for a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ is

$$\sigma_{t\bar{t}} = 831.76^{+19.77}_{-29.20}\ (\text{scale}) \pm 35.06\ (\text{PDF} + \alpha_s)\ \text{pb} \tag{1.35}$$

for a top quark mass of 172.5 GeV [34, 35]. An overview of the combined Tevatron and various LHC measurements of the top quark pair production cross section compared to the theoretical prediction is presented in Fig. 1.6.

Single Top Production

The production of single top quarks is more rare compared to the pair production, as it is induced by the weak interaction through a Wtb vertex. Single top quarks can be produced through the t channel, in association with a W boson and in the s channel. The names t and s channel refer to the Mandelstam variables t and s which describe the momentum transfer and the center-of-mass energy, respectively. The basic Feynman diagrams in leading order for all three processes are shown in Fig. 1.7. At the LHC, the t channel is the dominant production mode for single top quarks 1.7(a) with a cross section of

$$\sigma_{t\ \text{ch.}} = 216.99\ ^{+6.62}_{-4.64}\ (\text{scale}) \pm 6.16\ (\text{PDF} + \alpha_s)\ \text{pb} \tag{1.36}$$

at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ [37–39]. The top quark is produced through the exchange of a W boson between a bottom quark and a light-flavored quark in the initial state, changing both quark flavors. This light-flavored quark is usually emitted in a more forward

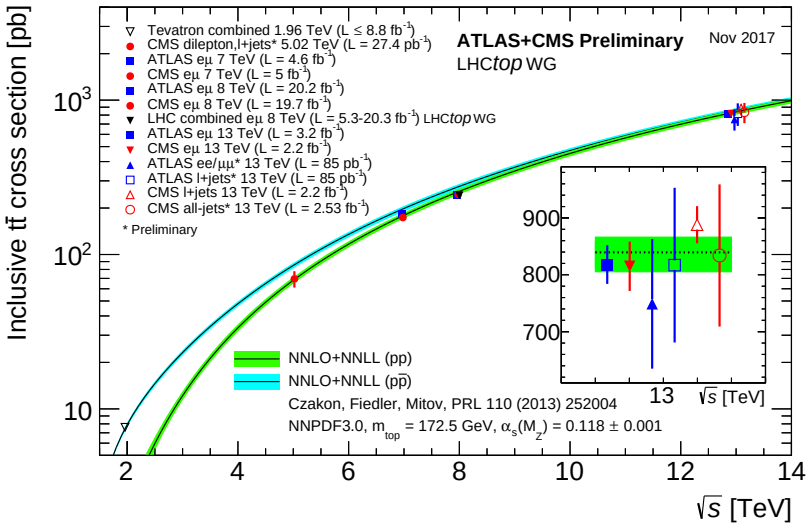


Figure 1.6.: The experimental status of top quark-antiquark pair production cross section measurements at hadron-hadron colliders. Various measurements have been carried out by different collaborations for different center-of-mass energies. All measurements are in excellent agreement with the predictions from theory [36].

direction compared to the other quarks. The cross section prediction is obtained at next-to-leading order (NLO) accuracy in QCD, i. e., including all additional Feynman diagrams that are suppressed by a factor α_s. The scale uncertainty of the prediction is estimated by independently varying the nominal factorization and renormalization scales, which are set to the top quark mass, up and down by a factor of two, where combinations in which both scales differ more than a factor of two are not taken into account. The final scale uncertainty is then quoted as the envelope of all different cross section results with varied scales. For the estimation of the PDF and α_s uncertainties, the envelope of different PDF sets have been used, following the PDF4LHC prescription [40, 41]. First results at next-to-next-to-leading order (NNLO) have also been published recently [42, 43], but until the uncertainties on these NNLO predictions become

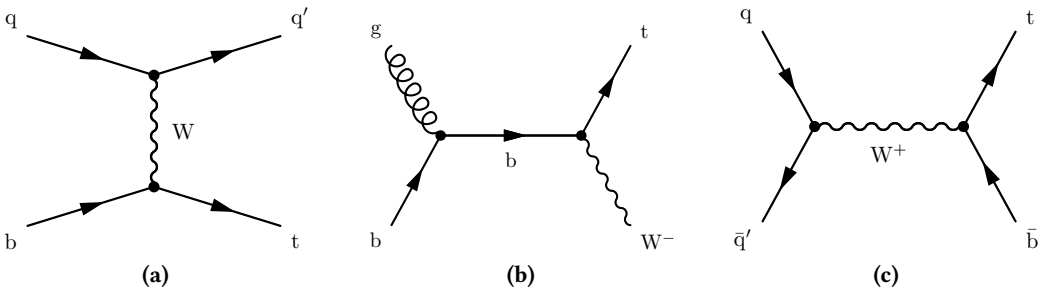


Figure 1.7.: The three main possibilities to produce a single top quark at the LHC through a Wtb vertex. The most common one is the t channel (a), followed by the tW-associated production (b) and the s channel (c).

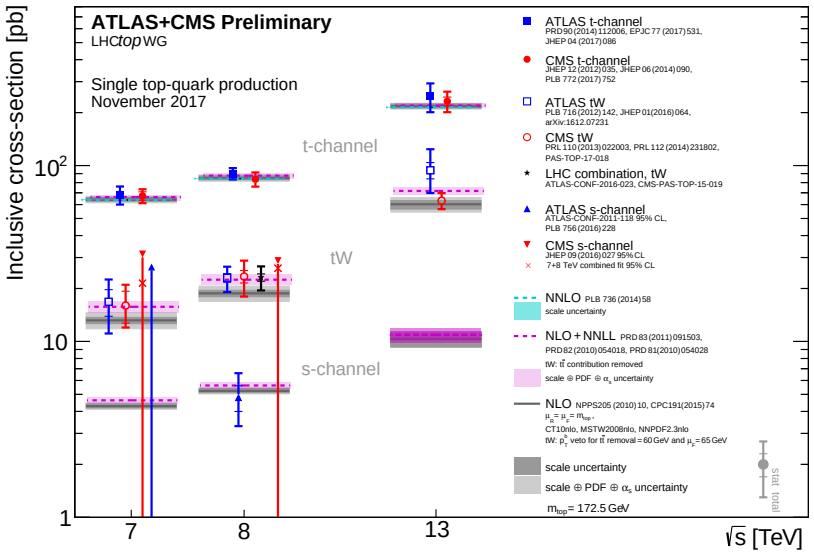


Figure 1.8.: Overview of the experimental status of the single top quark production cross section in the three different channels for a center-of-mass-energy of 7, 8 and 13 TeV [36]. All measured cross sections are in agreement with the predictions from theory. The analysis of the t-channel cross section by the CMS Collaboration will be extensively discussed in Section 6.

available for more different parameter values, the NLO result of Eq. 1.36 is used as prediction by measurements of the ATLAS and CMS Collaborations.
Another way to produce a single top quark is in association with a W boson through a space-like bottom quark 1.7(b). The cross section at $\sqrt{s} = 13\,\text{TeV}$ is [37, 39]:

$$\sigma_{\text{tW}} = 71.7 \pm 1.80\ (\text{scale}) \pm 3.40\ (\text{PDF} + \alpha_s)\ \text{pb}. \tag{1.37}$$

The last and rarest channel is the production through the s-channel process with a virtual W boson 1.7(c), where the cross section at the LHC is only

$$\sigma_{s\ \text{ch.}} = 10.32\ ^{+0.29}_{-0.24}\ (\text{scale}) \pm 0.27\ (\text{PDF} + \alpha_s)\ \text{pb} \tag{1.38}$$

for a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ [37, 39].
Single top production has been observed at the LHC for different center-of-mass energies in the t channel and the tW-associated production. The s-channel production has been only observed at the Tevatron, while at the LHC there is only evidence so far. Measurements of the three channels at the LHC are presented in Fig. 1.8.

1.4.2. Decay

The top quark is the only quark in the SM that does not form bound states because its lifetime of about 10^{-25} s is shorter than the typical hadronization time scale [44]. This unique property allows the study of, e. g., spin properties of top quarks, which would otherwise be washed-out by the process of hadronization. As the top quark mass exceeds the W boson mass, it can decay

through a real W boson and change its flavor into a down-type quark. This can either be a down, strange or bottom quark, but since the CKM martix element V_{tb} is almost equal to 1, the top quark decays nearly exclusively into a bottom quark. Therefore, the decay of the top quark is characterized by the subsequent decay of the W boson, which happens either into a quark and an antiquark (branching ratio 67.4%), called hadronically decaying top quark, or into a charged lepton and the corresponding neutrino (32.6%), called leptonically decaying top quark. In case of top quark pair production, the final state can either be fullhadronic, dileptonic or semileptonic, in case one top quark decays hadronically and the other top quark leptonically.

1.5. The Higgs Boson

With the existence of a Higgs field it should also be possible to observe an excitation of this field, a Higgs boson. The mass of this boson is a free parameter of the SM, therefore it is experimentally challenging to directly search for it. First searches from phenomenology started in the early 1970s [45], but left out a wide range of possible masses. More direct constraints on the mass came later from the Large Electron-Positron and Tevatron colliders [46–48]. Finally, on the 4th of July 2012, both the ALTAS and CMS Collaborations at the LHC announced the observation of a new boson with a mass of around 125 GeV [1, 2]. Although it was not clear by the time if this was the long-searched-for Higgs boson, succeeding measurements of properties and the increased amount of data leave almost no other explanation [49].

1.5.1. Production

An SM Higgs boson can be produced in different ways at the LHC. These include production modes involving fermions and bosons. The different main production cross sections depending on the center-of-mass energy are shown in Fig. 1.9 and each production mode is explained in the following. The gluon-gluon fusion (ggF) process is the most dominant production mode at the LHC. It is also the only mode in which the Higgs boson is produced in isolation and not with any byproduct. Since the Higgs boson couples directly to the mass, it cannot be produced via gluons directly. Instead, an additional quark loop is needed. The corresponding Feynman diagram is shown in Fig. 1.10(a). The top quark has the largest contribution SM inside the loop since its mass is approximately forty times higher than the mass of the second heaviest quark, the bottom quark. Although this production channel has the highest cross section, the experimental signature is challenging for some Higgs boson decays due to several larger backgrounds with similar final-state particles.
In the vector boson fusion production mode (VBF) two vector bosons are radiated from the two initial state quarks. Both fuse to form together the Higgs boson (Fig. 1.10(b)). These can be either two neutral bosons or two W bosons of opposite charge. The cross section of this production mode is already an order of magnitude lower than the cross section of the ggF process. The VBF channel has the feature of two additional jets in the final state in the forward direction which makes it experimentally more usable.
The associated vector boson production (VH), or Higgsstrahlung, is the third largest Higgs boson production mode. Two initial-state quarks form a virtual W or Z boson which radiates

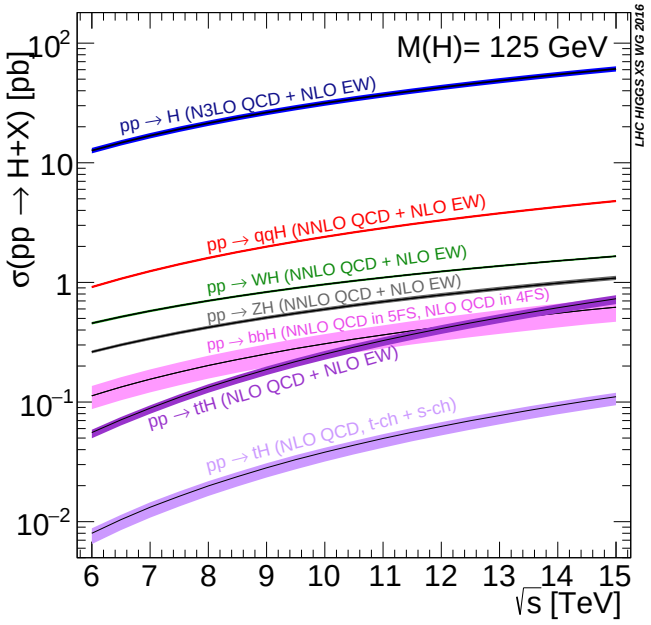


Figure 1.9.: Theoretical cross sections for different Higgs boson production modes at proton-proton colliders depending on the center-of-mass energy. A Higgs boson mass of 125 GeV is assumed. The predicted cross sections span over several orders of magnitude from the production of a Higgs boson via gluon-gluon fusion (pp → H) to the production of a Higgs boson in association with a single top quark (pp → tH) [50].

off a Higgs boson (Fig. 1.10(c)).
The last of the main production modes is the associated production with fermion pairs, mainly with top (ttH) and bottom quark pairs (bbH) (Fig. 1.10(d)). Although the coupling of the Higgs boson is much larger for top quarks than for bottom quarks, the cross sections of both modes are about the same at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$.
An additional production mode involving associated production with single top quarks (tH) will be discussed extensively in Section 1.6.

1.5.2. Decay

Since the Higgs boson couples directly to every massive particle and indirectly via loops to massless particles it can decay in many different ways [50]. Figure 1.11 shows the different possible branching ratios for a SM Higgs boson.
In 58.1% of all cases, the Higgs boson decays into a bottom quark pair. While this channel is the most dominant one, it is also experimentally the most challenging one due to the signature being similar to QCD processes. Other fermionic channels include the decay into tau leptons (6.3%), charm quarks (2.9%) and muons (<0.1%). Although the decay into a top quark pair would be favored, it is kinematically not allowed for the mass of a SM Higgs boson.
In the bosonic channel the Higgs boson decays mostly into a pair of gauge bosons. Most dominant here are W bosons (21.5%), followed by gluons (8.2%) and Z bosons (2.6%). The branching

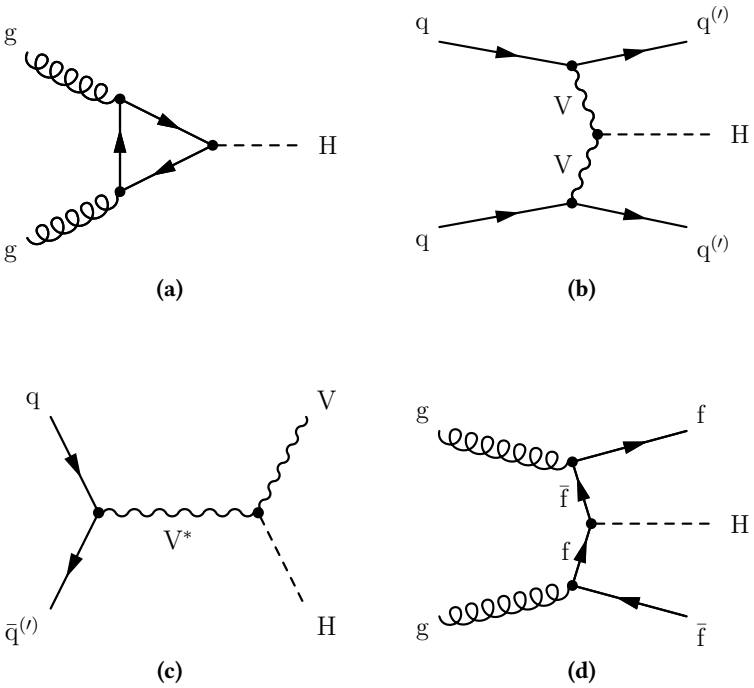


Figure 1.10.: The four main production modes for a Higgs boson at the LHC. The dominant channel is gluon-gluon fusion (a), but it is also the experimentally most challenging one due to the overwhelming background. Rarer, but also more promising, are the vector boson fusion (b) and the associated production with a vector boson (c). The rarest of the four main production modes is the associated fermion pair production (d).

ratio into a pair of photons is only 0.2%, but this mode has a clean signature.
Experimentally observed are the bosonic decay modes of the Higgs boson into photons, W bosons and Z bosons [50]. The only observed fermionic decay modes are the decay into tau leptons [51] and the decay into bottom quarks [52, 53].

1.5.3. Couplings

To answer the question whether the observed Higgs boson is really the Higgs boson predicted by the SM, the coupling of the Higgs boson to other particles can be studied. Since many different theories predict additional Higgs bosons, each with different couplings, a study of these coupling can be used to falsify or exclude such theories. For this reason the so-called kappa framework was developed [55]. One assumptions made in this framework is that the Higgs boson is a single and narrow resonance. In that case the cross section σ times branching ratio $\mathcal{B}$ of a process involving the Higgs boson H can be written as

$$(\sigma \cdot \mathcal{B})_{i \to H \to f} = \frac{\sigma_{i \to H} \cdot \Gamma_{H \to f}}{\Gamma_H}, \tag{1.39}$$

with the initial state i, the finale state f and the (partial) width Γ. Any deviations from an SM-like behavior can be parameterized with a coupling modifier κ for each given coupling. For example,

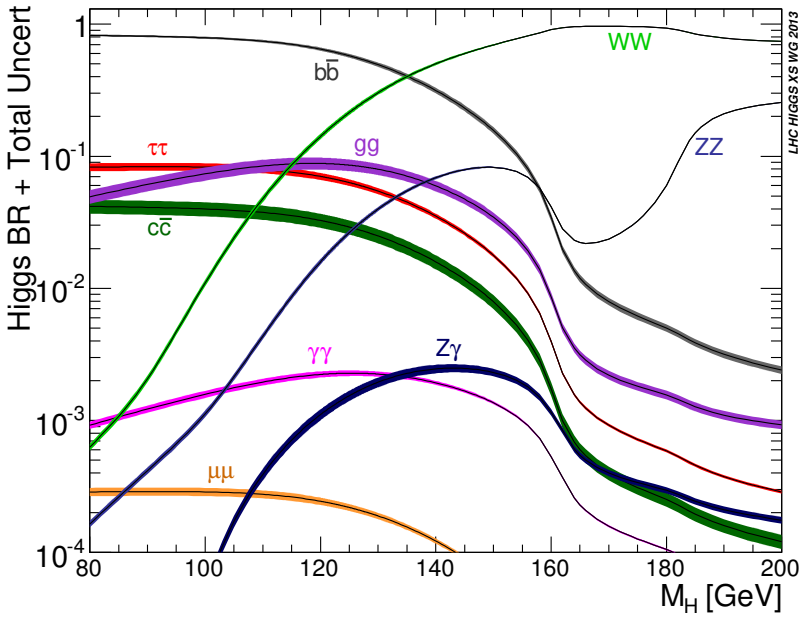


Figure 1.11.: The predicted branching ratios of the Higgs boson depending on the Higgs boson mass. Since the Higgs boson couples directly to mass, the branching ratio into heavier particles is higher compared to lighter particles. Although no direct decay to massless particles like gluons or photons is allowed, they can couple to the Higgs boson through virtual loops of heavier particles like the top quark or the W boson [54].

the Higgs boson production through gluon-gluon fusion and the subsequent decay into a pair of bottom quarks is parameterized by

$$\sigma_{\mathrm{ggF}} \cdot \mathcal{B}(\mathrm{H} \rightarrow \mathrm{bb}) = \kappa_{\mathrm{g}}^2 \cdot \sigma_{\mathrm{ggF}}^{\mathrm{SM}} \cdot \kappa_{\mathrm{b}}^2 \cdot \mathcal{B}_{\mathrm{H} \rightarrow \mathrm{bb}}^{\mathrm{SM}}, \tag{1.40}$$

where the scaling factors are either $\kappa_i^2 = \sigma_i/\sigma_i^{\mathrm{SM}}$ or $\kappa_i^2 = \Gamma_i/\Gamma_i^{\mathrm{SM}}$ depending on whether they affect the production or decay of the Higgs boson. By construction κ_i is 1 for all massive particles in the SM. In general the modifications can be more complex as for instance in the given example above there exists no direct coupling of the Higgs boson to gluons. In the SM, the effective coupling arises from loops of fermions and therefore, the kappa factor of gluons is decomposed into $\kappa_{\mathrm{g}}^2 = 1.06 \cdot \kappa_{\mathrm{t}}^2 + 0.01 \cdot \kappa_{\mathrm{b}}^2 - 0.07 \cdot \kappa_{\mathrm{t}}\kappa_{\mathrm{b}}$ [49]. The kappa framework works in leading order but can in principle be extended to next-to-leading order since most higher-order corrections factorize with the rescaling of couplings. Additional assumptions can be made such as that couplings to fermions κ_{f} and vector bosons κ_{V} do not depend on the specific type of particle. The current status of exclusion limits on those two modifiers is shown in Fig. 1.12.

1.6. Associated Production of a Higgs Boson with Top Quarks

The most interesting way to investigate the interplay between the Higgs boson and the top quark is the study of associated production modes. Similarly to the sole top quark production,

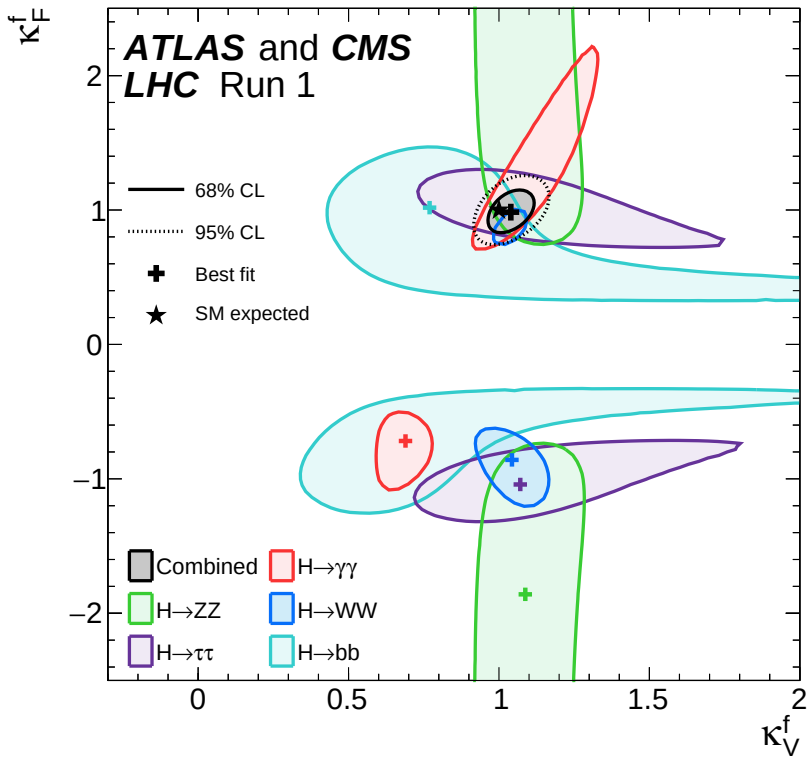


Figure 1.12.: The experimental status of constraints on the coupling modifiers of the Higgs boson to fermions κ_F and vector bosons κ_V after the combination of the Higggs boson measurements from the ATLAS and CMS Collaborations at Run I of the LHC [49].

it is possible to produce a Higgs boson in association with top quark pairs ($t\bar{t}H$) and single top quarks (tH). This allows direct access to the measurement of the top-Yukawa coupling as those processes all involve a direct coupling of the Higgs boson to the top quark, while the signature of the final state is still manageable for an experiment, unlike the gluon-gluon fusion production mode.
The most common associated production mode is the process involving top quark pairs (see the exemplary Feynman diagram in Fig. 1.10(d)) with a cross section of

$$\sigma_{t\bar{t}H} = 506.5\ ^{+29.3}_{-46.6}\ (\text{scale})\ ^{+20.8}_{-18.2}\ (\text{PDF} + \alpha_s)\ \text{fb} \tag{1.41}$$

at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ with a Higgs boson mass of $m_H = 125.09\,\text{GeV}$ [50]. This particular production mode has been only recently observed for the first time by the ATLAS and CMS Collaborations and is the first direct evidence for the coupling of the Higgs boson to top quarks, or more generally, to up-type fermions [56, 57]. The cross section of this process is proportional to κ_t^2 and therefore allows for a direct measurement of the magnitude $|\kappa_t|$, but not the relative sign.
This degeneracy can be resolved through associated Higgs boson production with single top

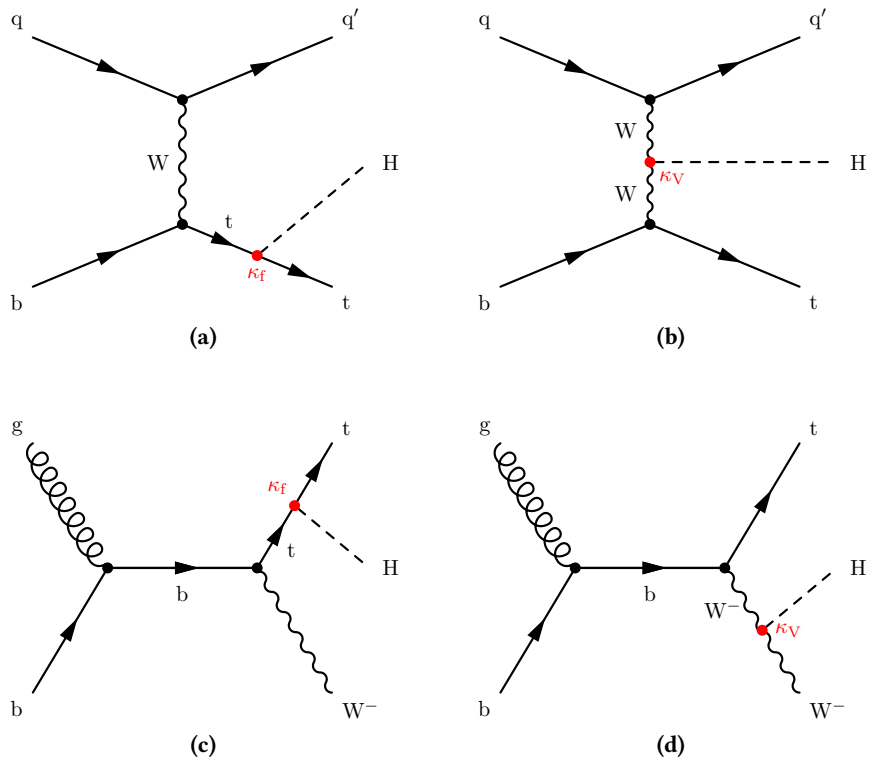


Figure 1.13.: Relevant Feynman diagrams contributing to the associated production of single top quarks and a Higgs boson (tH), in particular tHq ((a) and (b)) and tHW ((c) and (d)). The Higgs boson in both processes can be either radiated off from the top quark ((a) and (c)), depending on the coupling modifier κ_f, or from the W boson ((b) and (d)) with the coupling modifier κ_V. The overall amplitude of each process strongly depends on the difference of both coupling modifiers.

quarks [58–62]. Similar to the pure single top production, there is a distinction between the t channel, the tW-associated production and the s channel, which are denoted by their final state particles: tHq, tHW and tHb, respectively. The cross sections for all three channels at $\sqrt{s} = 13\,\text{TeV}$ are [50]:

$$\sigma_{\text{tHq}} = 74.26\,^{+4.83}_{-10.92}\,(\text{scale})\,^{+2.60}_{-2.75}\,(\text{PDF} + \alpha_s)\,\text{fb}, \tag{1.42}$$

$$\sigma_{\text{tHW}} = 15.17\,^{+0.74}_{-1.02}\,(\text{scale})\,\pm 0.96\,(\text{PDF} + \alpha_s)\,\text{fb}, \tag{1.43}$$

$$\sigma_{\text{tHb}} = 2.875\,^{+0.069}_{-0.052}\,(\text{scale})\,\pm 0.063\,(\text{PDF} + \alpha_s)\,\text{fb}. \tag{1.44}$$

As the cross section for the tHb process is quite low compared to tHq and tHW production, this process is neglected for the scope of this thesis and the following discussion.
In contrast to $t\bar{t}H$ production, the Higgs boson can either couple to the top quark or the W boson, as single top production always involves a Wtb vertex. The relevant Feynman diagrams for tHq and tHW are shown in Fig. 1.13. The cross sections of both production modes can be

parameterized with the scaling factors κ_t and κ_W under the assumption $\kappa_f = \kappa_t$ and $\kappa_V = \kappa_W$ [63]:

$$\sigma_{tHq} = (2.63 \cdot \kappa_t^2 + 3.58 \cdot \kappa_W^2 - 5.21 \cdot \kappa_t \kappa_W) \cdot \sigma_{tHq}^{SM}, \tag{1.45}$$

$$\sigma_{tHW} = (2.91 \cdot \kappa_t^2 + 2.40 \cdot \kappa_W^2 - 4.22 \cdot \kappa_t \kappa_W) \cdot \sigma_{tHW}^{SM}. \tag{1.46}$$

The cross section not only depends on the square of the coupling modifiers, but also on an interference term, which is introduced due to the two different diagrams contributing to the given process.
As the magnitude $|\kappa_t|$ can already be constrained by other processes, the most interesting scenario to study is the SM case with a flipped sign, i. e., $\kappa_t = -1$, which is referred to as the *inverted top coupling* (ITC) scenario in the following. A first direct search for the ITC scenario with the tHq process was performed by the CMS Collaboration at $\sqrt{s} = 8\,\text{TeV}$, but was not able to confirm or exclude the scenario yet [64]. Although this scenario is disfavored through measurements of the $\gamma\gamma$ final state with top quarks contributing in the necessary loop to couple photons and the Higgs boson, other, yet undiscovered, particles may also contribute to this loop.
In a more general way, the associated production of a Higgs boson with single top quarks can be used to search for a CP-violating coupling of the Higgs boson [62]. The effective Lagrangian of this non-SM Higgs boson can be written as

$$\mathcal{L} = -\overline{\psi} \left(\cos(\alpha) \kappa_{Htt} g_{Htt} + i \sin(\alpha) \kappa_{Att} g_{Att} \gamma_5 \right) \psi X_0, \tag{1.47}$$

with the corresponding field X_0, the CP-mixing phase α, the coupling strength $g_{itt} = \frac{m_i}{v}$ and the coupling modifier κ. The indices Htt and Att denote the scalar and pseudoscalar component, respectively. This parameterization has the benefit of easily transferring between a pure CP-even state ($\alpha = 0°$) and a pure CP-odd state ($\alpha = 90°$), while also recovering the SM case for $\alpha = 0°$ and $\kappa_{Htt} = 1$. Furthermore, by setting $\kappa_{Htt} = 1$ and $\kappa_{Att} = {}^2/_3$, the SM Higgs boson gluon-gluon fusion cross section can be reproduced for any value of the CP-mixing angle α, therefore a possible CP-mixing would only be visible in associated production modes. The predicted cross sections for the associated production with single top quarks (tX_0) and top quark pairs ($t\bar{t}X_0$) is shown in Fig. 1.14. Similar to the case of modified couplings κ, the cross section of the associated production with top quark pairs is degenerate, while the production mode with single top quarks is sensitive to the whole range of the CP-mixing angle α. In addition, the SM case with $\alpha = 0°$ is the least sensitive scenario, making any significant excess above the SM prediction a possible sign of physics beyond the SM.
The CP-nature of the Higgs boson can also be studied with angular distributions of the Higgs boson decay to two Z bosons, where both Z bosons decay leptonically. While a recent measurement of the CMS Collaboration is in agreement with the SM prediction [65], the latest measurement of the ATLAS Collaboration observes small deviations from the SM [66].

1.7. Indications for Physics beyond the Standard Model

Despite the success of the SM in predicting and describing subatomic particles and their interactions, some phenomena are not part of this theory. Furthermore, there are indications that the SM of the present day is not a complete theory of fundamental interactions.

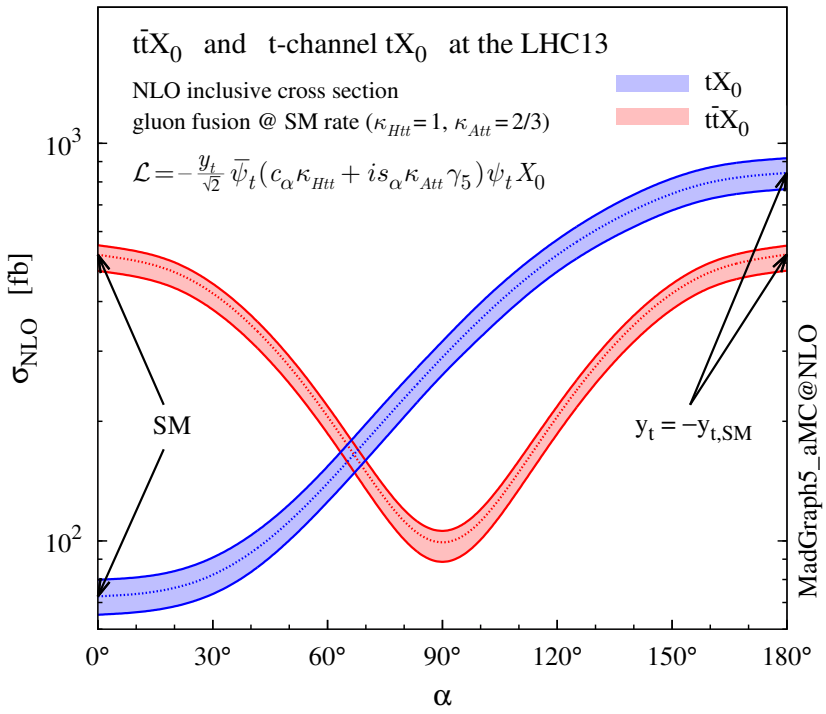


Figure 1.14.: The predicted cross section for the associated production of a CP-violating Higgs boson with top quark-antiquark pairs and single top quarks as a function of the CP-mixing angle α. While the associated production with top quark-antiquark pairs is not able to distinguish a CP-mixing angle of 0° and 180°, this degeneracy is lifted through the associated production with single top quarks [62].

The most striking issue is that the SM completely neglects an interaction that is dominant in everyday life: gravity. There is no term in the Lagrangian of the SM accounting for this type of interaction. Gravity is well described on a macroscopic scale by the theory of general relativity by Albert Einstein [67]. If one tries to incorporate gravity into the SM in a similar way like QED or QCD, non-renormalizable divergences occur. Such a quantum field theory of gravity would introduce an additional massless boson with spin 2, the graviton. However, completely neglecting gravitational interactions is not a problem on a microscopic scale, as the other forces are stronger by several orders of magnitude.
Measurements of rotation velocities from spiral galaxies suggest that a large fraction of matter inside the galaxy is actually not visible [68,69]. This concept of *dark matter* is further supported by observations of microlensing around galaxies [70,71] and the spectrum of the cosmic microwave background [72,73]. There is no suitable candidate in the SM which can explain this behavior in a satisfying way. Furthermore, from the evidence of the accelerating expansion of the Universe the concept of *dark energy* arises. Latest measurements came to the conclusion that dark energy accounts for a total of 69% of the energy in the Universe, while dark matter contributes with 26%. Only 5% of the total energy of the Universe is described by the known baryonic matter [73].

Observation of neutrino-flavor oscillations [74, 75] have proven that neutrinos possess a non-zero mass, which is in contrast to the predictions of the SM. Since the oscillation frequency is proportional to the square of the neutrino mass difference, no oscillation is possible if all neutrino flavors would have zero (or the same) mass. Some experiments even report significant excess above the expected neutrino-flavor oscillation, which can be explained by the existence of sterile neutrinos [76–83].
The SM also has no explanation for the baryon asymmetry in the Universe. The Big Bang created equal amounts of matter and antimatter, but an imbalance was created afterwards which resulted in a matter-dominated Universe. This mechanism is called baryogenesis. Baryon asymmetry is predicted in the SM via CP-violation in the CKM-matrix. However, the magnitude of CP-violation is not sufficient to describe the observable imbalance of matter and antimatter in the Universe [84, 85], which makes a possible CP-violating Higgs boson, as outlined in the previous section, even more interesting.

2. Statistical Analysis

In order to be able to make accurate predictions and comparisons between theory and experiment, a solid foundation of statistical inference is necessary. In this chapter, the multivariate analysis techniques are explained, which are crucial for the analyses of the investigated processes in this thesis. Afterwards, the statistical methods employed in this thesis are introduced.

2.1. Multivariate Analysis

In high energy physics analyses, it is crucial to have a good separation of the signal process from the various background processes. For a few processes this can be achieved by simple conditions on certain variables, e. g., the number of jets in an event or the transverse momentum of a particle. Although this may work well for certain processes, a more sophisticated approach to separate signal and background processes is often necessary. Nevertheless, almost all analyses utilize this kind of selections to enhance their signal-to-background ratio. The methods employed in this thesis are based on the concept of multivariate analysis (MVA), meaning the simultaneous evaluation of multiple variables to classify events into categories. Through this way, correlations of variables can be exploited to gain even more separation power compared to the simple ensemble of variables. The specific methods used in this thesis are explained in the following. Each method requires first to be trained on a data set where the desired outcome is known to define the algorithm. This is done with simulated data sets. Afterwards, the method can be applied to a measured data set. These methods need to be trained and applied to a phase-space region in which the data is well described by simulation. Otherwise, a mismodeling can result in unexpected outcome. To implement these methods in the analyses, either the TMVA (Toolkit for Multivariate Data Analysis) package [86] or the self-developed MiST (Machine learning in Single Top) framework is used. The latter combines state-of-the-art machine learning tools such as Keras [87] and TensorFlow [88] with a relatively easy data handling for high energy physics analyses.

2.1.1. Boosted Decision Trees

A boosted decision tree (BDT) is a combination of many binary classifiers, in this case simple decision trees, each of its own with limited classification abilities. The weighted combination of those decision trees can result in a significant increase in the classification power.
A simple decision tree consists of a starting point, called root node, from which the incoming data is classified based on a single property of the data. The two branches themselves lead to new nodes on which the data is classified based on a different property. The procedure is repeated until a certain criterion is fulfilled and the data is finally classified as signal- or background-like depending on the end node, called leaf. The principal structure of a decision tree is shown

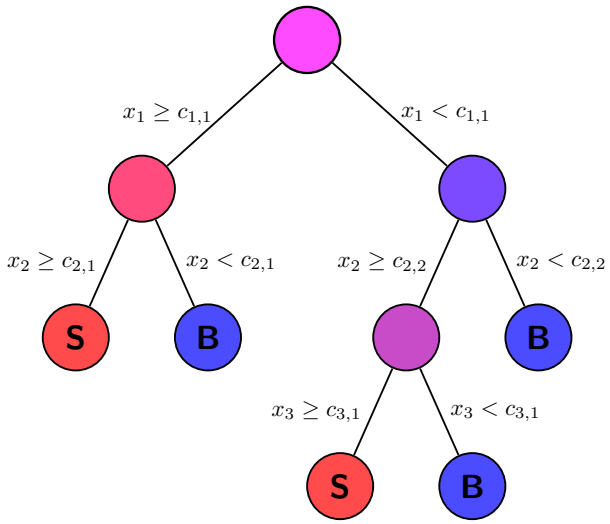


Figure 2.1.: The basic principle of a decision tree. At each level, a separation on the value of a single variable is applied, splitting the data into two categories. Consecutive cuts on different variables are used until the maximum separation is achieved.

in Fig. 2.1. The cut value at a given split node is determined by maximizing the separation S between signal and background

$$S = G_{\text{mother node}} - G_{\text{daughter node 1}} - G_{\text{daughter node 2}} \tag{2.1}$$

with the Gini coefficient G [89]. The Gini coefficient describes the imbalance of a data set and is defined as

$$G = P(1-P)\sum_{i=1}^{N} w_i, \tag{2.2}$$

where P is the purity of the data set and w_i the weight associated to entry i of a data set with N entries. Given N_s number of signal and N_b number of background events, the purity is expressed as

$$P = \frac{\sum_i^{N_\text{s}} w_\text{s}}{\sum_i^{N_\text{s}} w_\text{s} + \sum_i^{N_\text{b}} w_\text{b}}, \tag{2.3}$$

where w_s are the weights of signal and w_b are the weights of background events. The robustness of the classification is increased by employing many different decision trees into a so-called forest and evaluating the average.
In addition to using multiple decision trees, the performance can be enhanced by adapting the training procedure such that misclassified events receive a higher weight in the training iteration of the next tree. This method is called boosting. In general, many different algorithms

for boosting decision trees exist. The method used in this thesis is the adaptive boosting (AdaBoost) [90] algorithm. The rate of misclassified events in the m-th tree is defined as

$$r_m^{\text{mis}} = \frac{\sum_i w_i \delta_i^{\text{mis}}}{\sum_i w_i} \tag{2.4}$$

with the weight w_i ($= 1/N$ in the first iteration) and the parameter δ_i^{mis}, which is 1 for misclassified events and 0 for a correct classification. The boost weight of a single tree m is then defined as

$$\alpha_m = \ln \frac{1 - r_m^{\text{mis}}}{r_m^{\text{mis}}}, \tag{2.5}$$

from which the weight of the subsequent tree is calculated:

$$w_{m+1} = w_m \cdot \mathrm{e}^{\alpha_m \delta_i^{\text{mis}}}. \tag{2.6}$$

After the evaluation of all individual trees, the sum of the weights is normalized to unity. The final BDT score is then constructed as the weighted sum of the individual trees:

$$\text{BDT} = \sum_{m=1}^{N_{\text{tree}}} \alpha_m \delta_m, \tag{2.7}$$

where δ_m is the binary classification of the m-th tree for a given event.

2.1.2. Artificial Neural Networks

The concept of artificial neural networks (ANN) is based on the interconnection of neurons, similarly to the brains of animals and even humans. Different realizations of ANN exist, but one of the most common is the multilayer perceptron, which is also used in this thesis. The principal structure is shown in Fig. 2.2.
A neural network is organized in different layers, each with a specific number of neurons. Each neuron in a layer is connected to all neurons of the previous and following layer. The first layer is called the input layer and consists of one neuron for every input of the network. Each connection between two neurons is associated with a weight. Therefore, the neurons in the second layer, called hidden layer, receive the input

$$y_j = \sum_{i=1}^{N_{\text{input}}} w_{ij} x_i, \tag{2.8}$$

with the value x_i at input neuron i, the weight w_{ij} at the connection and the result y_j at hidden neuron j. Before the result is processed further in the network, the value of y_j is mapped with an activation function from an arbitrary range to a well-defined interval. In general, many different functions can be used, but the functions need to be non-linear to distinguish the network from a simple matrix multiplication and monotonous to allow a simple minimization of the error function. One of the most common is the hyperbolic tangent function

$$\tanh(x) = 1 - \frac{2}{\mathrm{e}^{2x} + 1}, \tag{2.9}$$

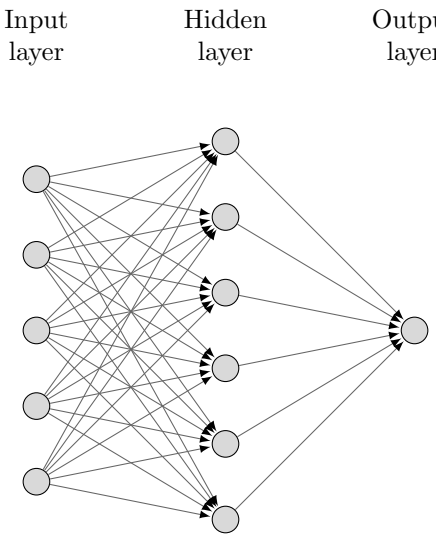


Figure 2.2.: The basic setup of an artificial neural network. Each input variable gives rise to a neuron in the input layer, which are all connected to every neuron in the hidden layer. In the same way every neuron in the hidden layer is connected to the output neuron.

which transforms the input x to an interval of -1 to 1. All neurons of the hidden layer are then mapped in the same way to the single neuron of the last layer, the so-called output layer, to build the final neural network variable:

$$y_{\text{ANN}} = \tanh\left(\sum_{j=1}^{N_{\text{hidden}}} w_j \cdot \tanh\left(\sum_{i=1}^{N_{\text{input}}} w_{ij} x_i\right)\right). \tag{2.10}$$

The weights need to be determined in the training step of the network to achieve the desired outcome. For each training iteration the network result is compared to the target, which is either $+1$ or -1 for signal and background events, respectively. The set of weights $\vec{w}$ is then adjusted in such a way that an error function

$$E(\vec{x}_i|\vec{w}) = \sum_{i=1}^{N} \ln\left(y_i \delta_i + \epsilon\right) \tag{2.11}$$

for a given number of events N is minimized. Here, $\vec{x}_i$ are the input values for event i, y_i is the achieved response of the network, δ_i the desired output and ϵ a regularization factor. This method of adjusting the weights of the network is called back propagation [91]. Finding the minimum of the error function is an optimization problem, which can be solved by different approaches. For the ANN in this thesis, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [92–95] is used to find the best possible weights. The algorithm is a quasi-Newton method in which the Hessian matrix does not need to be computed directly, which saves a significant amount of time in a non-linear optimization problem.

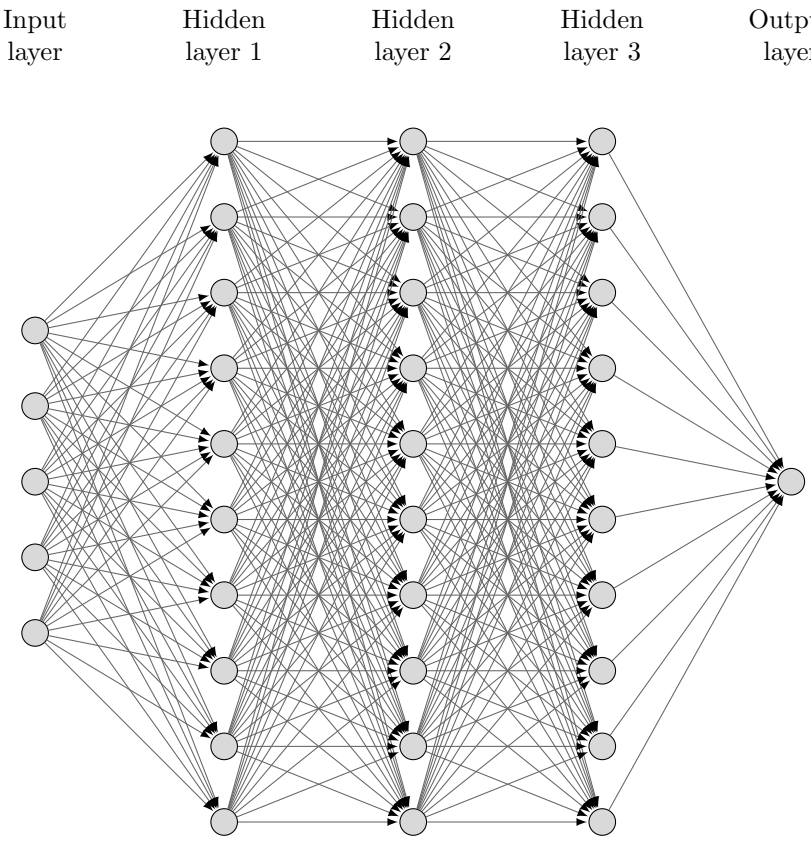


Figure 2.3.: Exemplary structure of a deep neural network. The structure shown here is similar to that of a standard neural network, but consists of significantly more neurons and layers. In general, deep neural networks can be even more complex, including connections between non-neighboring nodes and constitutional layers, but this illustration is chosen to resemble the network used in this thesis.

2.1.3. Deep Neural Networks

The determination of the minimum of the error function is the limiting factor for the complexity of a neural network, as the time to compute the error function quickly rises with additional layers and neurons. Thus, the classical ANN discussed in the previous section is limited to three layers and the separation power is similar to that of a BDT, which is in general much faster. The evaluation of the error function is a rather simple operation for modern central processing units (CPUs). Due to the development of graphics processing units (GPUs) in recent years, the training time of neural networks was reduced by several orders of magnitude as GPUs are more suited for fast and parallel floating point operations. This allowed neural networks to increase their complexity in the numbers of layers and neurons. To distinguish those complex networks from the shallow ANN with only one hidden layer, they are referred to as deep neural networks (DNNs). An exemplary structure of a DNN is shown in Fig. 2.3. The additional neurons and layers of a DNN compared to an ANN make it possible to derive more subtle features from the data set, which can significantly increase the separation power. In addition, more than one

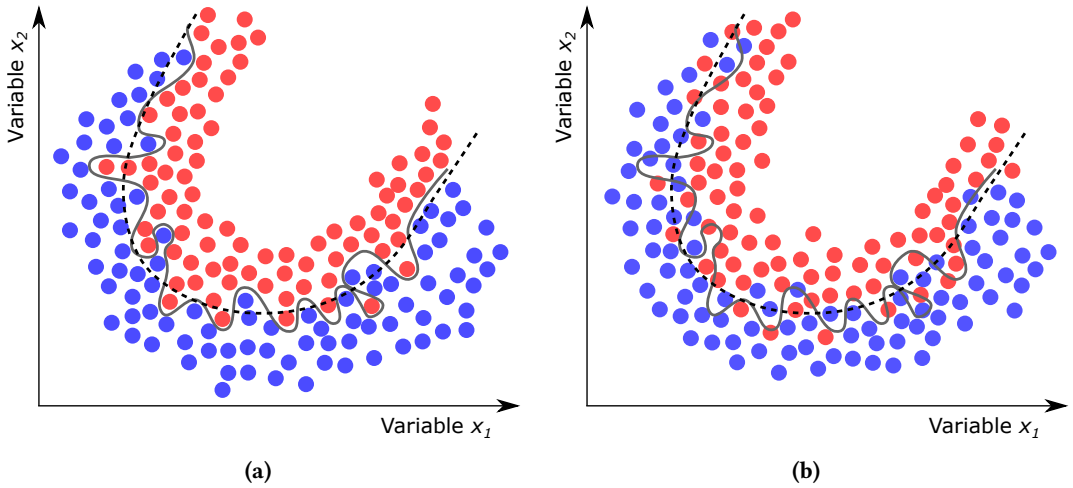


Figure 2.4.: An example of the overtraining effect. The training (a) and testing data set (b) are shown for two variables x_1 and x_2, involving signal (red) and background events (blue). The classifier learns statistical fluctuations of the training data set as features of the data (solid line), while the true separation is different (dashed line). As a consequence, applying this training to an independent testing sample results in worse performance compared to the performance on the training data set. Adapted and modified from [96].

output neuron can be used allowing multiclassification. However, all these advantages usually require a much larger data set for the training of the DNN compared to ANNs or BDTs.

2.1.4. Training and Validation

Each multivariate classifier is first trained with a subset of a simulated data set, which is then discarded from the further analysis to ensure a statistically independent result. A common problem in the training procedure is the so-called overtraining or overfitting. The problem usually occurs in cases of not enough training data or too many parameters of a classifier. In case of overtraining, the classifier learns the statistical fluctuation of the training data set as a feature, increasing the separation power when evaluating the classifier on the training data set, but decreasing it on a statistically independent data set. This effect is illustrated in Fig. 2.4. To check the training of a multivariate classifier, the separation power on the training data set is compared to a second, independent data set. A similar performance is desired and this can be quantified based on the corresponding distributions through, e. g., a (two-sample) Kolmogorov–Smirnov (KS) test [97, 98].
Different countermeasures can be carried out to counteract overtraining. In general, more complex multivariate models are more susceptible to overtraining. Thus, there is always a trade-off between optimizing the training result, while not learning statistical fluctuations. In case of BDTs, a common way is to limit the overall complexity, i. e., limit the depth in each individual tree and the overall number of trees. For neural networks, a common approach is the regularization of the error function as shown in Eq. 2.11 to penalize large weights. In addition for complex networks, such as DNNs, the dropout method can be used [99, 100]. In the training

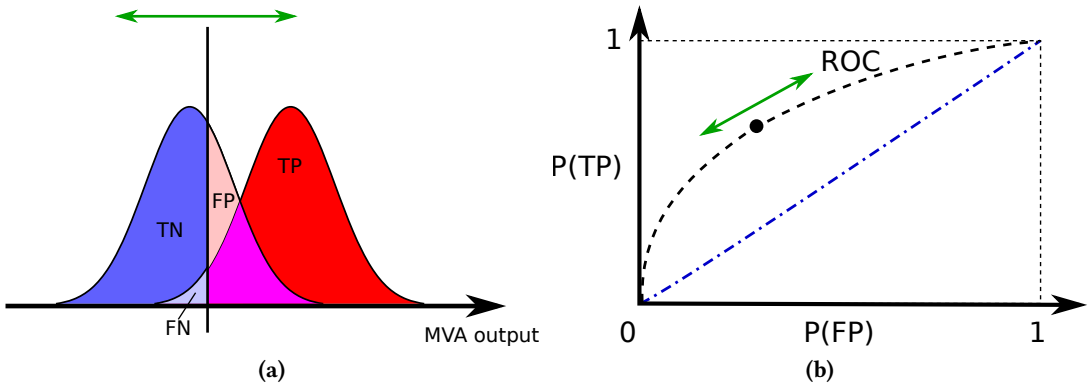


Figure 2.5.: The fractions of true-positive (TP) and false-positive (FP) results are obtained by the integral of the signal and background distribution from a given classifier value to infinity (a). Similarly, the values for the fraction of false-negative (FN) and true-negative (TN) results are determined by the integral from minus infinity to the given classifier value. The ROC curve is then obtained by scanning the classifier output range for the rate of TP and FP classifications (b). Adapted from [101].

of the DNN, every neuron has a given probability to be active, otherwise it is set to zero. This method ensures that a complex network does not only rely on a few single neurons and that the separation power is distributed across all neurons.
The performance of a classifier is quantified with the receiver operating characteristic (ROC) curve, in which the signal efficiency is shown against the background rejection or more generally speaking the true-positive (TP) rate against the true-negative (TN) rate. The curve is determined by scanning the whole range of the classifier output and by calculating both values for different values of the classifier. Instead of the TN rate, sometimes the false-positive (FP) rate is used which only mirrors the corresponding curve as $\mathrm{TN} = 1 - \mathrm{FP}$. The procedure is illustrated in Fig. 2.5. A ROC curve for a perfect classifier achieves a value of 1 for TP and TN, or in case of FP a value of 0. The best possible working point for the classifier is the one closest to this optimal value. The overall performance of the classifier can be determined by the area under the curve (AUC). A value of 1 is the score of a perfect classifier, while a value of 0.5 corresponds to an effective lower bound as this is identical to a random decision.

2.2. Statistical Methods

The statistical methods described in this section are either employed by the implementations in the THETA framework [102] or the COMBINE package [49, 103, 104], which is based on the RooFit program [105].

2.2.1. Maximum Likelihood Estimation

Under the assumption that a measurement $\vec{x}_i$ underlies a probability density function f with a given set of unknown parameters $\vec{a}$, the unknown parameters can be estimated with a maximum-

likelihood fit. To determine which set of parameters $\vec{a}$ is most likely to give the statistically independent set of results $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N$, a maximization of the likelihood function

$$L(\vec{a}) = \prod_i^N f(\vec{x}_i|\vec{a}) \tag{2.12}$$

is performed, where $f(\vec{x}_i|\vec{a})$ is the probability density function and N is the number of independent measurements. Since the analyses in this thesis rely on the counting of events in histograms, the Poisson distribution is used. The parameter of interest is the so-called signal strength modifier μ, which is a scale factor for the number of signal events with respect to the expected number of signal events. Given a histogram with multiple bins, the combined likelihood is defined as

$$L(\text{data}|\mu, \theta) = \prod_i^{\text{bins}} \frac{(\mu s_i(\theta) + b_i(\theta))^n}{n!} \cdot \mathrm{e}^{-(\mu s_i(\theta) + b_i(\theta))}, \tag{2.13}$$

with the number of observed events n and the number of predicted signal and background events s_i and b_i in bin i. The number of predicted signal and background events depends on a set of nuisance parameters θ accounting for systematic uncertainties of the measurement. To determine the maximum of the likelihood function, it is more advantageous to transform the likelihood first with the natural logarithm as this will turn the product into a sum, which is easier to calculate. In addition, instead of maximizing the function a minimization of the negative function is performed as minimization is a common task in computing. Thus, the negative log-likelihood (NLL) function is used:

$$\text{NLL} = -\ln L(\vec{a}) = -\ln \prod_i^N f(\vec{x}_i|\vec{a}) = -\sum_i^N \ln f(\vec{x}_i|\vec{a}). \tag{2.14}$$

2.2.2. Systematic Uncertainties

The predicted number of signal and background events depends on various experimental and theoretical parameters. These parameters are determined with finite precision and an uncertainty on the parameter is known, e. g., from auxiliary measurements. Uncertainties of such parameters are propagated to the analysis as systematic uncertainties, which are incorporated into the likelihood function as nuisance parameters. Depending on how the systematic uncertainty affects the prediction, a distinction between rate and shape uncertainties is made.
A rate uncertainty only affects the overall normalization of a template, i. e., the predicted histogram, by scaling each bin with the same factor. This is achieved by extending the likelihood function with a log-normal prior:

$$\pi(n) = \frac{1}{\sqrt{2\pi} n \sigma_n} \cdot \mathrm{e}^{-\frac{(\ln n - n_0)^2}{2\sigma_n^2}}, \tag{2.15}$$

with the number of events n, the mean number of events n_0 and the uncertainty σ_n. The log-normal distribution avoids unphysical results, such as $n < 0$, compared to a Gaussian distribution.

The case is different for uncertainties that affect the shape of a template where each bin is scaled independently. For this kind of uncertainty, a template morphing method is used [106, 107]. In addition to the nominal histogram, up- and down-shifted templates are provided for each systematic uncertainty, which correspond to a shift of one standard deviation in the respective direction. Based on those three values for each bin, a polynomial function is fitted to inter- and extrapolate the discrete values to a continuous function. The overall rate of the different templates is not conserved in general, therefore shape uncertainties can also include changes of the overall normalization.

2.2.3. Hypothesis Testing and Exclusion Limits

For the search of new particles in high energy physics, the signal-to-background ratio is usually relatively small and it is not possible to simply conclude whether a predicted signal is realized in nature or not. A more quantitative way to decide this is hypothesis testing, where the signal-plus-background model 's+b' ($\mu = 1$) is evaluated against the background-only model 'b' ($\mu = 0$). Both models are evaluated with a test statistic q. According to the Neyman-Pearson lemma [108], the most powerful test statistic is given by the ratio of the two likelihoods corresponding to the different hypotheses. For the purpose of this thesis, a slightly modified version of this test statistic is used [109]:

$$q_\mu = -2 \ln \frac{L(\mu, \hat{\theta}_\mu)}{L(\hat{\mu}, \hat{\theta}_{\hat{\mu}})}, \tag{2.16}$$

with $\hat{\mu}$ and $\hat{\theta}_{\hat{\mu}}$ being the values that maximize the likelihood globally, while $\hat{\theta}_\mu$ maximizes the likelihood for a given μ. The underlying probability density function f of the test statistic q_μ, derived from Monte Carlo toy experiments, can be integrated to obtain the p-value for the signal-plus-background and background-only model:

$$p_{\text{s+b}} = \int_{q_\mu^{\text{obs}}}^{\infty} f(q_\mu | \mu, \theta) \, \mathrm{d}q_\mu \tag{2.17}$$

$$1 - p_{\text{b}} = \int_{q_\mu^{\text{obs}}}^{\infty} f(q_\mu | \mu = 0, \theta) \, \mathrm{d}q_\mu . \tag{2.18}$$

The p-value can be interpreted as the probability to obtain a value of $q_\mu \geq q_\mu^{\text{obs}}$ for a given μ in the signal-plus-background hypothesis. A significance value α is defined, which needs to higher than the measured value of $p_{\text{s+b}}$ for a given μ, otherwise the signal-plus-background hypothesis is rejected with a confidence level (C.L.) of $1 - \alpha$ [110]. The whole procedure is demonstrated in Fig. 2.6. For a low expected number of signal events, the interpretation of p-values can be misleading. In this case, the probability density functions of the signal-plus-background and background-only hypotheses are almost the same, resulting in false sensitivity due to fluctuations. For this reason, a modified version of the p-value is used, the so-called CL_s limit [111, 112]. It is defined as the ratio of Eq. 2.17 and 2.18:

$$\text{CL}_\text{s} = \frac{p_{\text{s+b}}}{1 - p_{\text{b}}}. \tag{2.19}$$

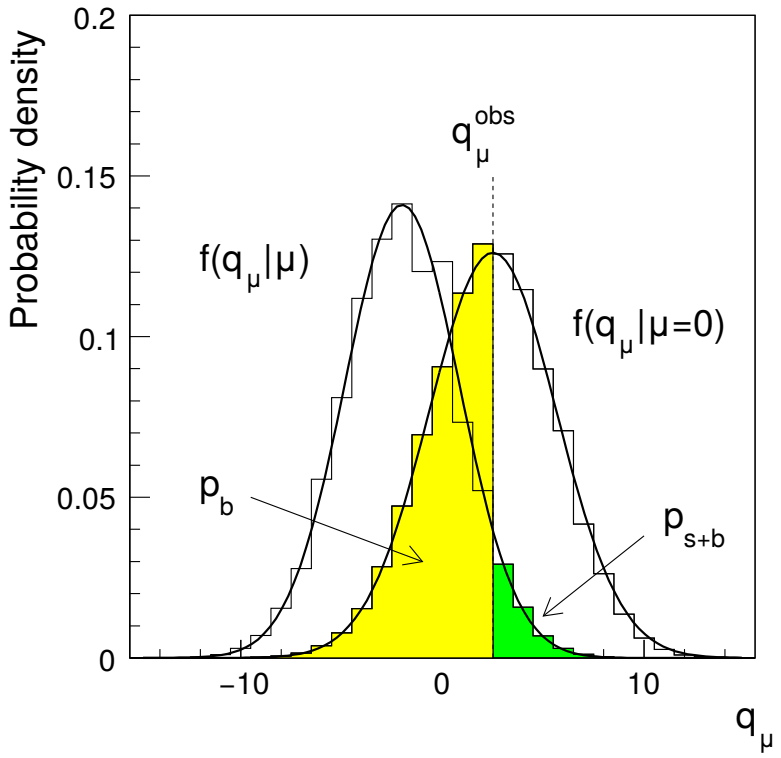


Figure 2.6.: Distribution of the two test statistic probability density functions $f(q|\mathrm{b})$ (background-only hypothesis) and $f(q|\mathrm{s}+\mathrm{b})$ (signal-plus-background hypothesis). To determine the p-values for both hypotheses, the integral of the probability density function is calculated starting from minus infinity to the observed value of the test statistic q_μ^{obs} in case of the background-only hypothesis and from q_{obs} to infinity for the signal-plus-background hypothesis. Illustration adapted from Ref. [109].

It is convention in high energy physics to exclude models with a C.L. of 95%, meaning that the signal strength modifier μ is adjusted until a value of $\mathrm{CL_s} = 0.05$ is reached, excluding all models with higher μ.
The exclusion limit based on real experimental data is called the observed exclusion limit. However, optimizing the analysis towards better results on data is not desired as it will bias the result. To avoid such a potential bias, even if this does not happen intentionally, the analysis is optimized and tested only on the prediction and the expected upper limit, determined from toy experiments. The analysis only determines the observed limit if the analysis concept and workflow is settled.
Expected limits can be obtained by performing multiple toy experiments and deriving the median of the resulting q_μ^{exp} distribution. In addition, uncertainties on the expected limit are calculated by the one and two standard deviations of the distribution. In order to derive a valid distribution of q_μ^{exp}, many toy experiments are necessary, requiring a large amount of computation power. This method is known as the full $\mathrm{CL_s}$ limit. Alternatively, an asymptotic formula based on the so-called Asimov data set can be used [109]. The method is based on

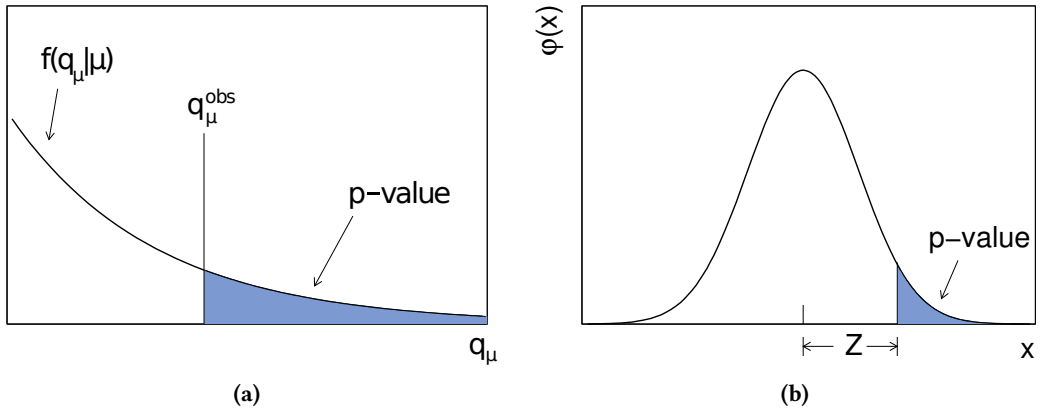


Figure 2.7.: The obtained p-value (a) can be translated into a significance in terms of standard deviations of a Gaussian distribution (b). Adapted from [109].

the work of Wald [113] and Wilks' theorem [114], which states that asymptotically the test statistic for a large data set follows an analytical χ^2 distribution. By employing this method, the computation-intensive calculation of toy experiments can be avoided.

2.2.4. Statistical Significance

In case of an excess of data compared to the background-only prediction, the possible signal needs to be quantified. In particle physics, the p-value is converted into the significance Z, defined as:

$$Z = \Phi^{-1}(1 - p). \tag{2.20}$$

Here, Φ^{-1} represents the quantile of the Gaussian distribution. The principle is illustrated in Fig. 2.7. The value of Z directly corresponds to the number of standard deviations for which a Gaussian-distributed variable would be away from the mean of the distribution in case of a valid background-only hypothesis. As neither the p-value nor the significance Z provide any specific value for which an observed signal can be claimed or not, it is conventional to set arbitrary barriers for Z. If the significance exceeds a value of three, evidence for the given signal can be claimed. An actual observation or discovery requires a significance of $Z > 5$. As for the exclusion limit calculation, the significance can be determined based on the prediction, called expected significance, or as observed significance based on the data.

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

The artificial production of heavy elementary particles, such as the top quark and the Higgs boson, is only possible at large particle accelerators due to the high amount of energy needed. A large variety of particle accelerators exists covering a large spectrum of collision energy and different accelerated particles. The largest and most energetic particle accelerator today is the Large Hadron Collider (LHC). Accelerating and colliding particles is only one part of the experimental apparatus, as the collision products have to be measured with high accuracy. This is achieved in the Compact Muon Solenoid (CMS) experiment, a multipurpose detector for analyzing a broad range of physics processes. The complete experimental machinery, starting from initial particles in the accelerator to the signals in the detector, is explained in following chapter.

3.1. The Large Hadron Collider

With a circumference of around 27 km [115, 116], the LHC is the largest machine ever built by mankind. It is a nearly ring-shaped particle accelerator for protons and nuclei with several pre-accelerators to reach new energy frontiers. The whole complex is located at the European Organization for Nuclear Research, which emerged from Conseil Européen pour la Recherche Nucléaire (CERN), in Geneva (Switzerland) [117]. After the shutdown of the Large Electron Positron (LEP) collider in 2000 and its decommissioning in 2001, the already existing tunnel was used for the LHC, which is built below the state territory of France and Switzerland, about 100 m below the surface. While the energy loss due to bremsstrahlung limits the maximum energy of circular lepton colliders such as LEP, the LHC can reach center-of-mass energies almost two orders of magnitude higher than LEP.
The acceleration process starts with a bottle of hydrogen, from which the protons are extracted by separating them from the electrons through an electric field. They are then passing the Linear accelerator 2 (LINAC 2) where the protons are accelerated through radio-frequency cavities up to an energy of 50 MeV [118]. This energy is further increased by injecting the protons into the Proton Synchrotron Booster ring (157 m circumference), where their energy can be raised up to 1.4 GeV [119], before they reach the Proton Synchrotron (PS) [120]. The PS with a circumference of 628 m is the oldest accelerator still in operation at CERN, started in 1959 as the particle accelerator with the highest energy at that time. Today it still serves as pre-accelerator for the LHC delivering protons and heavy ions at energies up to 25 GeV. From there, the particles are inserted into the last pre-accelerator before the LHC, the Super Proton Synchrotron (SPS) with a circumference of nearly 7 km [121]. At the SPS, where the W and Z

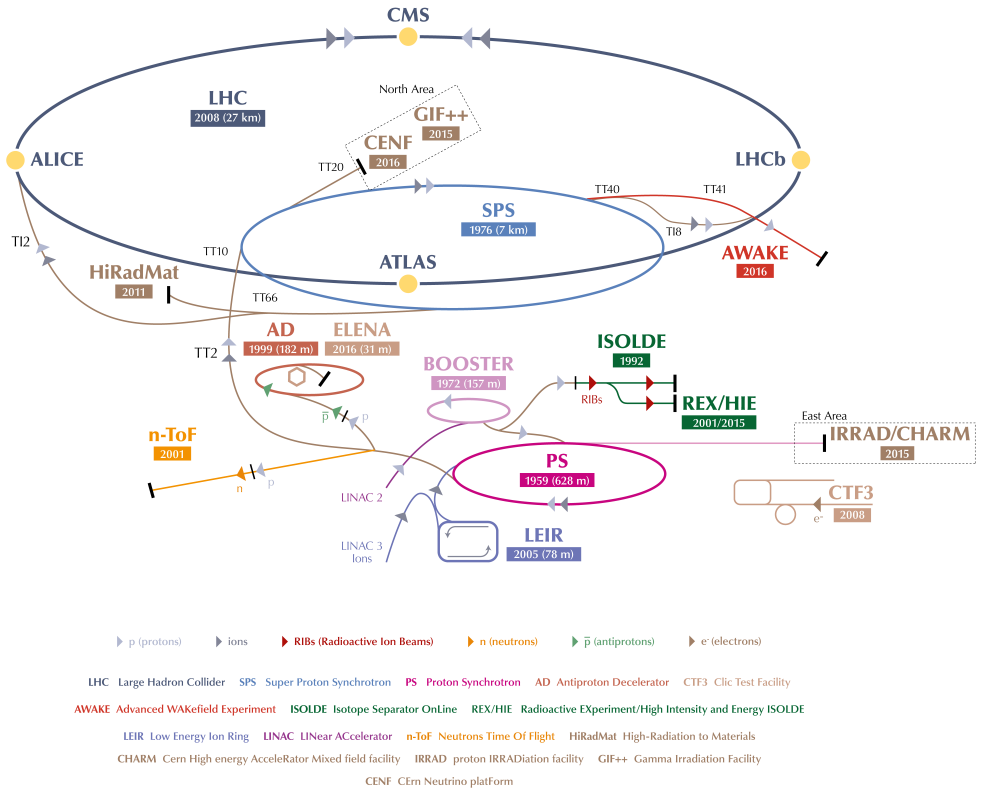


Figure 3.1.: Complete overview of the accelerator complex located at CERN. The entire path of the protons is shown, starting from LINAC 2 until the LHC ring. In addition, all other experiments involving the study of different particles are highlighted [124].

bosons were first discovered [122, 123], the particles can reach an energy of 450 GeV before they are injected into the main LHC ring. An illustration of the accelerating complex is provided in Fig. 3.1. The LHC ring consists of two beam pipes next to each other, where the particles are traveling either clock- or counterclockwise through the 27 km long accelerator. The beam itself is not a constant stream of particles, but instead consists of several packets of particles, called bunches. Each bunch contains of the order of 10^{11} particles and the ring can be filled with a total of 2808 bunches, separated by a time interval of 25 ns. The two beams are bent by 1232 dipole magnets on a curved trajectory through a magnetic field of up to 8.33 T. In order to reach such a high magnetic field the dipole magnets are cooled down by liquid helium to a temperature of 1.9 K to reach a superconducting state. Each dipole magnet weighs 35 t and is 15 m long. To counteract the widening of the beam through electromagnetic interactions of the particles among themselves, the beam is focused by 392 quadrupole magnets along the ring. The LHC is designed to provide beams of protons with an energy of 7 TeV, providing a total center-of-mass energy of 14 TeV.
After the particles have gained their maximum energy, the beams are bent to cross each other at

certain intersection points to allow interactions of particles from opposed bunches. A measure of the interaction rate is the instantaneous luminosity

$$L = f \cdot \frac{nN_1N_2}{4\pi\sigma_1\sigma_2}, \tag{3.1}$$

with the beam revolution frequency f, the number of bunches n, the number of particles N_i per bunch in beam 1 and 2 and the cross-section profile σ_i. The intensity of the beam decreases slowly over time due the interactions of particles and other effects, but the so-called fill will be kept several hours in the LHC before it is more advantageous to eject the current beam from the ring and start over with a new fill of higher intensity. To provide a measure of accumulated collisions, the integrated luminosity

$$L_{\text{int}} = \int L \, \mathrm{d}t \tag{3.2}$$

is used. The number of produced particles N of a given type can then be expressed as

$$N = L_{\text{int}} \cdot \sigma \tag{3.3}$$

with the production cross section σ of the process.
The LHC started operation in 2008, but because of a magnetic quench destroying several dipole magnets, the first data taking was delayed until 2010 with lowered energy. In the Run I of the LHC, collisions were produced at a center-of-mass energy of 7 TeV (2010-2011) and 8 TeV (2012). The delivered integrated luminosity during this run corresponds to 6.1 fb^{-1} and 23.3 fb^{-1} at 7 and 8 TeV, respectively. Afterwards the LHC was shut down for two years from 2013 to 2014 for improvements of the accelerator and detectors to prepare the machine for a higher center-of-mass energy of 13 TeV. The Run II of the LHC started in 2015, delivering a total integrated luminosity of 96 fb^{-1} until the end of 2017. The Run II of the LHC will continue until the end of 2018. An overview of the luminosity delivered by the LHC during its operation time is provided in Fig. 3.2.
The four main particle detectors of the LHC are located at the intersection points of the two beams to allow the investigation of collision products, each with different specialization:

- The goal of **ALICE** (A Large Ion Collider Experiment) is the observation of heavy nuclei interactions at high energies. Similar conditions have been present shortly after the Big Bang and could be used to better understand the effects of QCD through studying the quark-gluon plasma state at high energy densities.

- The **CMS** (Compact Muon Solenoid) experiment is a multipurpose detector designed to search for the Higgs boson, new physics beyond the SM and precision tests of SM observables at high center-of-mass energies.

- Similary, the **ATLAS** (A Toroidal LHC ApparatuS) detector covers a broad range of physics topics, which provides an independent measurement to the topics covered by the CMS experiment.

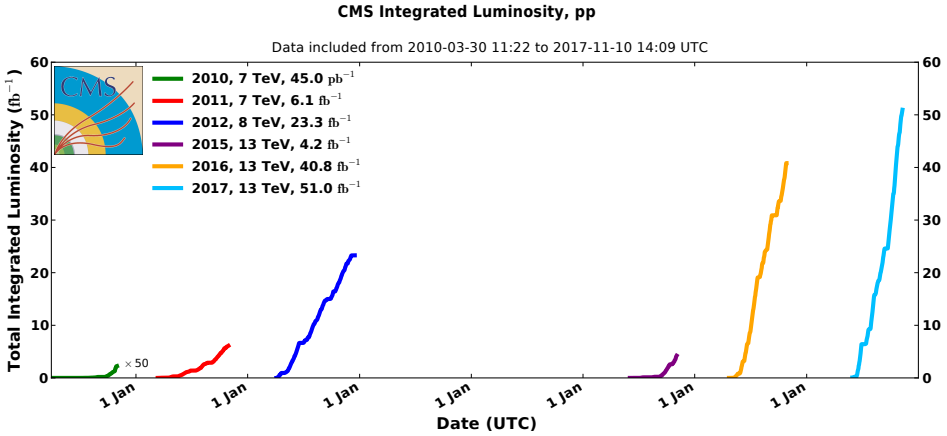


Figure 3.2.: The integrated luminosity recorded by the CMS experiment during the entire LHC operation time. Run I of the LHC involved collisions at a center-of-mass energy of 7 and 8 TeV in 2010-2012. After a shutdown in 2013 and 2014, the energy was raised to 13 TeV in 2015. In addition, the LHC increased the instantaneous luminosity to be able to provide a higher collision rate [125].

- More specialized in bottom quark physics is the **LHCb** (Large Hadron Collider beauty) experiment, which is built asymmetrically into one direction. Exemplary for LHCb are studies of CP-violation in b hadrons and search for rare decays.

In addition to those four major experiments there are also the LHCf (LHC-forward), MoEDAL (Monopole and exotic particle detector at the LHC) and TOTEM (Total cross section, elastic scattering and diffraction dissociation measurement at the LHC) experiments, which share their location with ATLAS, LHCb and CMS, respectively.

3.2. The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) experiment is a multipurpose detector with several different layers built around the interaction point of the two colliding beams [126]. It is located in a cavern below the surface at the so-called Point 5 of the LHC near Cessy (France). The cavern has a dimension of about 50 m×25 m×25 m. The detector itself is of cylindrical shape and is 28.7 m long with a diameter of 15 m. With all its components the CMS detector weighs 14 000 t, where the return yoke alone weighs 12 500 t. From the interaction point to the outside the CMS detector consists of a tracking system built of silicon, electromagnetic and hadron calorimeters, a superconducting solenoid and a muon system. A schematic overview of the detector and the subcomponents is shown in Fig. 3.3. To describe spatial quantities, a right-handed Cartesian coordinate system is used at the CMS experiment. It is defined such that the x axis points towards the center of the LHC and the y axis towards the surface. Thus, the z axis is aligned with the beam direction. Since gravity can be neglected it is advantageous to use spherical coordinates, as the interactions are invariant under rotations around the beam pipe. The orientation around the beam pipe, i. e., the x-y plane, is defined by the azimuthal angle ϕ. The polar angle θ describes

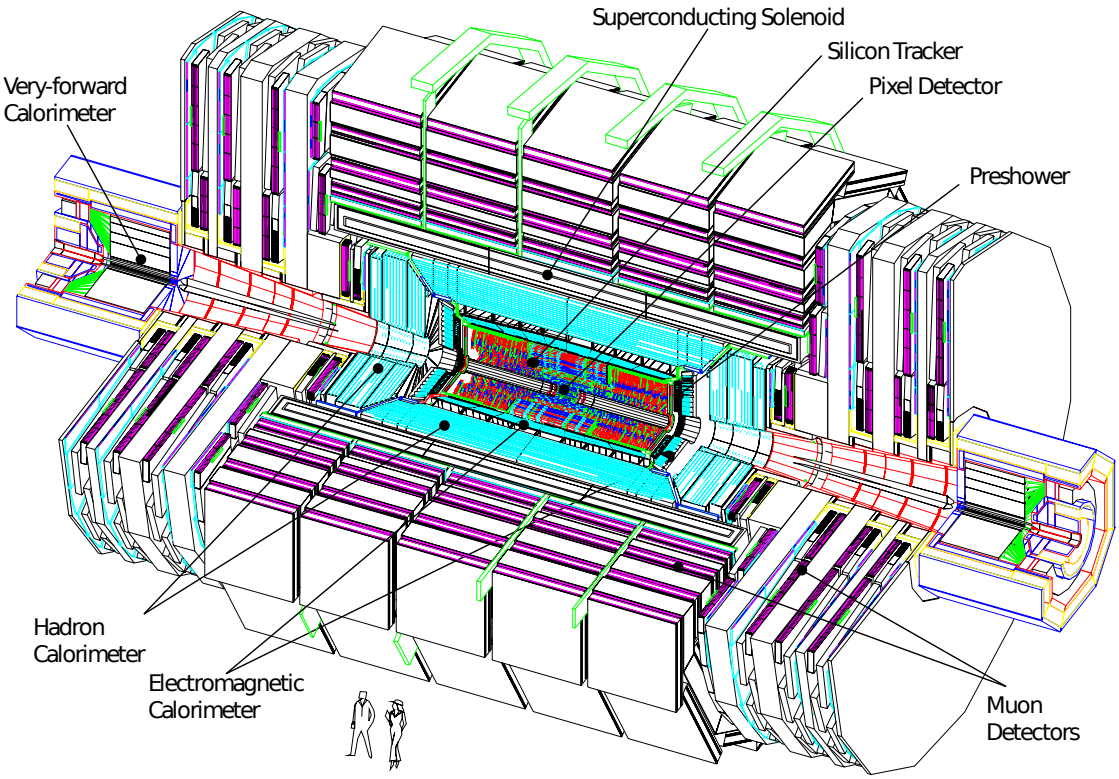


Figure 3.3.: An overview of the Compact Muon Solenoid (CMS) experiment located at the LHC. Closest to the interaction point in the middle of the detector are the pixel detector and the silicon tracker. They are surrounded by the electromagnetic and hadron calorimeter, supported with a preshower system in the endcap and a very-forward calorimeter in forward direction. The calorimeters are enclosed in a superconducting solenoid, providing a strong magnetic field. The outermost part of the detector is a system dedicated to the measurement of muons [126].

the angle with respect to the beam direction. It is also useful to define the Lorentz-invariant rapidity as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{3.4}$$

with the energy E and the momentum in z direction p_z. To avoid the dependence on energy and momentum, a related quantity, the pseudorapidity, can be defined as

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right), \tag{3.5}$$

which only relies on the polar angle θ. For massless particles, both quantities are equal. Pseudorapidity values range from zero for a particle perpendicular to the beam axis up to infinity for particles parallel in the beam direction. Negative values are used for the opposite direction. Together with the azimuthal angle ϕ, a combined angular separation ΔR in the η-ϕ plane is defined by:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{3.6}$$

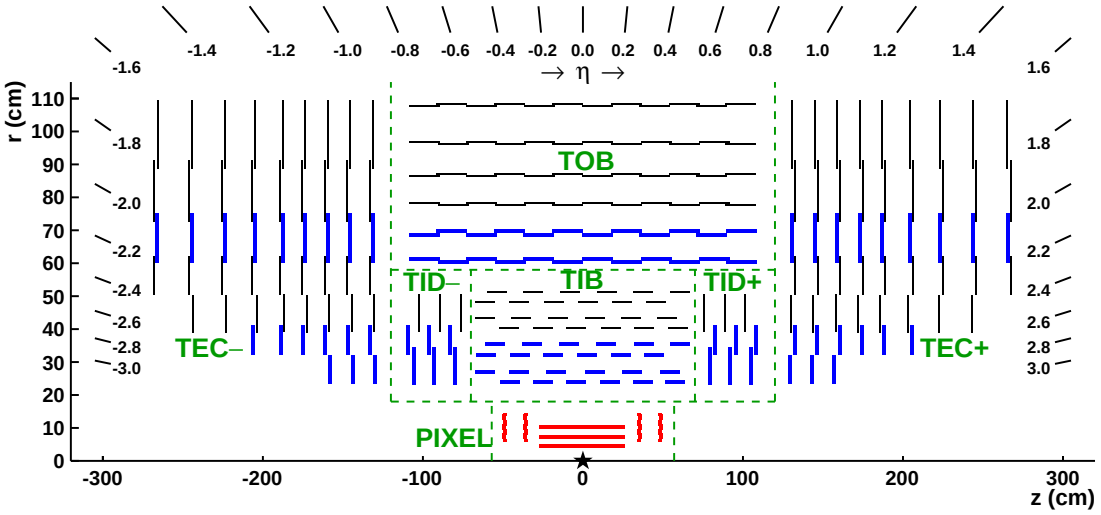


Figure 3.4.: Schematic overview of a segment of the tracker system, taken from [127]. Closest to the interaction point are the silicon pixel detectors (PIXEL), consisting of three barrel layers and two endcap layers. The amount of pixel layers was increased to four and three layers in the barrel and endcap, respectively. The pixel detectors are surrounded by tracking modules with multiple layers, arranged in the Tracker Inner Barrel (TIB), the Tracker Inner Disk (TID), the Tracker Outer Barrel (TOB) and the Tracker End-Cap (TEC).

In the following, the various subcomponents of the CMS detector are described in detail, starting from the innermost part.

3.2.1. Tracking System

Closest to the interaction point is the tracking system of the CMS detector. Its purpose is to measure hits of electrically charged particles traversing several layers of the tracker system. A track of the particle is then fitted from the various hits. The charged particles are bent in the perpendicular magnetic field, which allows to measure their transverse momentum and the sign of their charge from the curvature of the track. It is crucial to accurately measure all particle tracks to locate the position of the hard scattering process. The corresponding techniques and algorithms will be explained in Section 4.2.2.
The tracking system is based on the semiconductor silicon. A diode is built up from positively and negatively doped silicon and operated in reverse bias. Charged particles create electron-hole pairs while traversing the silicon detector material, which induce a current. This current is amplified and used to measure a signal. Silicon detectors allow more precise and faster measurements of particle hits, as the individual elements can be built smaller compared to e. g. wire chambers. However, they are much more expensive than other particle detectors. An overview of the complete CMS tracking system is provided in Fig. 3.4. The first part of the tracking system is made of silicon pixel modules, called the pixel tracker [128]. It consists of a barrel with three layers (BPIX) and two endcaps with two layers each (FPIX), covering an area up to $|\eta| < 2.5$. The cylindrical layers in the barrel are 53 cm long and have diameters of 4.3 cm, 7.3 cm and 10.2 cm. The endcaps are ±34.5 cm and ±46.5 cm away from the collision point. All

layers together consist of 66 million pixels, most of the size 150 µm by 100 µm, and are arranged on 1440 individual modules covering an area of 1 m^2. A spatial resolution up to 10 µm can be achieved by using the Lorentz drift of electrons and holes to optimally exploit charge sharing between different cells.
The pixel detector was replaced during the end-of-the-year shutdown of the LHC between 2016 and 2017 within the so-called Phase 1 upgrade [129]. Improvements to the pixel detector were necessary due to radiation damage of the old modules and to prepare the pixel tracker for the increased luminosity from 2017 on, resulting in a higher number of tracks per collision. Along with improvements of cooling, readout chips and powering, the number of layers was increased to four and three in the barrel and the endcap, respectively. For the additional layer in the barrel, it was necessary to replace the old beam pipe with a smaller one. The layers in the barrel now range from 3 cm up to 16 cm. With additional layers, the combinatorics for track-fitting can be reduced and a broader distance of individual layers allows a more precise result. The total number of pixels was increased to 124 million. Additional updates are planned in the Phase 2 upgrade for even higher luminosities [130].
The other part of the tracking system consists of silicon strip detectors, which surround the inner pixel detector. In total there are 15,148 strip detector modules with about 9.6 million readout channels covering an area of nearly 200 m^2. The geometry is similar to that of the pixel detector, as the different layers of the strip detector are organized in a barrel and endcap sections. There are ten layers of strip detectors in the barrel, which is built up by four layers in the Tracker Inner Barrel (TIB) and six layers in the Tracker Outer Barrel (TOB). The TIB is based at a distance between 20 cm and 50 cm from the interaction point and consists of two layers of double-sided (stereo) and two layers of single-sided modules, while the TOB is located from 55 cm to 110 cm and has also two layers of double-sided modules, but four layers of single-sided modules. Hits from particles can be determined with an accuracy of 30 µm to 50 µm. The endcaps are also organized in two different subsystems, the Tracker Inner Disk (TID) and Tracker End-Cap (TEC), each containing single- and double-sided modules.

3.2.2. Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is the next part of the CMS detector and built around the tracking system [131]. The purpose is to measure the energy of electromagnetically interacting particles by absorbing them. This is achieved through a cascade of electrons, positrons and photons, called a particle shower. When entering the detector material electrons and positrons emit photons through the effect of bremsstrahlung and photons convert into electron-positron pairs. This reaction chain is repeated until the energy of photons is lower than the mass of an electron-positron pair. The photons are then absorbed by the detector material, which will release this energy in form of scintillation light. Electrons and positrons also contribute to the energy deposit in the detector material when their energy is relatively low. The intensity of the scintillation light is proportional to the deposited energy and thus a direct measure of the energy from the single initial particle.
Requirements on a scintillator are a material capable of emitting scintillation light and a high density of the material, which increases the probability of interaction and, therefore, decreases the total length needed to contain the complete particle shower. This can either be achieved

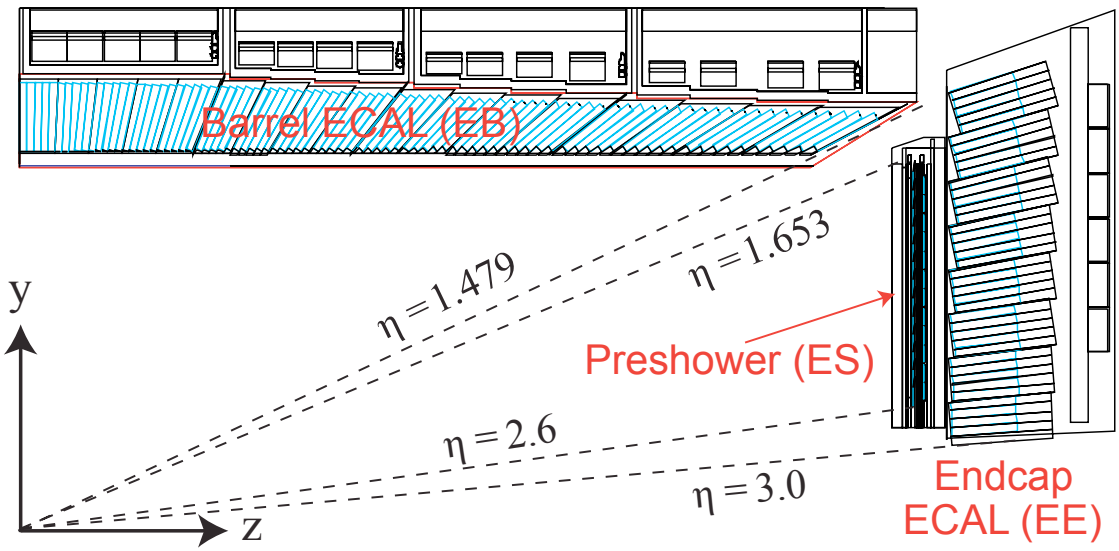


Figure 3.5.: Overview of one quadrant of the CMS electromagnetic calorimeter (ECAL). The ECAL is built up from $PbWO_4$ crystals, arranged in a Barrel ECAL (EB) and Endcap ECAL (EE) region. The endcap region is shielded by a preshower system (ES) to support identification of nonprompt photons from π^0 meson decays [132].

through a homogeneous calorimeter of a single material or a so-called sampling calorimeter with alternating layers of absorber and scintillation material. The ECAL used in the CMS detector is a homogeneous calorimeter consisting of 75,848 lead tungstate ($PbWO_4$) crystals. Advantageous of lead tungstate are a radiation length of $X_0 = 8.8\,\text{mm}$, describing the distance for an electron energy loss of $^1/_e$, and a Moilère radius of $R_M = 2.2\,\text{cm}$, characterizing the transverse dimension of the shower. In addition, the scintillation time is short enough to be read out while operating the LHC with a bunch spacing of 25 ns. Similar to the tracking system, the ECAL is separated into a barrel section (EB) and an endcap section (EE). The layout of the calorimeter is shown in Fig. 3.5. The crystals in the EB, arranged in 36 supermodules with 1,700 crystals each, have an area of 2.2 cm×2.2 cm facing the interaction point, however they are slightly inclined to avoid blind spots between the crystals. With a length of 23 cm, they cover more than 25 radiation lengths, providing enough material to capture the entire particle shower inside the crystal. The scintillation light is captured by photo diodes. The EB provides a coverage up to $|\eta| \leq 1.479$. The crystals in the EE are broader with a size of 2.86 cm×2.86 cm, providing a slightly worse resolution compared to the EB. To improve the rejection of nonprompt photons stemming from neutral pion decays, a preshower (ES) is located in front of the EE. The coverage of the ECAL is extended with the EE up to $|\eta| \leq 3.0$.
The energy resolution of an electromagnetic calorimeter is given by the following formula [133]:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{s}{\sqrt{E}}\right)^2 + \left(\frac{n}{E}\right)^2 + (c)^2, \tag{3.7}$$

where σ_E is the uncertainty on the measured energy E and s, n and c are coefficients to parametrize different effects. Since the shower evolution is a stochastic process, the coefficient s describes

the uncertainties of modeling the particle shower and the term itself has a weak energy dependence. The coefficient n describes the electric noise of the system and the resulting absolute uncertainty is independent of the energy of the initial particle itself. Any constant contributions are parametrized with the coefficient c, e. g., instrumental effects. It is important to keep the latter contribution small, especially in experiments such as the detectors at the LHC, as this term will dominate the calorimeter resolution in the high-energy range. For the CMS ECAL, these coefficients have been measured with an electron test beam which gave the following result [134]:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.3\%)^2. \tag{3.8}$$

3.2.3. Hadron Calorimeter

The hadron calorimeter (HCAL) of the CMS detector serves the purpose of measuring the energy of strongly interacting particles which traverse the previous subdetectors with no interaction or minimal energy loss [135]. In contrast to the ECAL, the HCAL is a sampling calorimeter built up on alternating layers of brass absorber and plastic scintillator. The mechanism of measuring the energy of particles is similar to the ECAL with an electromagnetic shower. While propagating through the absorber, the particles are scattered inelastically with the detector material creating additional hadrons such as pions. These secondary particles can also create additional particles, which results in a hadronic shower. Particles from this shower are then captured mostly by atoms of the absorber material, raising them into an exited state. The deexcitation of these atoms is then occurring under the emission of ultraviolet light, which is sampled by the scintillator layers. The wavelength of the light is then shifted via fibers to the optical spectrum and read out by photo diodes.
The CMS HCAL is organized in a hadron barrel (HB) and a hadron endcap (HE), complemented by a hadron outer (HO) barrel and a hadron forward (HF) calorimeter, which utilizes Cherenkov radiation in quartz fibers due to the high flux in forward direction. All parts are depicted in Fig. 3.6.

3.2.4. Superconducting Solenoid

The stable particles produced in collisions at the LHC are light compared to their energy, thus their velocity is a non-negligible fraction of the speed of light. To ensure that the trajectory of these fast-traveling particles can be bent to measure their momentum, a powerful magnetic field is required. This is achieved in the CMS experiment by a solenoid, providing a magnetic field up to 4 T [136, 137], although, for longevity, only a field strength of 3.8 T is used [138]. The solenoid consists of a superconducting coil with a diameter of 6 m and a length of 12.5 m, enclosing the HCAL of the CMS detector and producing a field of 3.8 T in the inside. This enormous magnetic field strength can only be reached by cooling the magnetic coil down to 4.5 K, where the NbTi conductor is in a superconducting state, allowing the current of 18 kA to flow without electrical resistance. The total energy stored in the coil corresponds to 2.3 GJ. The coil is surrounded by an iron return yoke, consisting of several layers, to guide the magnetic field lines, serve as absorber plates for the incorporated muon system and to increase the bending power. Summing

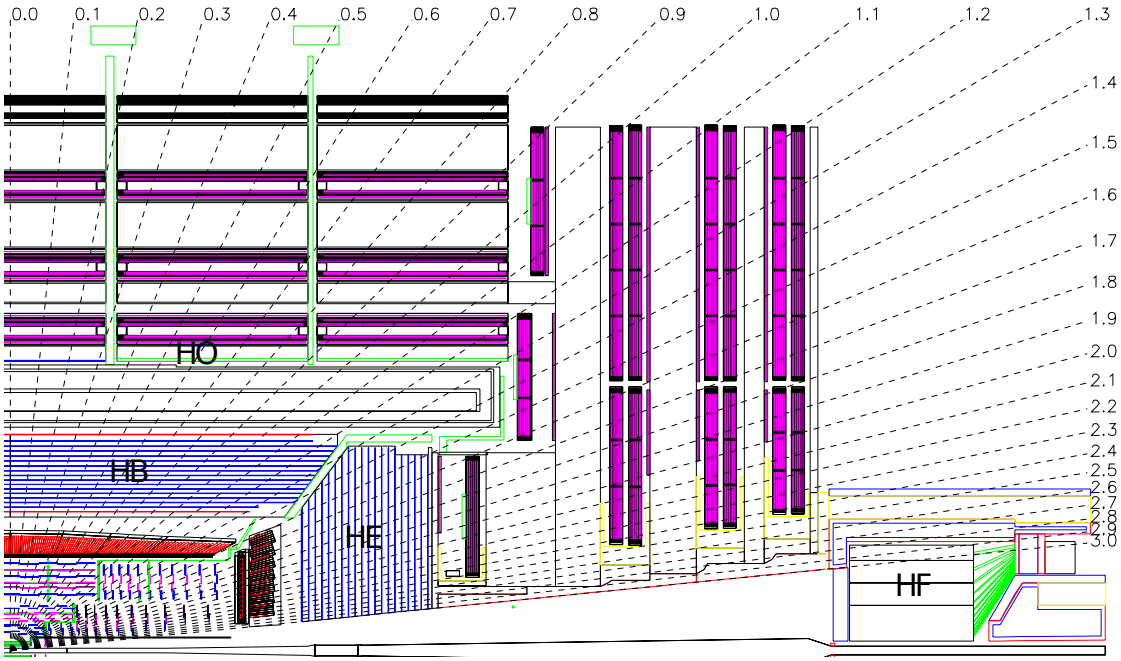


Figure 3.6.: An overview of the different components of the hadron calorimeter (HCAL). The hadron barrel (HB) and hadron endcap (HE) are surrounded by the superconducting solenoid. Any remaining energy is absorbed by the hadron outer (HO) barrel. A special hadron forward (HF) calorimeter is installed beginning at $|\eta| = 3$ [126].

up all layers, which enlarge the diameter of the CMS detector up to 14 m, the complete return yoke weighs 10,000 t. The yoke is built up from five rings of flat iron plates and three endcaps on each side. Every ring consists of three layers, where every layer is built up from twelve flat iron plates arranged around the interaction point. An illustration of the CMS magnet is provided in Fig. 3.7.

3.2.5. Muon System

The outermost part of the CMS detector is the muon subsystem, combining three different gaseous detectors to measure the properties of muons [139–142]. As muons produced in particle collisions at the LHC typically have an energy of the order of GeV, they are minimum ionizing particles and thus are able to penetrate all previous layers of the detector with minimal interaction. Conversely, this means that muons are the only particles able to reach the outermost part of the detector, making a dedicated particle identification unnecessary. The complete muon subsystem consists of drift tubes (DTs), cathode strip chambers (CSCs) and resistive plate chambers (RPCs). All these different categories of detectors are incorporated between the several layers of iron return yoke in the already familiar fashion of barrel and endcap geometry. A sketch of the different subsystems inside the detector is provided in Fig. 3.8. In the barrel, where the flux of muons is not too high, DTs are used to determine the position of a traversing muon. The DTs are filled with a combination of Ar and CO_2. Inside the drift tubes, a strong electric field is present between the walls and the wire in the center of the tube. When a muon travels

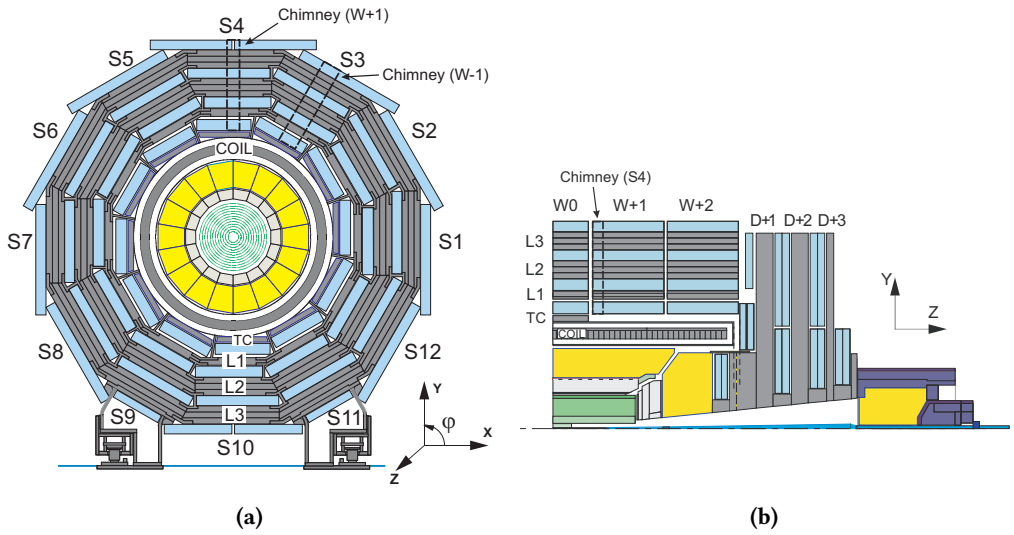


Figure 3.7.: Illustration of the CMS magnet in the transverse (a) and longitudinal (b) plane. The solenoid consists of five rings in the barrel, each built up from twelve segments. The barrel is enclosed with three endcap layers on each side [138].

through the gas it will separate electrons from the gas ions which are then accelerated by the applied electric field. This will create additional free electrons through impacts with other ions, resulting in an avalanche of electrons which can be measured as a current at the anode wire. The DTs are only located in the barrel section of the CMS detector and provide coverage up to $|\eta| < 1.2$ with a total of about 172,000 wires, each 2.4 m long. A DT chamber is built up from either three (MB1-3) or two (MB4) supermodules (SL). Each SL consists of four layers of anode wires and cathode walls. To allow a three-dimensional position measurement, the SL of the three inner layers are shifted by 90°: while for the two outer SL of a DT chamber, the wire is parallel to the beam, allowing a position measurement in the r-ϕ plane, the wires of the center SL are perpendicular to the beam, revealing the position in z direction. In case of the outermost layer with only two SL, only a measurement in the r-ϕ plane is provided. A spatial resolution of 100 µm can be achieved.
In contrast to the barrel of the CMS detector, the position of muons in the endcaps is determined by CSCs. A single CSC is a multiwire proportional chamber and is built in a trapezoidal geometry. Each CSC consists of seven panels made of epoxy with a thin copper coating serving a cathode. Each panel is interleaved with a layer of wires for the anodes, such that there are six layers of wires in each CSC. The working principle is similar to that of the DTs. The CSC modules are arranged cylindrically around the interaction point with the strips, on which the wires are mounted, pointing inwards, allowing a measurement of the radial coordinate. This endcap design allows a muon position coverage up to $|\eta| < 2.4$ with an overlap at $0.9 < |\eta| < 1.2$ with the DTs in the barrel. There are a total of 468 CSCs in operation at the CMS detector with 220,000 readout channels for cathode strips and 180,000 for anode wires.
Although the wire chamber technology of the DTs and CSCs provides good spatial resolution, the time and momentum resolution can be significantly improved by additional RPCs. An RPC

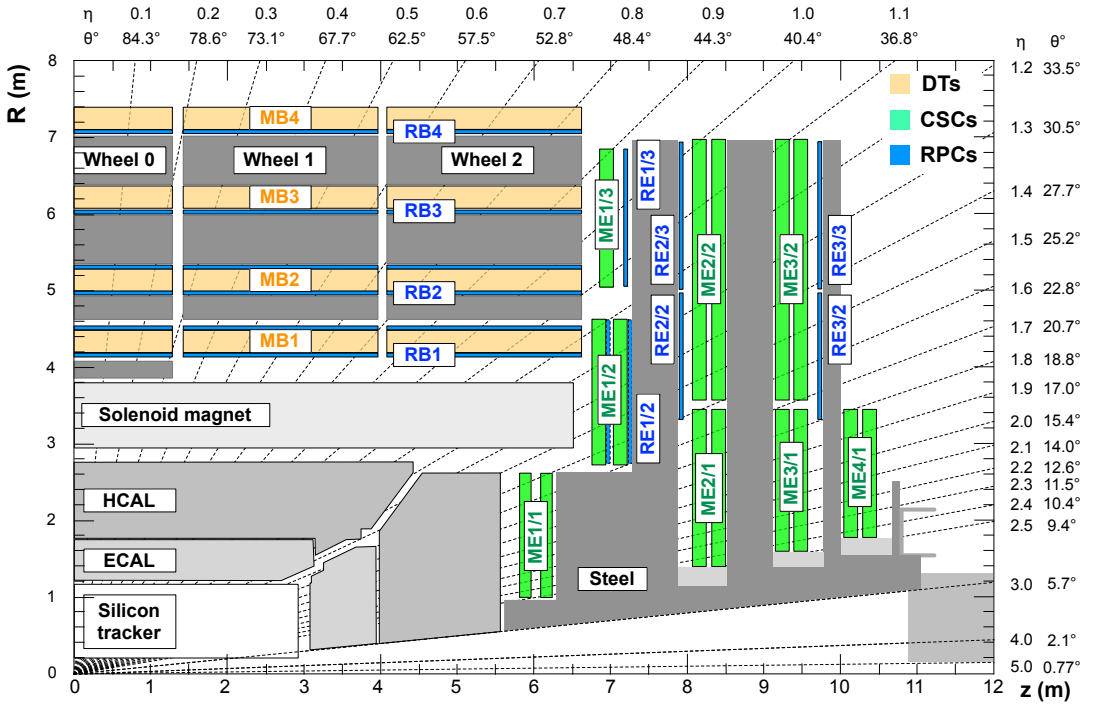


Figure 3.8.: The setup of the muon system of the CMS detector. Drift tubes (DTs) are located in the barrel region in between the return yoke of the magnet, while cathode strip chambers (CSCs) are installed in the endcap section of the detector. Resistive plate chambers (RPCs) support the muon system with fast information for the trigger [143].

consists of two high resistivity plastic plates with a gas between them. Similar to DTs and CSCs, muons create an electron avalanche inside the volume, which is read out by metallic strips. The hit pattern allows a determination of the muon's initial momentum. Most advantageous is the fast readout time of only a few nanoseconds, which allows the RPCs to be used as a trigger for interesting events. RPCs are installed in the barrel and endcap regions of the detector close to the other muon detectors to complement them. There are a total of six layers of RPCs in the barrel and three layers in the endcap, able to detect muons up to $|\eta| < 1.6$. The endcap section was upgraded during the first long shutdown (LS1) with a fourth layer of RPCs and an extended coverage up to $|\eta| < 2.1$.

3.2.6. Trigger System and Data Acquisition

The LHC operates at a bunch spacing of 25 ns, meaning a collision rate of 40 MHz. With protons colliding and the design luminosity of $L = 10^{34}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}$, this results in 25 interactions per bunch crossing on average. It is technically not possible to save all measured data from the detector for every collision to be analyzed at a later date. Instead, only events with interesting signatures are stored. This is achieved through a very fast decision from the trigger system. Based on certain requirements, such as a high-momentum muon present in the event, a decision whether the event should be stored for later usage is made within milliseconds. During this time the full

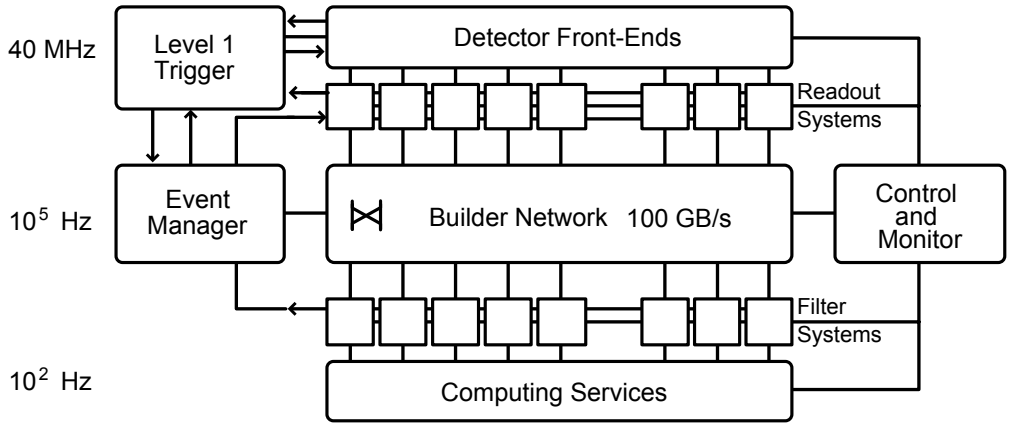


Figure 3.9.: Illustration of the trigger and data acquisition system of the CMS experiment [146]. Given the 25 ns bunch spacing of protons in the LHC, the initial data-taking rate is 40 MHz. In the first step, the rate is reduced to 100 kHz utilizing custom hardware to search for interesting signatures in the event. Computer algorithms on standard hardware can reduce the rate down to 400 Hz, at which point the rate is low enough to store all the remaining events for further analyses. In Run II of the LHC, the HLT rate can reach up to 1 kHz.

event data is held in pipeline buffers until the trigger system decides whether the event is worth keeping or it can be safely discarded. Through this preselection, the rate of collisions can be reduced by several orders of magnitude, making it possible to store all the remaining events for physics analyses.
The CMS trigger system consists of two different levels, the Level-1 (L1) trigger and the high-level trigger (HLT) [144–146]. The L1 trigger is implemented in custom hardware, such as field programmable gate arrays (FPGA) or application-specific integrated circuits (ASIC), and receives raw inputs from the calorimeters and the muon system. The initial collision rate of 40 MHz can be reduced with the L1 trigger down to a rate of 100 kHz. In contrast to the L1 trigger, the HLT is software-based and runs on standard computer farms. The full event data is coarsely reconstructed, which allows for more sophisticated trigger decisions. By employing the HLT, the event rate can be further reduced. This results in a final rate of about 1 kHz for Run II of the LHC.
The HLT of the CMS experiment is embedded in the data acquisition (DAQ) chain of the detector. The principle structure is shown in Fig. 3.9.

3.2.7. Computational Infrastructure

After the significant reduction of experimental data through the trigger system, the output of the DAQ system is directed to computer farms for storage, full reconstruction and analyses. There is no sufficient large enough computer center located at the CMS experiment for this task. Instead, a distributed computing model, the Worldwide LHC Computing Grid (WLCG) [147, 148], is employed. The WLCG is a grid-like infrastructure of computing centers across the globe, shared among all LHC experiments, to ensure sufficient computation power, storage and availability of data to every user. It is hierarchically structured in several layers, called tiers [149]. An overview of the WLCG is provided in Fig. 3.10. The only Tier-0 center is located directly at

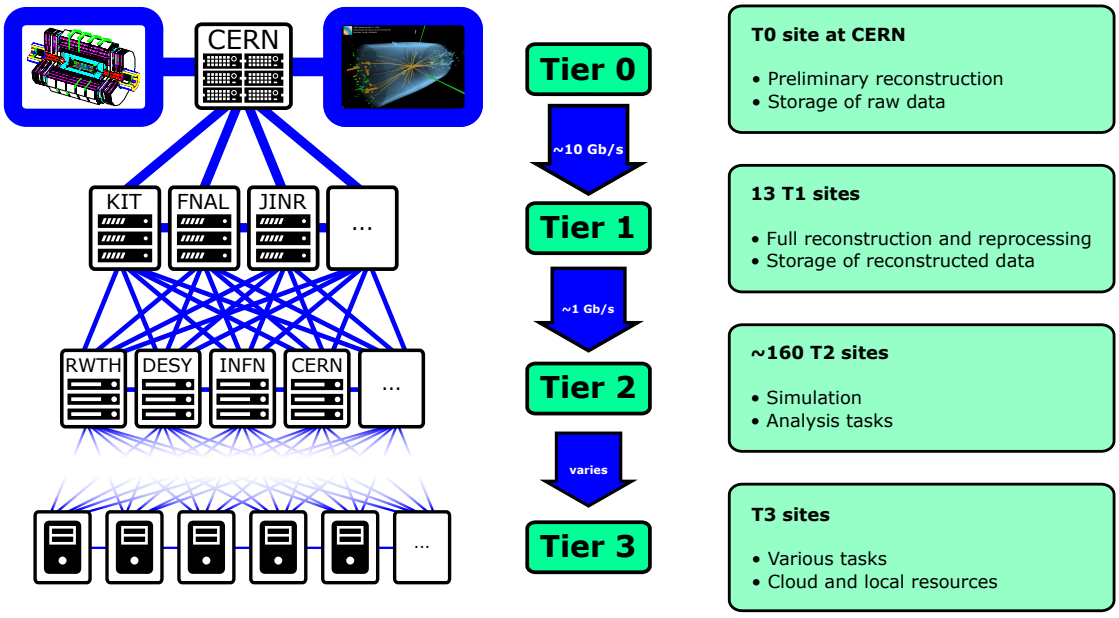


Figure 3.10.: Illustration of the computing infrastructure of the CMS experiment, which is part of the Worldwide LHC Computing Grid (WLCG). The WLCG is organized into different tiers. A single Tier-0 center located at CERN stores the raw data and performs a preliminary reconstruction. The data is then forwarded to several Tier-1 centers for full reconstruction. The grid further spreads out into many Tier-2 sites, capable of providing resources for simulations and analysis workflows. In addition, an inofficial Tier 3 can provide opportunistic resources, such as cloud services.

CERN. Here, the raw data from the experiments is directly stored on tape. The Tier-0 center is linked with several Tier-1 centers, large computing centers in different countries, with the Large Hadron Collider Optical Private Network (LHCOPN), providing a bandwidth of at least 10 Gb/s [150]. Main tasks of the Tier-1 centers are hosting copies of raw data, full reconstruction and fast data-serving for the next layer in the grid. Tier-2 centers are typically hosted at research institutes and universities, providing a large fraction of computation power, but usually do not have large storage capabilities compared to the Tier-1 centers. Thus, these centers carry out most of the simulation and processing needed for the final physics analyses. In addition, opportunistic resources, such as cloud services, can be added to the WLCG to increase the available computation power for flexible tasks. This so-called Tier 3 is not directly a part of the WLCG, but resources can be added at will.

4. Event Simulation and Reconstruction

To make a comparison between the recorded events from the experiments and the predictions from theory, an adequate simulation has to be provided, which is able to reproduce the outcome of the proton-proton collisions to the same level, in this case the electrical signals in the detector components of the CMS experiment. This is achieved by employing sophisticated simulations based on random numbers to sample probability density functions. These simulated events are then on an equal level as the recored data. Both of them then undergo the same processing sequence, in which the detector signals are combined to reconstruct basic elementary particles, such as electrons and muons. This chapter first gives an overview of the event simulation process, before the reconstruction of artificial and measured events is explained.

4.1. Simulation

For an analysis in high energy physics, one usually compares the observation from the experiment with the prediction obtained from a simulation. Collisions of elementary particles obey the rules of quantum mechanics, which has a non-deterministic nature. It is therefore not possible to predict the outcome of a single collision of particles. However, the properties of a larger set of collisions can be accurately described using the Monte Carlo (MC) method [151, 152], where single events are simulated by sampling the probability density functions from theory.
The simulation of events from the initial scattering to the electrical signals in the detector components is a complex task, as depicted in Fig. 4.1, which cannot be performed by a single computer program alone. Instead, a factorization approach is used, where each different step is simulated by a specialized program to achieve the highest accuracy.

4.1.1. Hard Scattering

The simulation of the hard scattering is of special importance as this involves the properties of the process one is interested in. The four-vectors of the interacting partons in the initial state are determined with the evaluation of a PDF for each parton. The PDFs are not known analytically, instead they are measured at a given four-momentum scale and can then be calculated for the requested energy (see Section 1.3). Different PDFs are available from various collaborations. The most common PDF sets are ABM [154], CTEQ [155], HERA [156], MSTW [157] and NNPDF [158–160]. The latter set has been extensively used in this thesis. The NNPDF set for two different scales is shown in Fig. 4.2.
To calculate the cross section of a process, all existing Feynman diagrams have to be considered to build up the matrix element. This is not possible as additional Feynman diagrams can always be added to an already existing subset by introducing radiation of gluons. As all these supplementary diagrams require at least one additional strong vertex, the probability compared to a

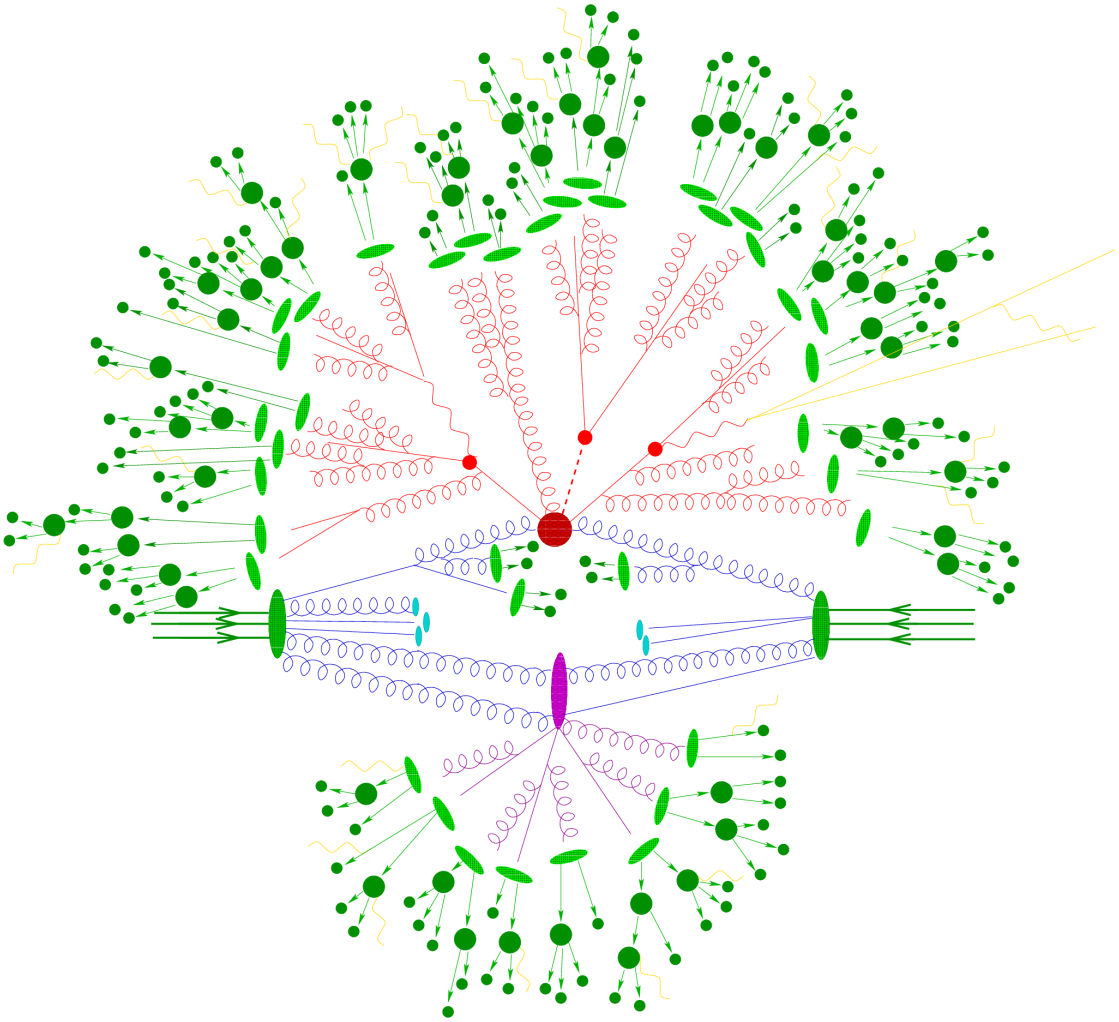


Figure 4.1.: Different simulation steps of proton-proton collisions. The hard scattering of initial partons (blue) occurs in the red circle in the middle, while a secondary interaction is happening at the purple shape. The outgoing partons (red and purple) undergo radiation of secondary particles, called parton shower, until their energy is low enough to form color-neutral hadrons. These hadrons can decay further, increasing the number of particles even more [153].

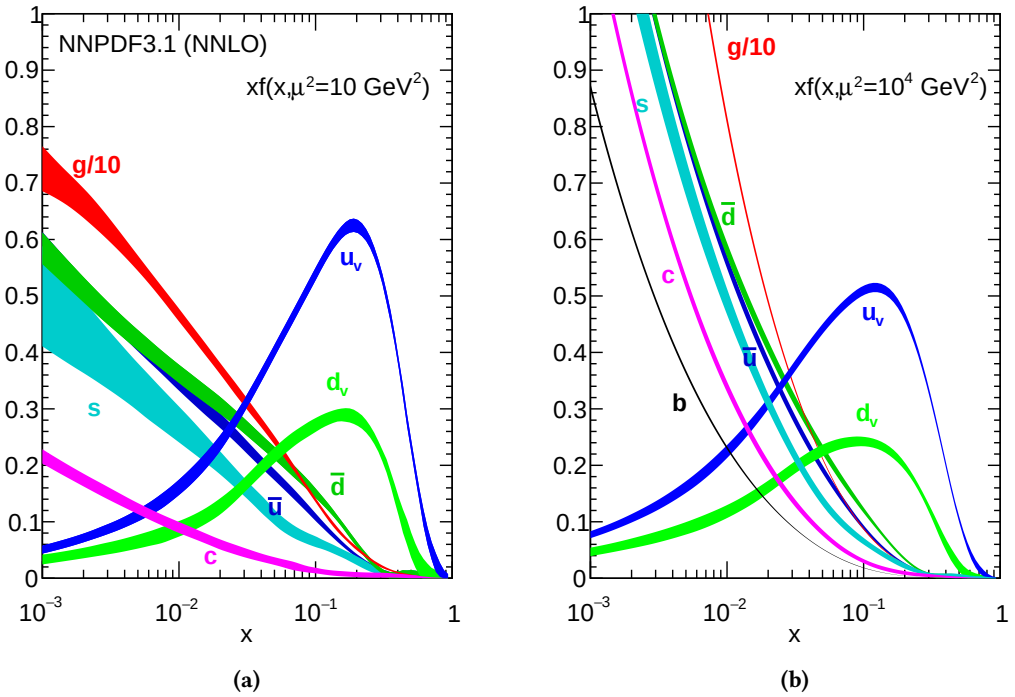


Figure 4.2.: The NNPDF parton distribution functions, which has been extensively used for the simulation of events in this thesis. The probability to find a certain valence quark, sea quark or gluon with a given fraction of the proton momentum x is shown for a momentum transfer of 10 GeV^2 (a) and 10^4 GeV^2 (b) [160].

simpler diagram is proportional to α_s. In case of high momentum transfer, α_s is small enough to treat this problem perturbatively and the resulting cross section can be expressed in a series of α_s:

$$\hat{\sigma} = \sigma_{LO}\left(1 + \sigma_1\left(\frac{\alpha_s}{2\pi}\right) + \sigma_2\left(\frac{\alpha_s}{2\pi}\right)^2 + \dots\right), \tag{4.1}$$

where σ_{LO} refers to the leading order (LO) cross section of the process, i. e., including only diagrams with the least number of strong vertices. The so-called next-to-leading order (NLO) cross section is obtained by including the next term in the series that is proportional to α_s. Corrections from NLO diagrams include real emissions, but also virtual corrections through emission and reabsorption (loop). This concept can be extended by including all diagrams with two additional strong vertices, resulting in the next-to-next-to-leading order (NNLO), and so forth. The same procedure can be also applied to electromagnetic corrections based on the fine-structure constant α_{em}, but for simulations for a hadron collider this is only a minor effect due to $\alpha_s \gg \alpha_{em}$. The computation of higher order diagrams increases significantly the simulation time, but can also introduce divergencies in the calculation and is thus not always straight forward.

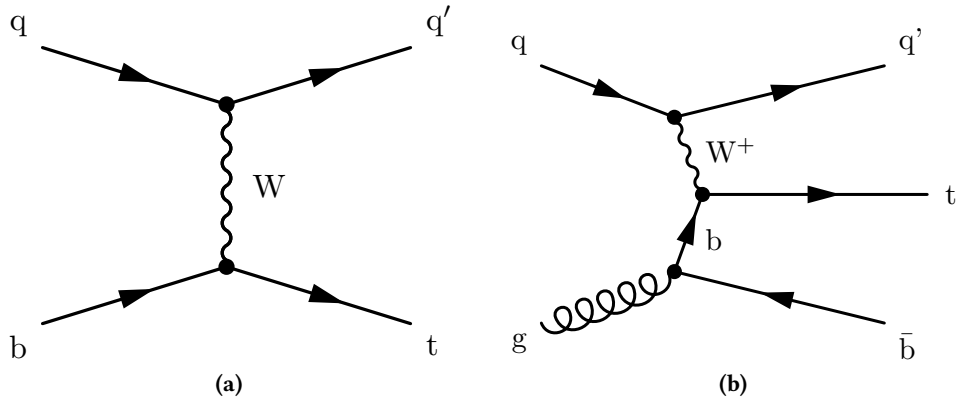


Figure 4.3.: Leading order Feynman diagrams for the single top quark production in the t channel for the five-flavor scheme (5FS) and four-flavor scheme (4FS). In the 5FS, a bottom quark is assumed to be present in the proton as a sea quark (a), while in the 4FS, the bottom quark must be created first from a gluon splitting (b).

4.1.2. Parton Shower

In the parton shower step, additional radiation of gluons with lower energy compared to the hard process is added to the simulation. These emissions can be added to every leg of the diagram and depending on whether it occurred in the initial or finale state it is called initial-state radiation (ISR) or final-state radiation (FSR). Subsequent emission of radiation is possible, resulting in many additional quarks and gluons, each with lower energy and increasing α_s. At higher energies, the shower can be computed with the matrix element, but not for any order. The evolution of the parton shower at lower energies is described with the Altarelli-Parisi splitting functions [26] and Sudakov form factors [161, 162]. To improve the accuracy of the parton shower, multiple matrix elements can be calculated, each with a different number of additional emissions. The matrix elements are then merged together and matched with the parton shower. Various merging and matching techniques are available, such as CKKW [163], MLM [164] and FxFx [165].

4.1.3. Processes with Bottom Quarks in the Initial State

A special case arises in the simulation of processes with bottom quarks in the initial state, such as the production of single top quarks, as the mass of the bottom quark exceeds the mass of the proton. The process can either be described using the so-called five-flavor scheme (5FS) or the four-flavor scheme (4FS) [166]. The difference of both schemes is shown exemplary for the single top t-channel process in Fig. 4.3. In the 5FS, the bottom quark is assumed to be massless and a constituent of the proton. Thus, a process like the single top t channel requiring one bottom quark in the initial state can be described as $2 \rightarrow 2$ process 4.3(a). The presence of a bottom quark as sea quark requires a second bottom quark to be present in the proton, which is done by performing a backward evolution in the parton shower. In the 4FS, the bottom quark is not considered a constituent of the proton and the corresponding PDF is set to zero. Instead, bottom quarks are created in pairs from a gluon splitting with high momentum, which replaces

the bottom quark in the initial state. In case of the exemplary single top t-channel production, this results in a $2 \rightarrow 3$ process 4.3(b). Here, the second bottom quark is described directly in the matrix element. This improves the accuracy of kinematic properties, but also makes the calculation harder due to the additional final-state parton.

4.1.4. Hadronization

With decreasing energy of each parton, the color-confinement of QCD forces the partons to form color-neural bound states. Since there is no exact solution known to the theory of QCD, which describes the process of hadronization, phenomenological models are used. One of the most popular approaches is the Lund string model [167]. In this model, the gluons are described with field lines, resulting in tubes of field lines between color-charged objects due to the self-interaction of gluons. If the energy of the string becomes too high, new quark-antiquark pairs are created in between until no isolated color charge remains. In contrast, the cluster hadronization model [168] strictly forces all gluons to split into quark-antiquark pairs. Hadrons are then formed by following the color-lines of partons to find suitable matches for a color-neutral state. In addition to the hadronization of partons, the decay of unstable hadrons is also simulated to achieve the same outcome as for real events.

4.1.5. Underlying Event and Pileup

So far, only the two interacting partons have been considered in the simulation, but in reality the situation at a hadron-hadron collider is more complicated. In addition, the remnants of the two protons, which do not take part in the hard scattering, also need to be considered as they are, in general, color-charged due to the missing partons. This so-called underlying event introduces additional activity in the event in terms of hadronization of the remnants.
As the probability for a single proton-proton collision is rather low, the experiment uses large numbers of protons in bunches at a high rate to increase the overall probability. However, this gives rise to the problem of pileup, multiple proton-proton interactions. A distinction is made between in-time and out-of-time pileup. Multiple interactions from the same bunch crossing, i. e., at the same time, are called in-time pileup and increase with rising instantaneous luminosity. The latter one, out-of-time pileup, is attributed to the response time of the detector, since a new collision happens every 25 ns. Both effects need to be applied to the simulation as well.

4.1.6. Monte Carlo Event Generators

The steps described above are implemented in different programs, where the whole chain of simulation is either provided by a single program or by interfacing two of them. Usually, the matrix element is generated with a dedicated program to achieve NLO precision and the output is directed to another program providing the parton shower and hadronization simulation steps. The programs are interfaced through a common data format, the Les Houches event (LHE) file format [169, 170]. These programs are called MC event generators and are capable of providing accurate predictions of a given process. A more extensive overview of MC event generators and the principles of the methods employed are provided in Ref. [171].

MadGraph5_aMC@NLO

The MadGraph5_aMC@NLO package is one of the two main matrix element and event generators used in this thesis [172]. It is the successor of MC@NLO [173] and MadGraph [174], which has been extensively used during the first years of the LHC, but with the drawback of providing only LO calculations. On the other hand, MadGraph5_aMC@NLO is able to provide generated events at LO and NLO precision. The generation of events is rather straightforward, as the user only has to provide a model with the parameters and define the involved particles. For the predictions at NLO, a possible double counting can occur as NLO contributions stemming from real emissions are identical to a hard emission from the parton shower. To avoid this problem, a dedicated matching of the matrix element calculations and the parton shower has to be performed. This is achieved through the MC@NLO method, where the overlapping parton shower contribution is subtracted. The method introduces negative weights to a fraction of events to reproduce the correct NLO predictions. Depending of the specific process, this can significantly reduce the effective number of simulated events and can lead to unphysical predictions, where distributions of observables contain negative entries.

POWHEG

The second main event generator is POWHEG, an abbreviation for positive weight hardest emission generator [175–177]. As the name implies, the hardest emission is already calculated in the matrix element, providing also simulations with NLO accuracy. This method requires interfacing POWHEG with a parton shower simulation that provides p_{T}-ordered showers. The principle of the hardest emission in the matrix element ensures that all simulated events can be used for an analysis. However, the drawback is that only predefined processes can be simulated and the implementation of new processes is not straightforward.

PYTHIA

PYTHIA [178–180] is a multipurpose event generator, providing complete simulation from matrix elements to hadronization, including also the underlying event and pileup. Since the matrix elements can only be calculated in LO, PYTHIA is mostly interfaced with NLO matrix element generators and takes care of the simulation from the parton shower onwards. The parton shower simulated by PYTHIA is p_{T}-ordered and the hadronization process utilizes the Lund string model. The underlying event is modeled with different tunes, i. e., different sets of parameters affecting the modeling of the parton shower and the hadronization, such as CUETP8M1 [181, 182], CUETP8M2T4 and CP5 [183].

Herwig++

Another multipurpose generator is Herwig++ [184], which is also mainly used in combination with other matrix element generators. The main differences to PYTHIA are an angular-ordered parton shower and the use of a cluster fragmentation model for the hadronization. The EE5C tune [185] is commonly used to model the underlying event. The Herwig++ generator is not

used in this thesis extensively and is only used to estimate an additional systematic uncertainty of MC simulations.

4.1.7. Detector Simulation

The last part of the whole chain of simulation programs is an accurate description of the detector response for generated events. The whole CMS detector is simulated with the GEANT4 package [186, 187], which describes the interaction of stable and hadronized particles with the detector material and the readout electronics. This lifts the simulated events on the same level as the measured data.

4.2. Reconstruction with the Particle-Flow Algorithm

Regardless of whether the events under investigation were simulated or actually recorded by the experiment, the signals from the detector response need to be reconstructed into physical or high-level objects. The core of the event reconstruction in the CMS experiment is the particle-flow (PF) algorithm [188–190], which combines the information of all subdetectors to reconstruct all stable particles in an event. The PF approach has been developed and used successfully by the ALEPH experiment at the LEP collider [191], but it is the first time this approach is used at hadron colliders. In order to make use of this technique, each detector component requires a fine granularity to resolve individual particles. Employing this approach significantly enhances the overall performance of the reconstructions and provides a powerful way for pileup mitigation. The algorithm starts by building *PF Elements*, consisting of reconstructed trajectories in the tracker and clustered energy deposits in the ECAL and HCAL. All these elements are then tested for a possible link, e. g., a track, that can be extrapolated to an energy deposit in the ECAL, a typical signature of an electron. Such combinations are called *PF Blocks*. The identification and reconstruction sequence is then applied to the list of blocks, starting with the easiest and most striking signatures: the muons. Afterwards, the corresponding blocks are removed from the list and the remaining blocks are delivered to the electron and photon reconstruction algorithm. After the remaining blocks have been processed and reconstructed into neutral and charged hadrons, a post-processing step is applied. Afterwards, high-level objects, such as jets, can be built. The power of the PF algorithm is illustrated in Fig. 4.4. In the following sections, each specific reconstruction algorithm is explained in more detail.

4.2.1. Tracks

Before the first elements of the PF algorithm can be built, the hits of charged particles in the various layers of the silicon pixel and strip detectors have to be combined to trajectories, called particle tracks. This needs to take into account possible misalignments of the individual tracker modules to achieve a resolution in the order of a few μm [192]. The tracks are reconstructed from the hits in the tracker subsystem using a Combinatorial Track Finder (CTF) [127] algorithm, which is based on Kalman filters [193–196]. The CTF algorithm utilizes an iterative approach, where tracks with high transverse momentum ($p_T > 0.8$ GeV) and close to the interaction region are reconstructed first. The corresponding hits are then removed from the event to reduce the

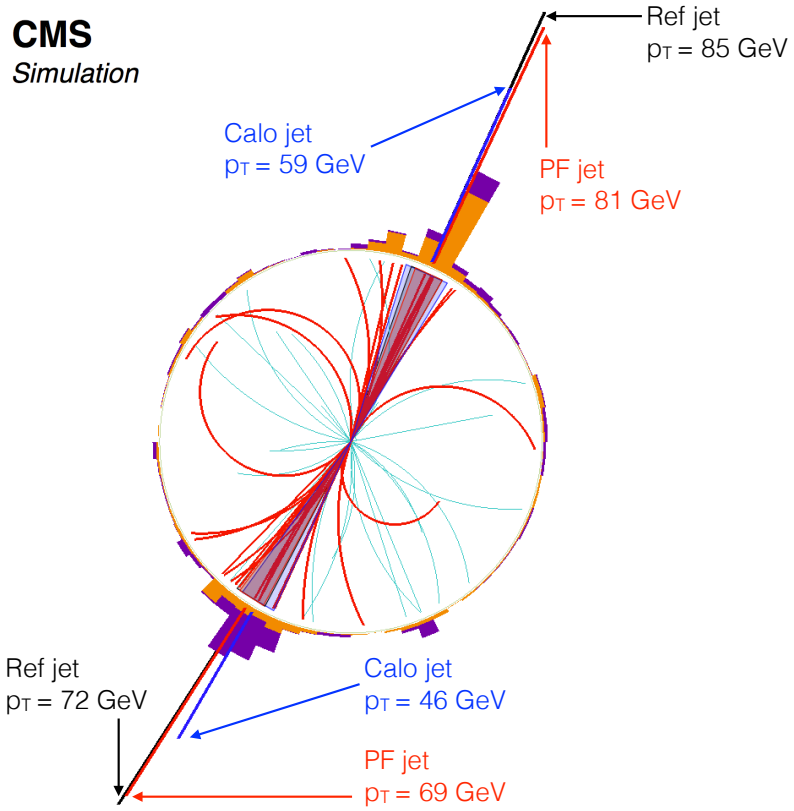


Figure 4.4.: Illustration of the improvement of jet reconstruction with the particle-flow (PF) algorithm with respect to the standard reconstruction using only calorimeter information. A transverse view of the CMS detector is provided for a simulated dijet event, showing the reconstructed tracks and energy deposits in the calorimeters. While the reconstructed p_T of both jets based only on calorimeter information (Calo jet) significantly differs from the reference values of the simulation (Ref jet), the PF algorithm provides an accurate description by also including the information of the tracker system (PF jet) [190].

combinatorics of subsequent iterations. Each iteration consists of four basic steps: First, a track seed is generated from only two or three hits. In the second step, Kalman filters are used to extrapolate the track candidates to additional layers and suitable hits are added to the candidate. A more sophisticated Kalman filter algorithm is then used in the third step to determine the best possible track for the given hits. In the last step, all tracks that do not fulfill certain quality criteria are discarded.

4.2.2. Vertices

With all tracks reconstructed, the next step is to find all vertices in the event, i. e., the position of tracks with the same origin. Of special interest is the determination of the primary vertex (PV) from the interaction point of the hard scattering. Nevertheless, all other vertices at the

beamline are also important as they indicate pileup contributions from other proton-proton interactions in the same bunch crossing. Only high-quality tracks are selected as input for the vertex reconstruction. The selected tracks are then clustered according to their distance to the beam spot in z direction. To be able to resolve as many vertices as possible, but not split a valid vertex into multiple track clusters, a deterministic annealing (DA) algorithm [197] is used to determine the best possible solution. Vertex candidates with at least two associated tracks to them are then fitted using an adaptive vertex fitting (AVF) algorithm [198–200] to determine the position of the vertex and assign a weight to each of the associated tracks, which corresponds to a probability to be compatible with the vertex. The PV of an event is defined as the vertex with the highest sum of p_{T}^2 from the originating tracks associated to reconstructed physical objects, such as jets and charged leptons.
In-time pileup effects are mitigated by employing the charged hadron subtraction (CHS) technique [201], where PF objects are removed from the collection if they are assigned to a vertex other than the PV of the event.

4.2.3. Muons

Muons are the first particles that are reconstructed through the PF algorithm as their signature is unique and they can thus be easily separated from the remaining candidates. The reconstruction of muons utilizes information from the tracker (tracker track) and the muon system (standalone-muon track), where the tracks are first reconstructed independently in each subdetector [202, 203]. Depending on which information is used as a starting point, two main different reconstruction strategies exist. In the global muon reconstruction (outside-in), a matching tracker track is searched for a given standalone-muon track and the associated hits of both tracks are fitted simultaneously using a Kalman filter. This approach significantly enhances the resolution for high-p_{T} muons, i. e., $p_{\mathrm{T}} > 200\,\mathrm{GeV}$. The complementary approach is the tracker muon (inside-out), which considers all tracks with $p_{\mathrm{T}} > 0.5\,\mathrm{GeV}$ as a muon candidate. Possible tracker muons are found if a track can be matched to a short muon segment built up from hits in the DTs or CSCs, while considering all possible interactions along the path. The tracker muon reconstruction is especially efficient for muons with $p_{\mathrm{T}} < 5\,\mathrm{GeV}$ since it relies only on small segments in the muon system. The reconstruction based only on standalone-muon tracks is also possible, but due to the worse resolution it is not used for the analyses in this thesis.

4.2.4. Electrons

The next particles to be reconstructed are electrons [134, 204, 205]. Due to their lower mass compared to muons, electrons can already lose a significant amount of their energy inside the tracker in terms of bremsstrahlung. This results in a broader shower in ϕ direction in the ECAL and increases the difficulty of the reconstruction. Similar so the muon reconstruction, two different approaches are used. The first method starts with energy deposits in ECAL crystals. Neighboring energy deposits are clustered together around a seed, which is the crystal with the highest energy deposit, to a so-called supercluster (SC). The seed of the SC serves as starting point for matching the cluster to tracks in the tracker. In case of electrons the standard track reconstruction with the CTF algorithm utilizing Kalman filters is not adequate, as hits cannot be

associated to the track due to changes in curvature because of bremsstrahlung. Thus, tracks with bad quality from the original track reconstruction are refitted using a dedicated Gaussian sum filter (GSF) [206,207], where energy loss of electrons in each layer of the tracker is approximated with a Gaussian distribution. This approach works well for electrons with high transverse momentum. The complementary approach starts with Kalman filter and GSF tracks and tries to link them with ECAL clusters with the help of multivariate techniques, which results in better resolution for low-p_T electrons.

4.2.5. Photons, Charged Hadrons and Neutral Hadrons

After the reconstruction of electrons and muons and removing the corresponding elements from the collection in the event, the remaining elements are assigned to photons and hadrons. A special case are isolated photons, where the reconstruction is similar to electrons, but without any associated tracks [208]. The remaining particles are mostly charged hadrons ($\pi^{\pm}$, $K^{\pm}$, protons), neutral hadrons (K^0_L, neutrons) and nonprompt photons stemming from π^0 decays. Remaining clusters in the ECAL and HCAL which cannot be assigned to any tracks are associated to photons and neutral hadrons. For the region of the tracker ($|\eta| \leq 2.5$), all ECAL clusters are assigned to photons and HCAL clusters to neutral hadrons. In the forward region ($|\eta| > 2.5$), this strict distinction cannot be made anymore due to the absence of tracking information. Thus, clusters in the ECAL and HCAL which can be linked are associated to charged or neutral hadrons stemming from the same source of hadronization, while the ECAL clusters without such a link are associated to photons. The remaining HCAL clusters are linked to remaining tracks and ECAL clusters under different hypotheses.
If the measured energy exceeds the momentum of the associated tracks under consideration of the energy resolution, this is interpreted as a photon or neutral hadron: In case the excess is smaller than the ECAL energy, but larger than 500 MeV, a photon is reconstructed. If the recalibrated ECAL energy is then still not compatible and larger than 1 GeV, an additional neutral hadron is reconstructed.
If the measured momentum is compatible with the measured energy, no additional neutral particles are reconstructed. Instead, the track of the charged hadron is refitted using also the calorimeter information. This improves the resolution and allows a smooth transition between the tracker-dominated regime at low energies and the calorimeter-dominated regime at higher energies.
When the measured energy is significantly lower than the momentum of the associated tracks (at least three standard deviations), all already reconstructed muons with bad quality are inspected for possible ECAL deposits. Any further disagreement is contributed to misidentified tracks, which are then sorted in decreasing order according to their uncertainty in the transverse momentum. The tracks are then sequentially removed until either the PF block has no more tracks or the discrepancy between energy and momentum is resolved.

4.2.6. Jets

After the reconstruction of basic particles, additional high-level objects can be defined. They may have no direct physical meaning themselves, but are more abstract and are closer to the

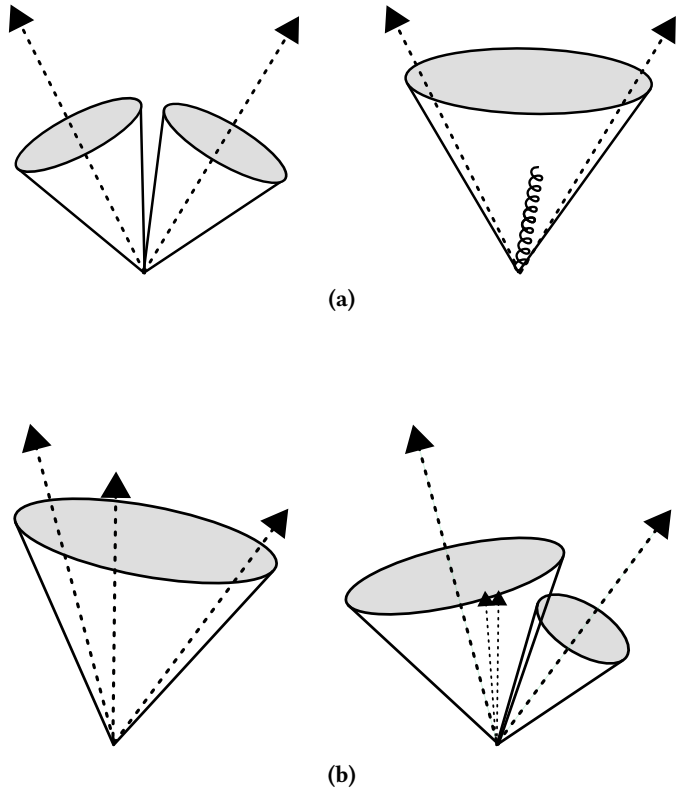


Figure 4.5.: Illustration of the concept of infrared safety and collinear safety of jet clustering algorithms. A infrared-safe clustering algorithm is robust against soft-gluon radiation in such a way that it does not change the outcome of the clustering (a). Similar, the algorithm should be robust against collinear splitting of particles (b).

experimental situation. One of these objects types is a jet, a collection of collimated hadrons stemming from the hadronization of partons. As a direct comparison between simulation and measurement at the level of partons is not possible at hadron colliders, the conclusions on partons are drawn on these artificial objects.
The clustering of hadrons into jets is mathematically defined and follows specific algorithms. In principle, there are two different kinds of jet algorithms: cone-based algorithms, which cluster all objects within a given radius, and sequential clustering algorithm, combining objects subsequent according to some criteria. All those methods need to fulfill two properties to be reliable for an environment such as the LHC: First, the clustered jet should not depend on the distribution of energy among a collinear splitting of hadrons (collinear-safe). The second condition is that the clustering is robust against additional soft emission of gluons (infrared safe). Both properties are illustrated in Fig. 4.5. The main jet clustering algorithm used in CMS is the anti-k_T algorithm [209], which is both collinear- and infrared-safe. Different objects are clustered according to the generic distance metric:

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

where p_T is the transverse momentum of object i or j, $\Delta_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}$ the difference between the objects in the y-ϕ plane and R a fixed-size parameter. The value of n determines the actual clustering algorithm. In case of the anti-k_T algorithm, a value of $n = -1$ is used. Values of $n = 1$ and $n = 0$ correspond to the k_T algorithm [210] and the Cambridge/Aachen algorithm [211], respectively. The distance of an object to the beam is given by

$$d_{B,i} = p_{T,i}^{2n}. \tag{4.3}$$

The clustering is done in iterative steps. First, all distances between objects themselves $d_{i,j}$ and between the objects and the beam $d_{B,i}$ are calculated. If the smallest of all distances is found between object i and j, they are reclustered into a new object k. On the other hand, if the smallest distance is found between an object and the beam, this specific object is removed from the list and considered a jet. This procedure is repeated until all objects have been clustered. The parameter R determines the size of the clustered jet in the y-ϕ plane. For the anti-k_T jets used at the CMS experiment, a distance parameter of $R = 0.4$ is used at $\sqrt{s} = 13\,\text{TeV}$, while at 7 and 8 TeV a value of $R = 0.5$ was used. Although the anti-k_T algorithm is not a cone-based clustering algorithm, the resulting jets are of circular shape. A comparison of the three different jet clustering algorithms is provided in Fig. 4.6. The clustering of jets can be very computing intensive for hadron-hadron collisions, as the distances between all objects have to be calculated in every iteration step. The complexity of the clustering can be reduced utilizing a nearest neighbor approach [212], implemented in the FastJet package [213].

Jet Energy Corrections

Since clustered jets are reconstructed from hadrons in the detector, they are vulnerable to effects stemming from pileup in the event. Furthermore, the CMS detector is not completely homogeneous, resulting in different response for the same particles in different regions of the detector. To counteract these circumstances, dedicated jet energy corrections are applied to improve the performance of jet measurements. This is realized in the CMS Collaboration with a factorized approach [214–216], where each level of corrections represents a different effect. The corrections are applied sequentially and in a fixed order. An overview of all corrections is provided in Fig. 4.7. The different jet energy corrections levels are:

- **L1 pileup:** The first correction applied is aimed at removing contributions from pileup interactions to the measured jet energy. This is achieved by comparing the simulations of QCD dijet events with and without additional pileup. The pileup contribution can be parameterized with an offset energy density [217]. Any residual difference is resolved with η-dependent corrections using a random-cone (RC) method in zero-bias events, i. e., events without any potential trigger bias.

- **L2L3 MC truth:** The difference between the energy of the reconstructed jet and the parton energy is resolved by comparing the response in a simulated QCD MC sample. Corrections are applied to yield a uniform response in different regions of pseudorapidity (L2Relative) and transverse momentum (L3Absolute).

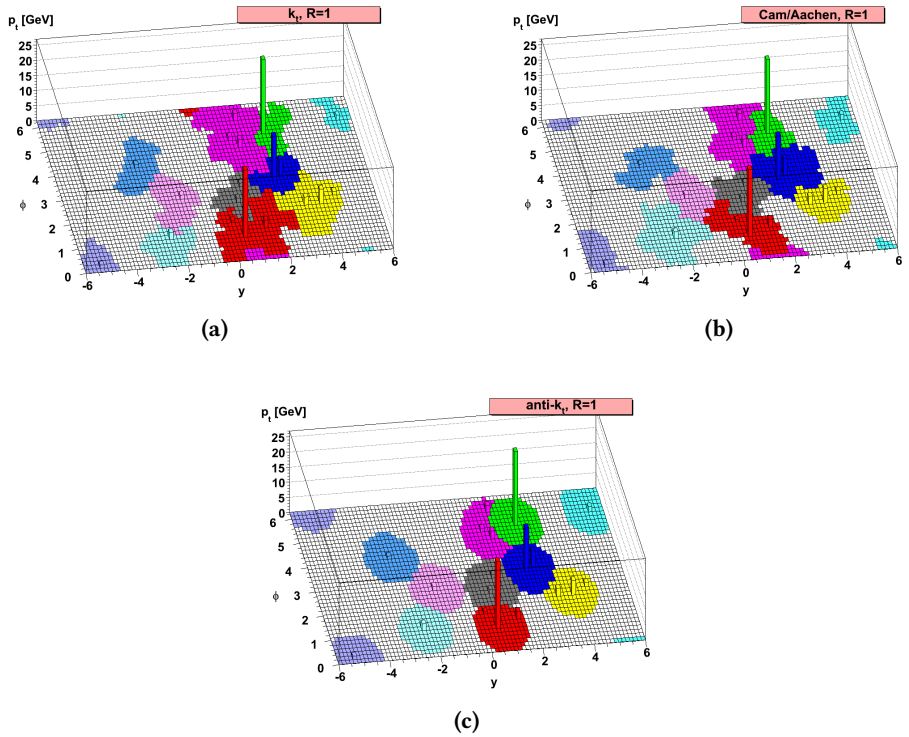


Figure 4.6.: Example of different jet clustering algorithms for the same simulated event. The k_T (a), Cambridge/Aachen (b) and anti-k_T algorithms (c) are shown, all with the same distance parameter of $R = 1$. While the jets obtained from the k_T and Cambridge/Aachen algorithm strongly depend on the noise in the detector, the anti-k_T algorithm provides robust, cone-shaped jets [209].

- **L2L3 residuals:** Any remaining difference in the response for jets in data and MC simulations are corrected with residual corrections. Deviations in the pseudorapidity are resolved using dijet events with similar p_T, where one of the two jets serves as reference in the barrel region (L2 residuals). For the absolute scale of the jet, i. e., response in p_T, corrections are derived from measurements of Z+jets and photon+jets events, where the additional gauge boson can be precisely reconstructed (L3Residuals).
- **L5 flavor:** Optional corrections can be applied due to differences in flavor response, but since they are not used in this thesis, they are only listed for completeness.

The non-continuous numbering follows historical reasons and refers to the jet energy corrections made in Run I for a center-of-mass energy of 7 and 8 TeV.

b Tagging

The processes studied in this thesis all involve top quarks. The top quark does not form any bound states in form of hadrons and decays into a bottom quark and a W boson in almost all cases. Thus, the decay products of the top quark are of special importance in identifying events

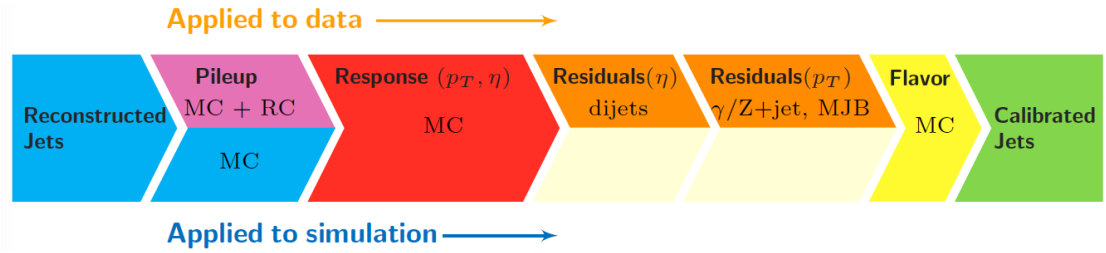


Figure 4.7.: Illustration of the jet energy corrections in the CMS experiment. The reconstructed jets undergo a pileup subtraction based on MC simulations, with residual corrections for jets in data based on a random-cone (RC) method. The response of the reconstructed jets is then corrected by comparing the energy to the corresponding parton in a simulation, depending on the pseudorapidity and transverse momentum. Remaining differences between data and simulation are corrected in data by measurements in dedicated events. Additional flavor-dependent corrections can be applied [216].

in which top quarks were produced. Jets stemming from bottom quarks feature a property which can be exploited to distinguish them from jets stemming from gluons or quarks of a different flavor. This property is the presence of a secondary vertex in the event at which the already hadronized bottom quark (B meson) is decaying into a charm or up quark. The secondary vertex can be distinguished from vertices along the beam line because of the long lifetime of B mesons, as the bottom quark cannot decay into a top quark and the off-diagonal CKM matrix elements V_{cb} and V_{ub} are small. Identifying jets from bottom quarks utilizing this information is known as b tagging [218, 219] and the corresponding algorithms are called b taggers. An illustration of this technique is provided in Fig. 4.8. The b taggers used in the CMS Collaboration utilize information about tracks and secondary vertices to determine a probability that a jet originated from a bottom quark.
Different variants are available, but the most common at $\sqrt{s} = 13$ TeV is the combined secondary vertex (CSVv2) algorithm. This algorithm is an improved version of the original CSV algorithm heavily used during Run I, which utilizes multivariate techniques to combine information about displaced tracks and secondary vertices. Tracks compatible with a K^0_S meson decay or with angular separation are rejected. Jets without a secondary vertex and selected tracks receive a default value for the classifier of -1. Three different neural networks are used with a different number of input variables depending on the number of secondary vertices found in the jet. These categories are:

1. **RecoVertex:** At least one secondary vertex has been found and the full set of input variables is used.
2. **PseudoVertex:** No secondary vertex has been found, but at least two tracks have a large impact parameter significance and the combined mass exceeds the mass of the K^0_S meson by at least 50 MeV.
3. **NoVertex:** Jets that do not qualify for the categories above, thus only track information is used.

All three neural networks are combined in a likelihood ratio. In addition, two independent training sets are performed, one with charm quarks as background and one with light-flavored

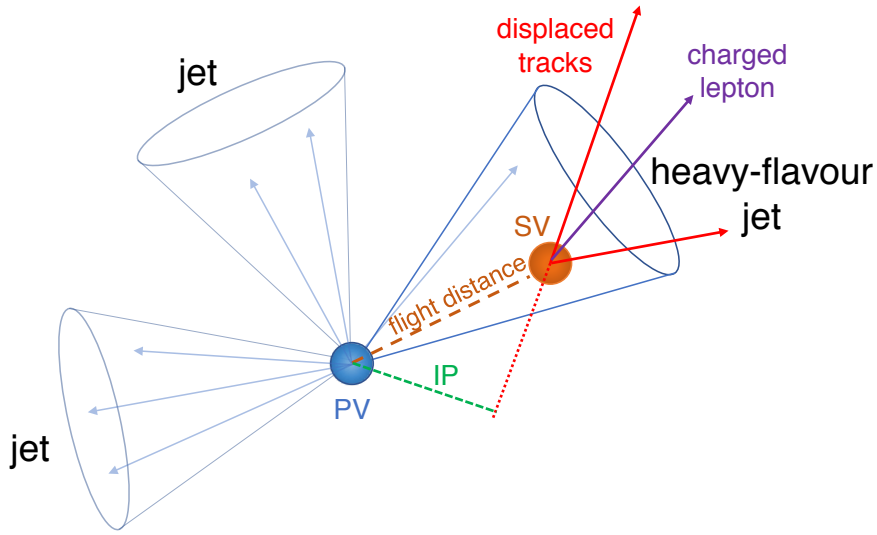
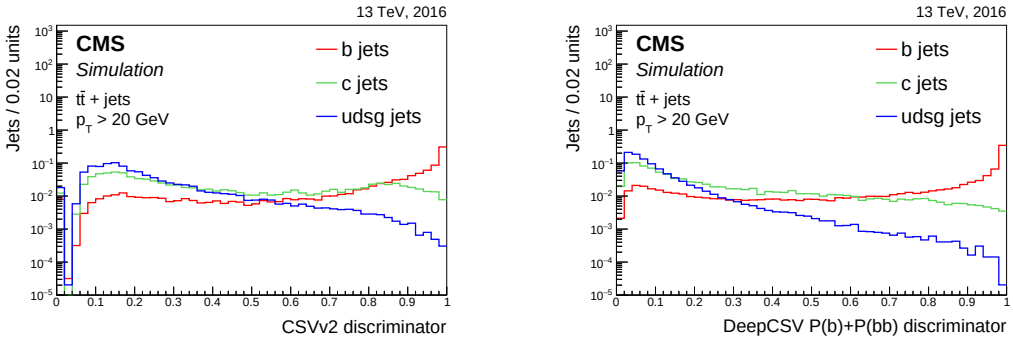


Figure 4.8.: An illustration of the principle of b tagging. The creation of a bottom quark in the hard scattering results in the fragmentation into a B meson, which has an unusual long lifetime. This results in a secondary vertex (SV), at which the B meson decays, with a flight distance away from the primary vertex and an impact parameter (IP). The displaced tracks stemming from the secondary vertex can be used to reconstruct the vertex and the b tagging algorithm provides a probability for the corresponding jet to originate from a bottom quark [219].

jets as background. The final classifier is then a linear combination of both with a relative weight of 3 (1) for light (charm) quarks, motivated by the flavor composition of top quark pair production. The main improvements with respect to the earlier version of the CSV algorithm are exploiting multivariate techniques, an extended set of input variables (19 in total for the RecoVertex category, compared to six in Run I) and consideration of the overall jet kinematics. Due to the recent improvements in the area of machine learning [220], the CSVv2 algorithm has been further improved during Run II into the so-called DeepCSV algorithm employing a DNN. The main improvements of the DeepCSV algorithm are an increased number of nodes and layers in the network and a simultaneous training in all vertex categories. In addition, the DNN has different output neurons (multiclassification) which correspond to the probability of jets in different flavor categories:

1. ***P*(b):** Exactly one b hadron inside the jet
2. ***P*(bb):** At least two b hadrons inside the jet
3. ***P*(c):** Exactly one c hadron and no b hadron inside the jet
4. ***P*(cc):** At least two c hadrons and no b hadron inside the jet
5. ***P*(udsg):** Jets from light quarks or gluons

With this categorization a more sophisticated separation is possible compared to a single discriminator. To provide the same result as the CSVv2 algorithm, the values of $P(\mathrm{b})$ and $P(\mathrm{bb})$



(a) **(b)**

Figure 4.9.: The distribution of both b tagging algorithms utilized in this thesis, the combined secondary vertex (CSV) algorithm (a) and DeepCSV algorithm (b). The response of the classifier is shown for true b jets (red), c jets (green) and light-flavored jets (blue), measured in data taken in 2016 at a center-of-mass energy of 13 TeV [219].

are added together. The discriminator of the CSVv2 and DeepCSV (P(b) + P(bb)) b tagging algorithms are shown in Fig 4.9(a) and Fig 4.9(b), respectively.
Although b tagging is a reliable way to identify jets stemming from bottom quarks, the aforementioned algorithms are not perfect. Like for any non-perfect classification there is always the trade-off between signal efficiency, here b tagging efficiency, and the false-positive rate, which corresponds in this case to the misidentification (mistag) rate, i. e., jets wrongly identified as jets from bottom quarks. Thus, the CMS Collaboration uses different working points for each b tagging algorithm. These working points are called loose, medium and tight, roughly corresponding to a mistag rate of 10%, 1% and 0.1%, respectively. A comparison of different b tagging algorithms in terms of b tagging efficiency and misidentification probability is provided in Fig. 4.10.

4.2.7. Missing Transverse Momentum

Another important quantity of an event is the missing transverse momentum [221]. The CMS detector is built in a way to measure or even absorb all interacting particles within the detector volume. As the interacting protons only have momentum in longitudinal direction, but not in the transverse plane, conservation of momentum predicts that the vector sum of all particles transverse momenta created in the collision should add up to zero. Any occurring difference can have two reasons. First, the CMS detector does not cover the entire solid angle of 4π and the reconstruction of particles is smeared by the detector resolution. However, these detector and reconstruction effects give only minor contributions. The second reason for missing transverse momentum are particles that escape the detection of the experiment. In the SM, this is the neutrino. For instance, the neutrino from a leptonic W boson decay, such as happening in the decay of the top quark, can carry a significant amount of momentum away from detection. Furthermore, many dark matter candidates are predicted to only interact weakly, thus escaping the

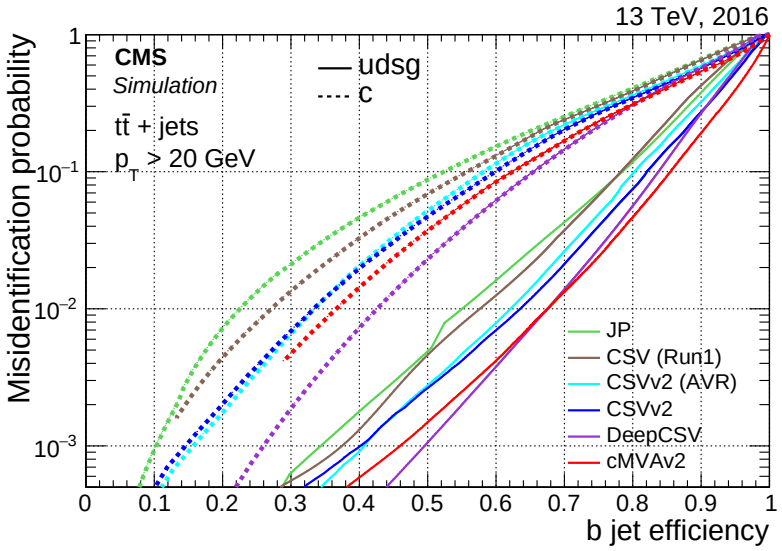


Figure 4.10.: Comparison of the performance from different b tagging alorithms, showing the b jet efficiency over the misidentification probability for c jets and light-flavored jets. Significant improvements are observable when comparing the CSV-based algorithm of Run II to earlier versions at Run I or simpler jet probability (JP) algorithms [219].

detector and leaving high missing transverse momentum. The missing transverse momentum is defined as

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} = \vec{p}_{\mathrm{T}}^{\,\mathrm{miss,raw}} + \vec{C}_{\mathrm{T}}^{\,0} + \vec{C}_{\mathrm{T}}^{\,\mathrm{I}} \tag{4.4}$$

$$= -\sum_{i=1}^{N_{\mathrm{PF}}} \vec{p}_{\mathrm{T},i} + \vec{C}_{\mathrm{T}}^{\,0} + \vec{C}_{\mathrm{T}}^{\,\mathrm{I}}, \tag{4.5}$$

where $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss,raw}}$ is the uncorrected missing transverse momentum from all reconstructed PF objects and $\vec{C}_{\mathrm{T}}^{\,0,\mathrm{I}}$ are the so-called Type-0 and Type-I corrections [222]. The Type-0 correction accounts for degrading in performance due to missing transverse momentum contributions from pileup interactions, while the Type-I correction propagates the jet energy corrections to the calculation of the missing transverse momentum.

5. Preselection of Physics Objects and Additional Reconstruction

Although the analyses described in this thesis use different data sets accumulated during the Run II of the LHC, all share some common quality requirements on the physics objects. These preselections ensure that all objects are originating from actual physics processes and are not due to problems or noise effects in the reconstruction. In addition, the reconstruction of the four-momentum of a leptonically decaying W boson is not straightforward due to the missing information of the longitudinal momentum of the neutrino. These topics are covered in the following sections.

5.1. Physics Objects

The physics objects from the PF reconstruction, described in the previous chapter, are the inputs for the analyses described in this thesis. To ensure a high quality of reconstructed objects, several requirements are needed to be fulfilled.

5.1.1. Primary Vertex

A primary vertex (PV) candidate must fulfill certain quality requirements to be considered in the analyses described in this thesis. The PV must be located within a cylinder of radius 2 cm and length 24 cm around the center of detector. In addition, the number of degrees of freedom of the PV must be greater than four.

5.1.2. Muons

Apart from the requirements of the PF algorithm, additional criteria are imposed for muon candidates following the recommendations by the Muon Physics Objects Group (MUO POG) of the CMS Collaboration [223]. These muon quality selections are condensed into an identification (ID) flag, which is true if the muon candidate fulfills all criteria and false otherwise [224]. The two ID selections for muons, called loose and tight ID, are summarized in Table 5.1. These requirements ensure the quality of the muon reconstruction. To reject events with nonprompt muons, e. g., muons originating from a semileptonic meson decay inside a jet, an isolation variable can be defined as

$$I_\mu = \frac{1}{p_{T,\mu}} \left(p_{T,CH} + \max \left(p_{T,NH} + p_{T,\gamma} - \Delta\beta \cdot p_{T,CH(PU)}, 0 \right) \right), \tag{5.1}$$

with the transverse momentum of the muon $p_{T,\mu}$, the transverse momentum of charged hadrons $p_{T,CH}$, neutral hadrons $p_{T,NH}$, photons $p_{T,\gamma}$ and charged hadrons from pileup $p_{T,CH(PU)}$ inside a

Table 5.1.: Properties of the muon needed to pass the criteria of the loose and tight muon ID. The loose ID only requires the PF muon to be either a tracker or global muon, whereas the tight ID only accepts global muons with well-fitted tracks.

	loose ID	tight ID
PF muon candidate	✓	✓
Global muon or tracker muon	✓	–
Global muon	–	✓
χ^2/n_{dof} from the muon track fit	–	< 10
Hits in the muon chamber	–	≥ 1
Number of muon stations with segments	–	≥ 2
Impact parameter d_{xy} with respect to the PV	–	$< 0.2\,\text{cm}$
Longitudinal distance d_z with respect to the PV	–	$< 0.5\,\text{cm}$
Number of hits in the pixel detector	–	≥ 1
Number of tracker layers with hits	–	≥ 6

cone of the radius 0.4 in the η-ϕ plane around the muon and an adjustable correction factor $\Delta\beta$. Since there is no way to distinguish if a neutral particle is created in the PV or stemming from pileup, the contribution of those particles is corrected. The value of $\Delta\beta$ is chosen to be 0.5 as this represents approximately the fraction of neutral to charged hadrons stemming from pileup interactions [225]. The isolation variable is a powerful way to suppress nonprompt muons from QCD multijet events. The efficiency of the muon ID and isolation measured in data from 2015 is shown in Fig. 5.1.

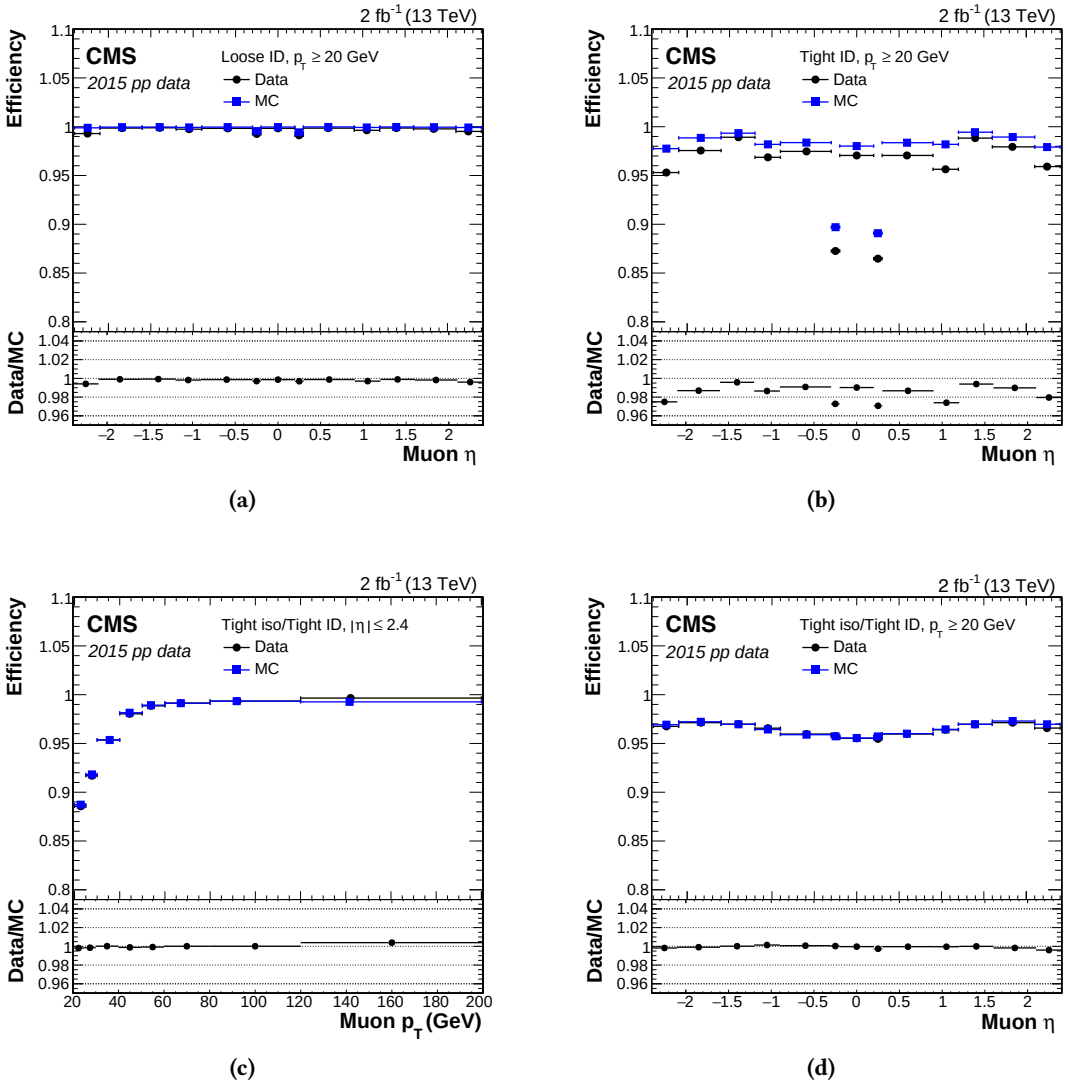


Figure 5.1.: The efficiency of muon ID and isolation requirements. The efficiency in simulation and data is shown for the loose (a) and tight muon ID (b) dependent on the pseudorapidity of the muon. Efficiencies for the combined selection of tight muon ID and tight muon isolation ($I_\mu < 0.15$) are shown dependent on the transverse momentum of the muon (c) and the pseudorapidity of the muon (d). Differences in efficiency of simulation and data are corrected in the respective analyses.

5.1.3. Electrons

Similar to muons, electron candidates have to fulfill several properties to be accepted by analyses described in this thesis. These criteria follow the recommendations from the E/gamma Physics Object Group of the CMS Collaboration [226]. These conditions are summarized in the electron ID flag, which correspond to an efficiency of about 95%, 90% and 75% for the veto, loose and tight ID, respectively [227]. To ensure constant efficiencies over different years of data, the requirements for the electron ID are re-evaluated whenever the conditions of the LHC or the experiment change. The criteria for the veto, loose and tight electron ID are summarized in Table 5.2 for the data taken in the years 2015, 2016 and 2017. The selection criteria on the impact parameter of the electron with respect to the primary vertex d_{xy} and d_z are part of the ID decision in 2015, but not in 2016 and 2017. Therefore, for data from these years, both selection criteria are applied in addition to the electron ID decision.
Compared to the muon isolation, the electron isolation is slightly differently defined:

$$I_{\mathrm{e}} = \frac{1}{p_{\mathrm{T,e}}} \left(p_{\mathrm{T,CH}} + \max \left(p_{\mathrm{T,NH}} + p_{\mathrm{T,\gamma}} - \rho \cdot A_{\mathrm{eff}}, 0 \right) \right), \tag{5.2}$$

with the average transverse momentum density ρ and the effective area A_{eff} to estimate the contribution from pileup interactions. The cone size for the electron is chosen to be only 0.3. Compared to the muon selection criteria, the electron isolation is already part of the electron ID and no additional selection is required.

Table 5.2.: The selection criteria for electrons to pass the corresponding ID in each year of the data taking. The individual values of each variable are different for the barrel and endcap region of the ECAL. The requirements on the impact parameters d_{xy} and d_z are only present in the ID for data taking in 2015 and are removed from the ID in 2016 onwards. The constants used for the 2017 electron ID are defined as $c_1 = \frac{1.12}{E_{sc}} + \frac{0.0368 \cdot \rho}{E_{sc}}$ and $c_2 = \frac{0.5}{E_{sc}} + \frac{0.201 \cdot \rho}{E_{sc}}$, where the E_{sc} is referring to the energy of the electron supercluster.

2015 data	$\lvert\eta_{sc}\rvert \leq 1.479$ veto ID	$\lvert\eta_{sc}\rvert > 1.479$ veto ID
Supercluster shower shape	< 0.0114	< 0.0352
$\lvert\Delta\eta(\text{sc, track})\rvert$	< 0.0152	< 0.0113
$\lvert\Delta\phi(\text{sc, track})\rvert$	< 0.216	< 0.237
Hadronic/EM energy	< 0.181	< 0.116
Electron isolation	< 0.126	< 0.144
$\left\lvert\frac{1}{E_{sc}} - \frac{1}{p_{track}}\right\rvert$ (GeV^{-1})	< 0.207	< 0.174
Expected missing hits	≤ 2	≤ 3
Passing conversion veto	✓	✓
$\lvert d_{xy}\rvert$ (cm)	< 0.0564	< 0.222
$\lvert d_z\rvert$ (cm)	< 0.472	0.921

2016 data	$\lvert\eta_{sc}\rvert \leq 1.479$		$\lvert\eta_{sc}\rvert > 1.479$	
	loose ID	tight ID	loose ID	tight ID
Supercluster shower shape	< 0.011	< 0.00998	< 0.0314	< 0.0292
$\lvert\Delta\eta(\text{sc, track})\rvert$	< 0.00477	< 0.00308	< 0.00868	< 0.00605
$\lvert\Delta\phi(\text{sc, track})\rvert$	< 0.222	< 0.0816	< 0.213	< 0.0394
Hadronic/EM energy	< 0.298	< 0.0414	< 0.101	< 0.0641
Electron isolation	< 0.0994	< 0.0588	< 0.107	< 0.0571
$\left\lvert\frac{1}{E_{sc}} - \frac{1}{p_{track}}\right\rvert$ (GeV^{-1})	< 0.241	< 0.0129	< 0.14	< 0.0129
Expected missing hits	≤ 1	≤ 1	≤ 1	≤ 1
Passing conversion veto	✓	✓	✓	✓

2017 data	$\lvert\eta_{sc}\rvert \leq 1.479$		$\lvert\eta_{sc}\rvert > 1.479$	
	loose ID	tight ID	loose ID	tight ID
Supercluster shower shape	< 0.0105	< 0.0104	< 0.0356	< 0.0305
$\lvert\Delta\eta(\text{sc, track})\rvert$	< 0.00387	< 0.00353	< 0.0072	< 0.00567
$\lvert\Delta\phi(\text{sc, track})\rvert$	< 0.0716	< 0.0499	< 0.147	< 0.0165
Hadronic/EM energy	$< 0.05 + c_1$	$< 0.026 + c_1$	$< 0.0414 + c_2$	$< 0.026 + c_2$
Electron isolation	< 0.133	< 0.0361	< 0.146	< 0.094
$\left\lvert\frac{1}{E_{sc}} - \frac{1}{p_{track}}\right\rvert$ (GeV^{-1})	< 0.129	< 0.0278	< 0.0875	< 0.0158
Expected missing hits	≤ 1	≤ 1	≤ 1	≤ 1
Passing conversion veto	✓	✓	✓	✓

5.1.4. Jets

Jets reconstructed from the PF algorithm rely on the correct measurement of energy clusters in the ECAL and HCAL. To reject jets only clustered from noise in these detector elements, a dedicated preselection of jets has to be performed to distinguish jets stemming from real physics interactions and these so-called noise jets. This is achieved by the PF jet ID [228], which involves different jet criteria based on the constituents of the clustered jet such as the number of constituents or the energy fraction of the ECAL and HCAL. In the central region of the detector, the tracker can be exploited to incorporate information about the charge from track measurements. The individual values for the jet ID are shown in Table 5.3, corresponding to the loose ID for data taken in 2015 and 2016. For the data taken in 2017, the values correspond to the tight ID, as the achieved efficiency for the tight ID is greater than 99% in all pseudorapidity regions. The efficiency is determined by the comparison of a noise-enriched minimum-bias selection with the selection of physical dijet events. A comparison of both selections for the different variables is shown in Fig. 5.2.
To correct for differences in the jet energy resolution between data and simulation, the predicted jet spectrum has to be smeared to describe the observation. This is achieved with a hybrid method, where the p_{T} of each jet is scaled either by a constant η-dependent scale factor if the jet can be matched to the corresponding generator-level jet or with a factor obtained from a stochastic smearing [232].

Table 5.3.: Requirements on reconstructed jets to be accepted by the analyses described in this thesis. The quality criteria are dependent on the pseudorapidity of the reconstructed jet to take the different subsystems of the detector into account, e. g., tracking information for charged particles is only available in the central region of the detector ($|\eta| \leq 2.4$). Changes with respect to the previous year are highlighted [229–231].

2015 data	$\lvert\eta\rvert \leq 2.4$	$2.4 < \lvert\eta\rvert \leq 2.7$	$2.7 < \lvert\eta\rvert \leq 3.0$	$\lvert\eta\rvert > 3.0$
Number of constituents	> 1	> 1	–	–
Neutral hadron fraction	< 0.99	< 0.99	–	–
Neutral EM fraction	< 0.99	< 0.99	< 0.90	< 0.90
Number of neutral particles	–	–	> 2	> 10
Charged hadron fraction	> 0	–	–	–
Charged EM fraction	< 0.99	–	–	–
Charged multiplicity	> 0	–	–	–

2016 data	$\lvert\eta\rvert \leq 2.4$	$2.4 < \lvert\eta\rvert \leq 2.7$	$2.7 < \lvert\eta\rvert \leq 3.0$	$\lvert\eta\rvert > 3.0$
Number of constituents	> 1	> 1	–	–
Neutral hadron fraction	< 0.99	< 0.99	**< 0.98**	–
Neutral EM fraction	< 0.99	< 0.99	**> 0.01**	< 0.90
Number of neutral particles	–	–	> 2	> 10
Charged hadron fraction	> 0	–	–	–
Charged EM fraction	< 0.99	–	–	–
Charged multiplicity	> 0	–	–	–

2017 data	$\lvert\eta\rvert \leq 2.4$	$2.4 < \lvert\eta\rvert \leq 2.7$	$2.7 < \lvert\eta\rvert \leq 3.0$	$\lvert\eta\rvert > 3.0$
Number of constituents	> 1	> 1	–	–
Neutral hadron fraction	**< 0.90**	**< 0.90**	**–**	**> 0.02**
Neutral EM fraction	**< 0.90**	**< 0.90**	**> 0.02 and < 0.99**	< 0.90
Number of neutral particles	–	–	> 2	> 10
Charged hadron fraction	> 0	–	–	–
Charged EM fraction	**–**	–	–	–
Charged multiplicity	> 0	–	–	–

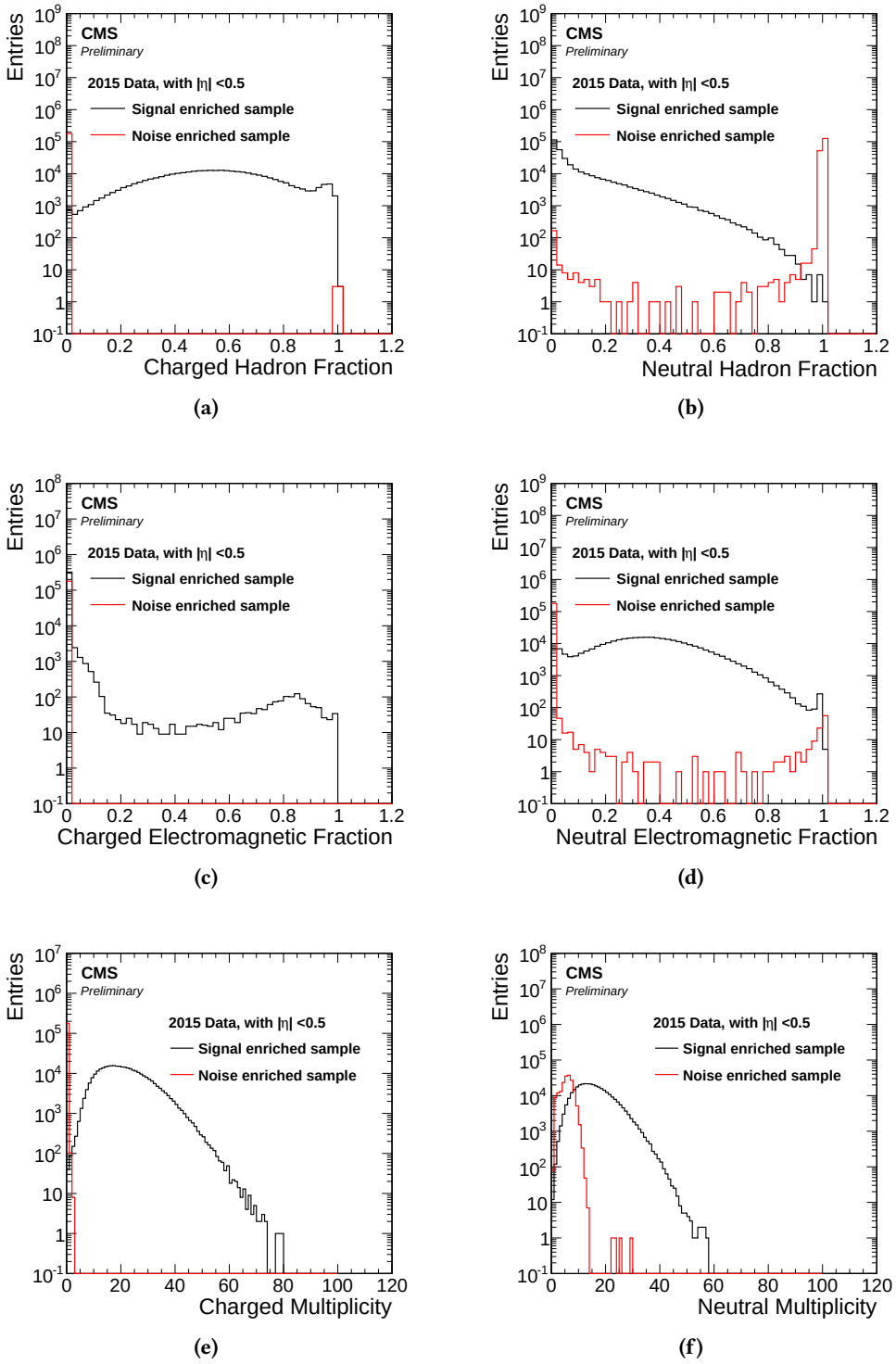


Figure 5.2.: Different variables required for a jet to pass the PF jet ID, in particular: the charged hadron fraction (a), the neutral energy fraction (b), the charged electromagnetic fraction (c), the neutral electromagnetic fraction (d), the charged multiplicity (e) and the neutral multiplicity (f). The distributions are shown for signal jets from physics processes and jets obtained from noise [228].

5.1.5. b Tagging

To obtain a reliable way to interpret the different values of b tagging algorithms, different working points of each algorithm are defined. These working points are defined in such a way that a selection of every value above a certain working point refers to a fixed mistagging probability. The different working points are loose, medium and tight and refer to a mistagging rate of about 10%, 1% and 0.1% for light-flavored or gluon jets, respectively. This results in a b tagging efficiency of 81% (loose), 63% (medium) and 41% for the CSVv2 algorithm in simulated top quark pair events [219]. The different working points of each year for the relevant b tagging algorithms used in this thesis are provided in Table 5.4.

Table 5.4.: The different working points of b tagging algorithms used in analyses in this thesis. For the data taken in the years 2015 and 2016, the CSVv2 algorithm is used, whereas for 2017 the DeepCSV algorithm is applied.

2015 data	loose	medium	tight
CSVv2	> 0.460	> 0.800	> 0.935

2016 data	loose	medium	tight
CSVv2	> 0.5426	> 0.8484	> 0.9535

2017 data	loose	medium	tight
DeepCSV	> 0.1522	> 0.4941	> 0.8001

5.2. W Boson Reconstruction

To be able to fully reconstruct a top quark, all of its decay products, i. e. the bottom quark and the W boson, have to be reconstructed first. While the bottom quark manifests itself as a jet in the detector, the W boson can decay either hadronically into a quark and an antiquark or leptonically into a charged lepton and the corresponding neutrino. More specifically, leptonically decaying W bosons in the context of this thesis means a W boson decaying into an electron or muon, where the electron or muon can either come directly from the W boson decay or from the subsequent decay of a leptonically decaying tau lepton, emerging from the W boson decay. Although the W boson decays directly into a quark and an antiquark in about $^2/_3$ of all cases, this decay mode is not considered for the reconstruction of a single top quark in this thesis as the background from QCD multijet events would be too overwhelming. Instead, only leptonically decaying W bosons are considered as stemming from a single top quark decay. However, this introduces the problem that the neutrino can only be detected indirectly in the transverse plane and the longitudinal information is missing to construct the full four-vector of the W boson. This problem can be solved by imposing a constraint on the W boson mass to the literature value of $m_\mathrm{W} = 80.385\,\mathrm{GeV}$ [3]. The mass of the W boson is defined by the kinematic properties of the decay products

$$\begin{aligned} m_\mathrm{W}^2 &= \left(E_\ell + \sqrt{(p_\mathrm{T}^\mathrm{miss})^2 + p_{z,\nu}^2}\right)^2 - (\vec{p}_{\mathrm{T},\ell} + \vec{p}_\mathrm{T}^\mathrm{miss})^2 - (p_{z,\ell} + p_{z,\nu})^2 \\ &= (80.385\,\mathrm{GeV})^2, \end{aligned} \tag{5.3}$$

with the energy of the lepton $E_\ell^2 = p_{\mathrm{T},\ell}^2 + p_{z,\ell}^2$. This equation can be solved for the longitudinal momentum component of the neutrino $p_{z,\nu}$, leading to a quadratic equation:

$$p_{z,\nu}^\pm = \frac{\Delta \cdot p_{z,\ell}}{p_{\mathrm{T},\ell}^2} \pm \sqrt{\frac{\Delta^2 \cdot p_{z,\ell}^2}{p_{\mathrm{T},\ell}^4} - \frac{E_\ell^2 \cdot (p_\mathrm{T}^\mathrm{miss})^2 - \Delta^2}{p_{\mathrm{T},\ell}^2}}, \tag{5.4}$$

with

$$\begin{aligned} \Delta &= \frac{m_\mathrm{W}^2}{2} + \vec{p}_{\mathrm{T},\ell} \cdot \vec{p}_\mathrm{T}^\mathrm{miss} \\ &= \frac{m_\mathrm{W}^2}{2} + p_{\mathrm{T},\ell} \cdot p_\mathrm{T}^\mathrm{miss} \cdot \cos(\Delta\phi), \end{aligned} \tag{5.5}$$

where $\Delta\phi$ is the angle between the lepton and the missing transverse momentum. Depending on the value of the discriminant, two different cases are possible. If the discriminant is positive, two different solutions are obtained from which the one with the smaller absolute value is chosen for $p_{z,\nu}$. If the discriminant becomes negative, two solutions with imaginary components are obtained. This can be caused by a non-perfect measurement of $p_\mathrm{T}^\mathrm{miss}$, giving the finite resolution of the detector. In this case, the $p_{x,\nu}$ and $p_{y,\nu}$ components are varied in such a way that the discriminant vanishes and only one real solution remains [233].
In addition to the full reconstructed W boson, a so-called transverse W boson mass can be defined as

$$m_\mathrm{T}^\mathrm{W} = \sqrt{(p_{\mathrm{T},\ell} + p_\mathrm{T}^\mathrm{miss})^2 - (p_{x,\ell} + p_x^\mathrm{miss})^2 - (p_{y,\ell} + p_y^\mathrm{miss})^2}. \tag{5.6}$$

This variable is helpful to distinguish processes with real W bosons, such as top quark production, from processes with virtual W bosons, which occur in QCD multijet events.
The reconstruction of a hadronically decaying W boson requires the combination of two jets and is therefore unambiguous compared to a leptonically decaying W boson. In this case, ambiguity only arises in events where the jet assignment is not well-defined, i. e., more than two jets exist in the event. This combinatorial problem is present in the search for single top production in association with a Higgs boson, where multivariate methods are used to find the best possible jet candidates for the hadronically decaying W boson (see Section 8.5.3).

6. Measurement of the t-Channel Single Top Quark Production Cross Section at $\sqrt{s} = 13$ TeV

Single top quark production was observed for the first time at the Tevatron experiments CDF and DØ in 2009 [234, 235], 14 years after the discovery of top quark pair production at the same collider. Single top quark production in the t channel has been also observed and measured at Run I of the LHC by the CMS Collaboration at a center-of-mass energy of 7 TeV [236, 237] and 8 TeV [238]. Compatible results have also been obtained by the ATLAS Collaboration [239–241] during Run I.

The t-channel single top quark production mode is not only the channel with the highest cross section among all single top quark production mechanisms in proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV, but also the channel with the most striking signature, i. e., the recoiling jet in the forward region of the detector. Therefore, studying single top quark production in the t channel is the best way to access the single top quark sector with the first data recorded by the CMS experiment at the Run II of the LHC in 2015. The cross section of t-channel single top quark production is measured with data from 2015 and additional quantities, such as the CKM matrix element V_{tb}, are calculated from the measured cross section. Two different data sets are analyzed in this chapter: one data set recorded in the summer of 2015 at a bunch spacing of 50 ns with an integrated luminosity of $42\,\mathrm{pb}^{-1}$ and the full data set at a bunch spacing of 25 ns, corresponding to an integrated luminosity of $2.2\,\mathrm{fb}^{-1}$.

This chapter describes in detail every aspect of the cross section measurement. The first section provides an overview of the general analysis strategy, followed by a section dedicated to an introduction to the signal and background processes along with their individual signatures. The third section is dedicated to the event selection to derive a signal-enriched phase space. In the fourth section, the simulation of events and corrections to simulated events are discussed. A special treatment for the background stemming from QCD multijet events is discussed in the fifth section. The measurement of the t-channel single top quark cross section with the early data of 2015 is presented in the sixth section. In the seventh section the improvements of the analysis with the full data set of 2015 with respect to the early measurement are discussed. One particular improvement, the multivariate analysis, is explained in detail in the eighth section. The ninth section is devoted to the fitting procedure to extract the cross section. Systematic uncertainties influencing the accuracy of the measurement are discussed in the tenth section. The final results are provided in the eleventh section, followed by a conclusion and an outlook in the twelth and last section of this chapter.

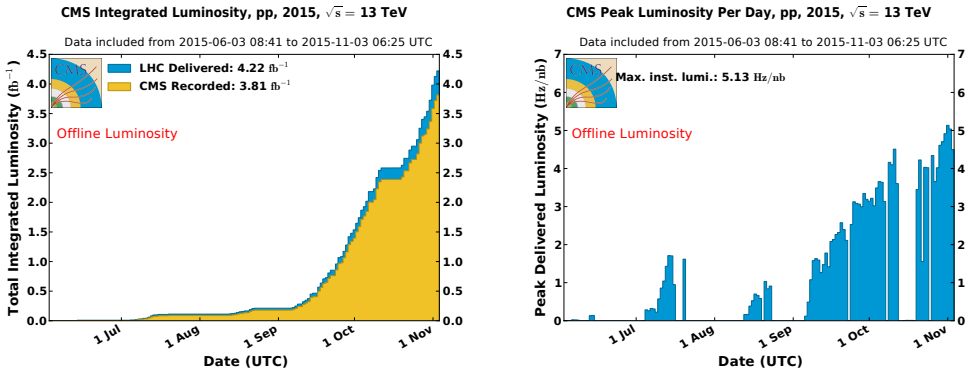


Figure 6.1.: The cumulated integrated luminosity of the data taken in 2015 (a) and the instantaneous luminosity (b) [125]. Because of problems with the magnetic field of the CMS detector, the data set available for measurements is reduced from 3.8 fb^{-1} to 2.2 fb^{-1} The measurement presented in this chapter is performed on an early data set of summer 2015, corresponding to an integrated luminosity of 42 pb^{-1}, and on the full data set of 2015.

6.1. Analysis Strategy

The goal of this analysis is to measure the inclusive cross section of t-channel single top quark production at a center-of-mass energy of $\sqrt{s} = 13$ TeV, both for the combined and separated cross section of top quark and top antiquark production. The analyzed data set corresponds to an integrated luminosity of 2.2 fb^{-1} and was recorded during the first year of proton-proton collisions at $\sqrt{s} = 13$ TeV, i. e., in 2015 [242]. This data set is only a subset of the total recorded data, but due to technical problems only this subset ensures a good quality of the accumulated data. The integrated and instantaneous luminosity over time of this year are shown in Fig. 6.1. A preliminary result is obtained with the first data recorded in summer 2015 at a bunch spacing of 50 ns. Because of technical problems of the LHC and the magnet cooling system of the CMS experiment at that time, the data set only corresponds to an integrated luminosity of 42 pb^{-1}.
To perform this analysis, the relevant signal and background processes are selected based on their final state and their cross section. Events are then selected based on the properties of the final state of the signal process to derive a signal-enriched phase space and additional background-dominated regions, based on the number of jets and b-tagged jets in an event. The simulation of the selected processes is then corrected by known mismodeling of observables and detector efficiencies to match the observation in data. A data-driven approach is then employed to estimate and model the difficult-to-handle background process of QCD multijet events. Signal and background processes are further separated by using multivariate analysis tools, in particular an artificial neural net. The cross section is then extracted using a maximum-likelihood estimation on the output distribution of the neural network.
The total cross section of single top quark and single top antiquark production is directly related to the V_{td}, V_{ts} and V_{tb} CKM matrix elements. With the assumption that the off-diagonal matrix elements are small,

$$|V_{\text{td}}|^2 + |V_{\text{ts}}|^2 \ll |V_{\text{tb}}|^2, \tag{6.1}$$

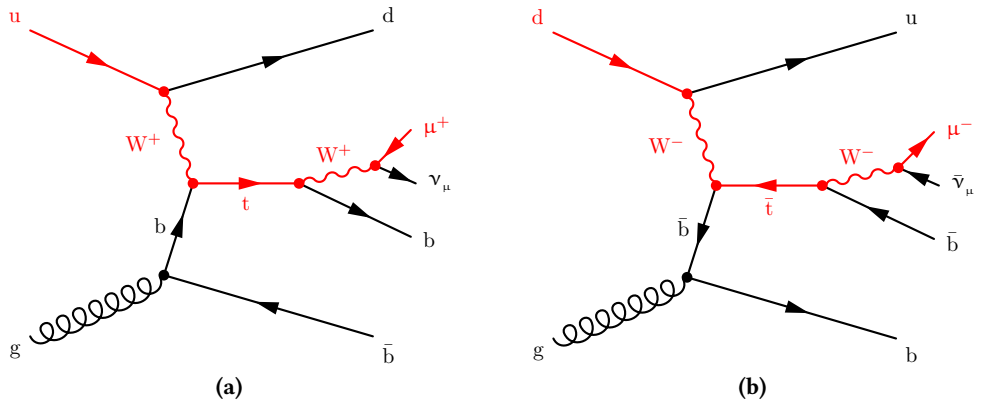


Figure 6.2.: Leading-order Feynman diagrams for single top quark (a) and single top antiquark (b) production in the t channel in the 4FS description. The electric charge of the muon from the top quark or top antiquark decay can be directly related to an up- or down-type quark in the initial state of the process. Therefore, t-channel single top quark production can be used to probe the initial state PDF.

the contribution of V_{td} and V_{ts} can be neglected. Furthermore, if all top quarks decay to bottom quarks, the cross section of single top quark production in the t channel $\sigma_{t\text{-ch.}}$ is directly proportional to the matrix element V_{tb} and the relation can be inverted:

$$|f_{LV} \cdot V_{tb}| = \sqrt{\frac{\sigma_{t\text{-ch.}}^{\text{measured}}}{\sigma_{t\text{-ch.}}^{\text{theory}}}}, \tag{6.2}$$

where $\sigma_{t\text{-ch.}}^{\text{measured}}$ is the measured cross section, $\sigma_{t\text{-ch.}}^{\text{theory}}$ is the predicted cross section and f_{LV} is a form factor for a possible anomalous Wtb coupling [243]. In the SM, f_{LV} is equal to 1. This approximation does not rely on the unitarity of the CKM matrix, meaning that a possible fourth generation of quarks is not forbidden.
The independent measurement of the single top quark and single top antiquark production cross sections can be used to determine the up- and down-type quark ratio of the proton, as illustrated in Fig. 6.2. By considering only leptonic W boson decays, the sign of the electric charge of the lepton is the same as for the top quark or top antiquark due to the charge current mediated by the W boson. The same can be applied to the production vertex, where either a top quark is produced (W^+ boson) or a top antiquark (W^- boson). Because of charge conservation, the initial-state parton can then be identified either as up- or down-type quark, providing a method to probe the corresponding PDFs of the proton. The quantity of interest for this PDF measurement is the ratio of the single top quark and antiquark production cross sections in the t channel:

$$R_{t\text{ ch.}} = \frac{\sigma_{t\text{-ch.,t}}}{\sigma_{t\text{-ch.,}\bar{t}}}. \tag{6.3}$$

Building the ratio of both cross sections has the advantage of reducing the impact of systematic uncertainties due to correlations between both measurements compared to an individual treatment.

6.2. Event Topology

An important part of the analysis is the separation of the signal process from the various background processes. Since the cross section of the signal process at 13 TeV is relatively small compared to the minimum-bias cross section, i. e., events with only soft QCD interaction (217 pb vs. 80 mb), the final-state topology of the signal process can be used to determine a kinematic region in which the signal-to-background ratio is more in favor of the signal. An overview of the signal process and all relevant background processes is provided in Fig. 6.3. Other background processes, such as s-channel single top quark production or diboson production, have been found to be negligible for the event selection described in the following section because either their cross section is too low or their final-state topology is rejected by the event selection requirements.

6.2.1. Signal Process

The signal process of interest is the t-channel single top quark production. The LO Feynman diagram of this process in the 4FS is shown in Fig. 6.3(a), including the leptonic decay of the W boson. The single top quark production in the t channel features a distinct signature, the light-flavored quark recoiling against the W boson, which helps to distinguish this process from all other background processes. This light-flavored quark has a relatively high transverse momentum compared to, e. g., additional quarks or gluons from the parton shower and is emitted in a more forward direction in the detector. Since the top quark decays into a bottom quark and a W boson, which then decays leptonically into a charged lepton and the corresponding neutrino, the final state contains a jet originating from a bottom quark, a charged lepton and missing transverse momentum, as the neutrino cannot be detected directly. An additional jet from a bottom quark is present in the final state from the initial gluon splitting. In summary, the final state of the t-channel single top quark production has two jets originating from bottom quarks, one charged lepton and missing transverse momentum. However, the jet from the initial gluon splitting has a soft transverse momentum spectrum, and therefore often fails the detector acceptance.

6.2.2. Background Processes

Various background processes are considered in this analysis and will be explained in detail in the following.

Associated tW Single Top Quark Production

Another single top quark production mode is considered as background in this analysis, the associated tW single top quark production. An exemplary Feynman diagram in the 5FS is shown in Fig. 6.3(b). The final state consists of a jet from the bottom quark of the top quark decay and two W bosons, which can either decay leptonically or hadronically. If one of the two W bosons decays leptonically and the other W boson hadronically, the final state consists of three jets, a charged lepton and missing transverse momentum, and therefore resembles the final state of the signal process. Fortunately, the cross section of the associated tW production is about three

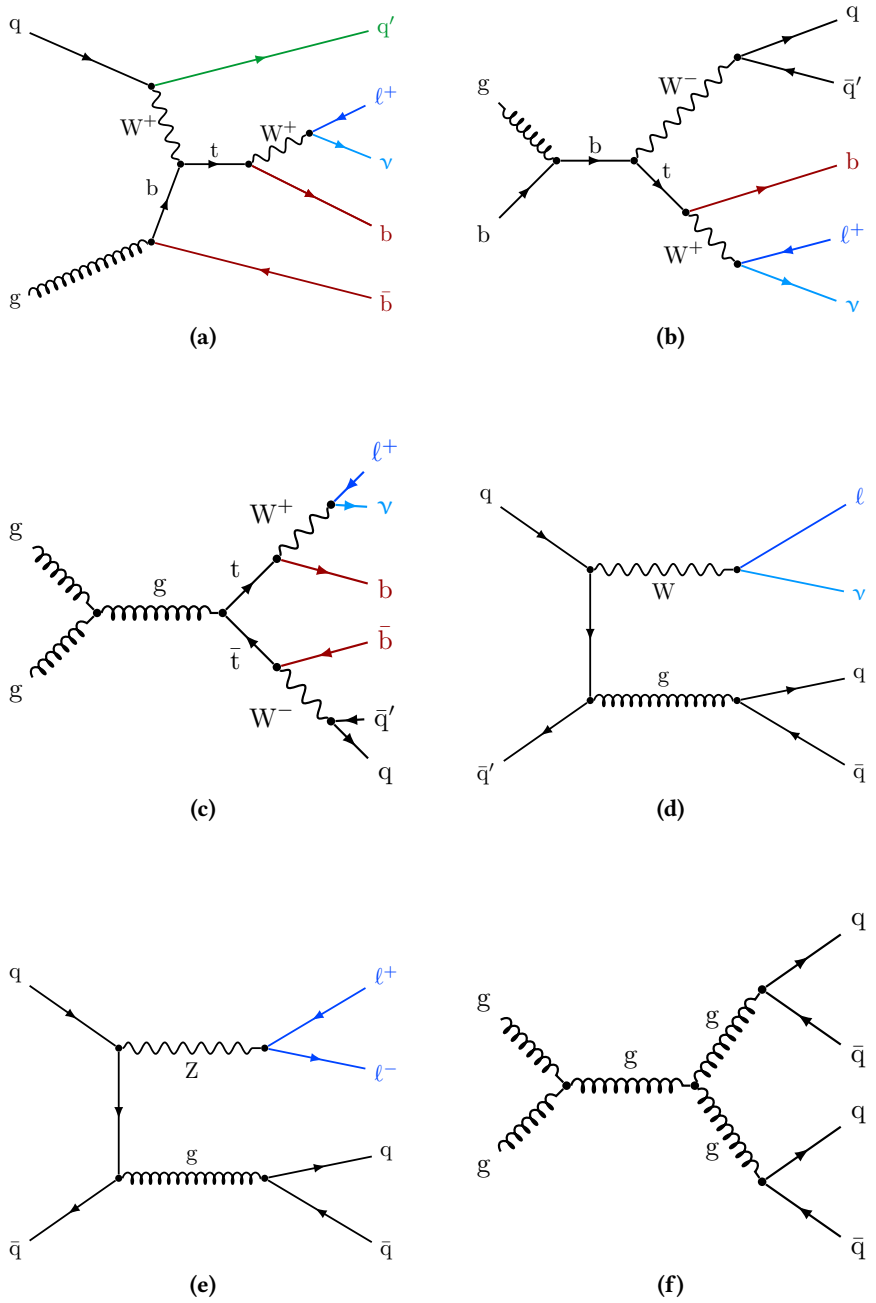


Figure 6.3.: Relevant signal and background processes for the measurement of the t-channel single top quark production cross section with their final-state decays. The signal process (a) features a distinct signature in the final state, the light-flavored jet in forward direction. Background processes are associated tW single top quark production (b), top quark pair production (c), production of W bosons (d) and Z bosons (e) in association with jets and QCD multijet production (f).

times lower than that of the single top quark production in the t channel. The contribution of the associated tW production can be further suppressed by utilizing b tagging information of jets, as only one of the three jets is originating from a bottom quark, compared to two jets for the t-channel production.

Top Quark Pair Production

The most dominant background for the measurement of t-channel single top quark production is the top quark pair production, as shown in Fig. 6.3(c). The composition of the final state is similar to that of the associated tW single top quark production, but with two jets from bottom quarks, instead of only one jet. Each of the two W bosons can again either decay leptonically or hadronically, but at least one of them has to decay leptonically in order to mimic the final-state topology of the signal process. To suppress contributions from top quark pair production, which is about four times more likely than the single top t-channel production, a requirement on the maximum number of jets in an event can be imposed. Along with requiring at least one charged lepton, the contribution from fullhadronic top quark pair production is completely negligible. Dileptonic top quark pair production is suppressed in a similar way by rejecting events with more than one charged lepton, but this requirement cannot reject all contributions from this decay mode, as the reconstruction of both leptons may not work properly in some events. The largest contribution arises from semileptonic top quark pair production, where some jets do not fall into the detector acceptance and the remaining objects can thus look similar to that of the final state of the signal process.

Production of Electroweak Bosons in Association with Jets

The second most dominant background does not involve top quarks and is the associated production of electroweak bosons with jets, in particular for this analysis the production of W bosons (W+jets), followed by minor contributions of Z bosons (Z+jets). Exemplary Feynman diagrams for the W+jets and Z+jets background processes are shown in Fig. 6.3(d) and in Fig 6.3(e), respectively. The vector bosons are considered to decay leptonically to create charged leptons, in this case either a charged lepton and the corresponding neutrino for W bosons or a charged lepton-antilepton pair for Z bosons, and the jets are originating from an additional gluon in the event. The Z+jets background contribution can be suppressed by imposing a requirement on the missing transverse momentum in an event or similar quantities, as there is no neutrino present in the final state. More problematic to reject are contributions from W+jets events, as the final state is similar to that of the signal process. In order to repel such events, a b tagging requirement on the jets of the gluon splitting can be imposed. Compared to single top quark or top quark pair production, the production cross section of the associated vector boson production is at least one magnitude higher and they still contribute to the relevant background processes because of imperfections of the detector and reconstruction algorithms.

QCD Multijet Events

Although QCD multijet events do not feature a charged lepton in the hard interaction, as illustrated in Fig. 6.3(f), this background cannot be neglected as its cross section is several orders of magnitude higher than that of any other process discussed before. A rare leptonic signature can be created from a leptonic W boson decay inside a meson of the hadronization process or jets can be misidentified as leptons, called fakes. As these rare circumstances, along with the high likelihood of such events overall, are difficult to model, a dedicated treatment of this background is employed and explained in Section 6.5.

6.3. Event Selection

Based on the final-state topology of signal and background processes, an event selection is applied, which not only aims for rejecting most of the background events to define a signal-enriched phase space (signal region), but also for defining additional regions in which the modeling of background processes can be checked for any deviations from the predictions (control region).
Every event must pass the online high-level trigger (HLT) path `HLT_IsoMu20_v*` to be accepted by the analysis. This HLT path searches for an isolated muon in the event with a transverse momentum of at least 20 GeV. To match this online selection, a dedicated criterion is applied on the reconstructed offline muon. Each event must contain exactly one tight muon, where a tight muon is required to have a transverse momentum of at least 22 GeV to take the turn-on of this HLT path into account, to pass the tight muon ID as described in Section 5.1.2 and by having a relative isolation of at most 0.06. With this selection criterion, most events with nonprompt muons stemming from QCD multijet events can be rejected. To suppress contribution from processes with multiple charged leptons, i. e., processes with a leptonically decaying Z boson or two leptonically decaying W bosons, events with additional loose leptons are rejected, which can either be muons or electrons. A loose muon is defined by having a transverse momentum of at least 10 GeV, passing the loose muon ID as defined in Section 5.1.2 and to have a relative isolation of at most 0.15. Electrons that have a transverse momentum of at least 20 GeV and passing the electron veto ID defined in Section 5.1.3, are considered loose electrons. Because of inefficiencies in the transition gap between the ECAL barrel and endcap, electrons within $1.4442 < |\eta_{sc}| < 1.566$ are excluded from the analysis. Jets are selected by requiring a transverse momentum of at least 40 GeV and an absolute pseudorapidity of at most 4.7. In addition, all jets must fulfill the jet ID requirements described in Section 5.1.4 and have a distance of $\Delta R > 0.4$ with respect to the tight muon in the η-ϕ plane. For b-tagged jets, the requirement on the transverse momentum is the same, but the jets have to be located in the central region of the detector ($|\eta| < 2.4$) where b tagging is possible because of the presence of the tracker system. To ensure a high rejection of light-flavored jets or jets from gluons, or equivalent a low mistagging rate, the CSVv2 b tagging algorithm is used at its tight working point, as described in Section 5.1.5. To reject remaining contributions from QCD multijet events, a requirement of 50 GeV on the reconstructed transverse W boson mass is enforced.
The selection criteria for leptons are shared among all regions, but the number of selected jets

Table 6.1.: The event selection for the different categories used in this analysis, intended to reject most contributions from background processes. The main signal region for the analysis is the 2-jets–1-tag region.

	2-jets–0-tag (2j0t)	2-jets–1-tag (2j1t)	3-jets–1-tag (3j1t)	3-jets–2-tags (3j2t)
Trigger match for `HLT_IsoMu20_v*`	✓	✓	✓	✓
Number of tight muons	1	1	1	1
Number of additional loose muons	0	0	0	0
Number of loose electrons	0	0	0	0
Number of selected jets	2	2	3	3
Number of b-tagged jets	0	1	1	2
Transverse W boson mass	> 50 GeV	> 50 GeV	> 50 GeV	> 50 GeV

and selected b-tagged jets defines the actual signal and control regions. As explained in the previous section, the second b jet from the initial gluon splitting has a rather soft transverse momentum spectrum. Therefore, this jet often fails the selection criteria and the signal region is required to have exactly two jets, from which exactly one them is b-tagged. The signal region is called 2-jets–1-tag, or short, 2j1t. In addition, the relative contribution from top quark pair production is significantly higher when three jets with two b-tagged jets are required. Thus, the 3-jets–2-tags (3j2t) region is a control region for the top quark pair background. The same applies to a selection of three jets, with one of them being b-tagged. This 3-jets–1-tag (3j1t) region also serves the purpose of constraining the background of top quark pair production. An additional region with exactly two jets, but explicitly no b-tagged jet, is defined to validate the QCD multijet event background estimation (2j0t), explained in Section 6.5. The event selection criteria for all regions are summarized in Table 6.1.

The top quark is reconstructed by combining the four-vectors of its decay products, i. e., the W boson, following the reconstruction explained in Section 5.2, and the jet associated to the bottom quark stemming from the top quark decay. The jet-to-quark assignment is trivial for regions where only one jet is supposed to be the stemming from a bottom quark (2-jets–1-tag and 3-jets–1-tag), but for the other two regions, the assignment is treated differently. In the 3-jets–2-tags region with exactly two b-tagged jets, the b-tagged jet with the higher value of the CSVv2 b tagging algorithm is used to reconstruct the top quark as this b-tagged jet is the most likely b jet. For the 2-jets–0-tag region, where no b-tagged jet is selected, the jet with lowest absolute pseudorapidity is assigned to the top quark decay. Consequently, the light-flavored jet in this region is the other, more forward jet. In the 3-jets–1-tag region, the more forward of the two untagged jets is assigned to the light quark, as the light quark is expected to be emitted in forward direction.

6.4. Simulation of Events and Corrections

The t-channel single top quark production signal is simulated using version 2.2.2 of the MADGRAPH5_aMC@NLO event generator [172] in the 4FS. The top quark pair background is modeled with POWHEG version 2.0 [176, 177], while the background from associated tW single top quark production is simulated with POWHEG version 1.0 [244] with the 5FS. All simulations use a top quark mass of 172.5 GeV. Background processes including vector boson production, i. e., W/Z+jets, are modeled with the MADGRAPH5_aMC@NLO event generator. The simulation samples for W/Z+jets production utilize the FxFx merging technique [165] to incorporate up to two additional partons at the matrix element level. A simulation sample of QCD multijet events is created with PYTHIA version 8.2 [180]. For the parton shower, the hadronization and the underlying event simulation PYTHIA with version 8.2 is used with the tune CUETP8M1 [181, 182]. The PDF set used for all simulation samples is the NNPDF30_nlo_as_0118 PDF [159]. An extensive list of all MC simulation samples is provided in Appendix A.1. Additional pileup interactions are added on top to the hard scattering, parton shower and hadronization to resemble the data-taking conditions.

6.4.1. Number of Pileup Interactions

The simulation of pileup interactions strongly depends on the conditions of the data-taking period, e. g., the instantaneous luminosity during data-taking. As the production of simulation samples is time consuming, simulation samples are produced in advance of the data-taking with a preliminary pileup profile. Although the kinematic distributions of most objects are unaffected by the actual pileup conditions due to the high requirements on the transverse momentum, a significant discrepancy can be observed in the number of primary vertices per event. This discrepancy is resolved by reweighting each event by a specific weight to reproduce the distribution observed in data. The weights are obtained by comparing the distribution in simulated samples and the distribution of data with a minimum-bias selection [245, 246]. The effect of this reweighting procedure on the number of primary vertices is shown in Fig. 6.4.

6.4.2. Muon Efficiencies

The reconstruction and selection of muons is affected by various, non-perfect reconstruction steps. The efficiencies of all these steps are combined in an overall muon reconstruction efficiency. Individual efficiencies for the muon ID, isolation and trigger are taken into account. The overall muon efficiency ϵ is defined as:

$$\epsilon = \epsilon_{\text{ID}} \cdot \epsilon_{\text{Iso|ID}} \cdot \epsilon_{\text{Trigger|Iso}}, \tag{6.4}$$

where ϵ_{ID} is the efficieny of the muon ID, $\epsilon_{\text{Iso|ID}}$ the efficiency of the muon isolation for the given ID and $\epsilon_{\text{Trigger|Iso}}$ the muon trigger efficiency for the given isolation. Efficiencies for the tight muon ID are provided by the MUO POG [223, 247]. For the muon isolation, only scale factors for values of 0.25 and 0.15 are centrally available, but not for the chosen isolation value of 0.06. Thus, the muon isolation efficiency and the subsequent muon trigger efficiency have been

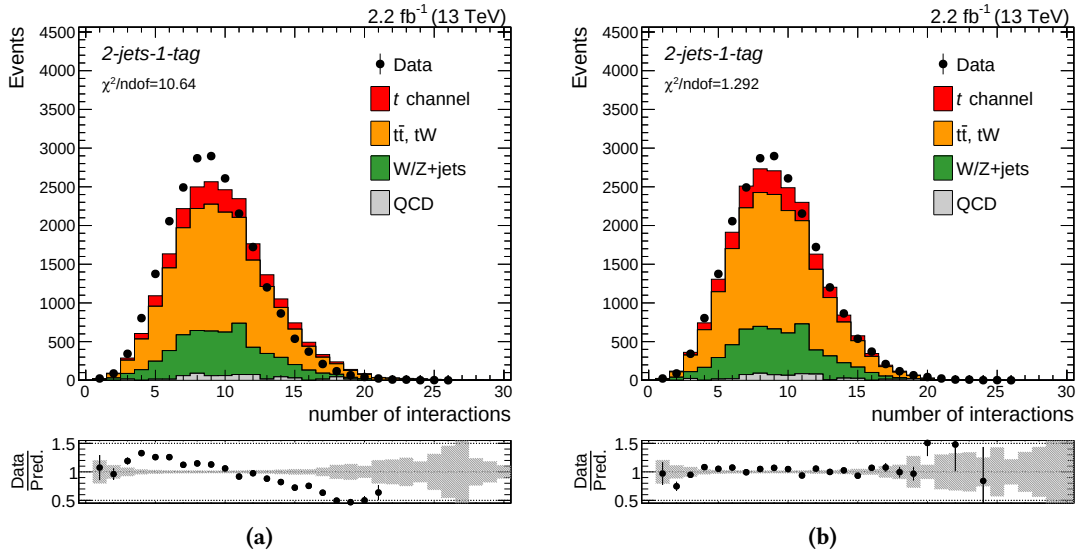


Figure 6.4.: Effect of the corrections applied to simulated events due to a different number of pileup interactions. The number of interactions in data is not well described (a) by the simulation. The mismodeling is corrected by applying an event weight to each simulated event to recover the distribution observed in data (b). A strong improvement of the prediction is observed. Both distributions are scaled to the integrated luminosity.

measured for this analysis with a tag-and-probe method [248–250]. The effect on the transverse momentum of the tight muon is illustrated in Fig. 6.5.

6.4.3. b Tagging Efficiencies

As the performance of b tagging algorithms is different for simulation and data, the performance in simulation has to be adjusted to match the observation in data. To recover the correct event yields, the simulated events are reweighted on an event basis, taking into account different efficiencies for every process [251]. The probability to obtain an event with N_i b-tagged jets, i. e., jets passing a certain working point of the b tagging algorithm, and N_j untagged jets in the correct b tag multiplicity category is given for MC simulations by

$$P_{\text{MC}} = \prod_{i \in \text{b-tagged}}^{N_i} \epsilon_i \prod_{j \in \text{untagged}}^{N_j} (1 - \epsilon_j) \tag{6.5}$$

and similarly, the probability in data is given by

$$P_{\text{Data}} = \prod_{i \in \text{b-tagged}}^{N_i} \text{SF}_i \epsilon_i \prod_{j \in \text{untagged}}^{N_j} (1 - \text{SF}_j \epsilon_j), \tag{6.6}$$

where ϵ is the process-dependent efficiency for a given jet to be b-tagged and SF are scale factors to correct discrepancies between data and simulation. These scale factors are centrally

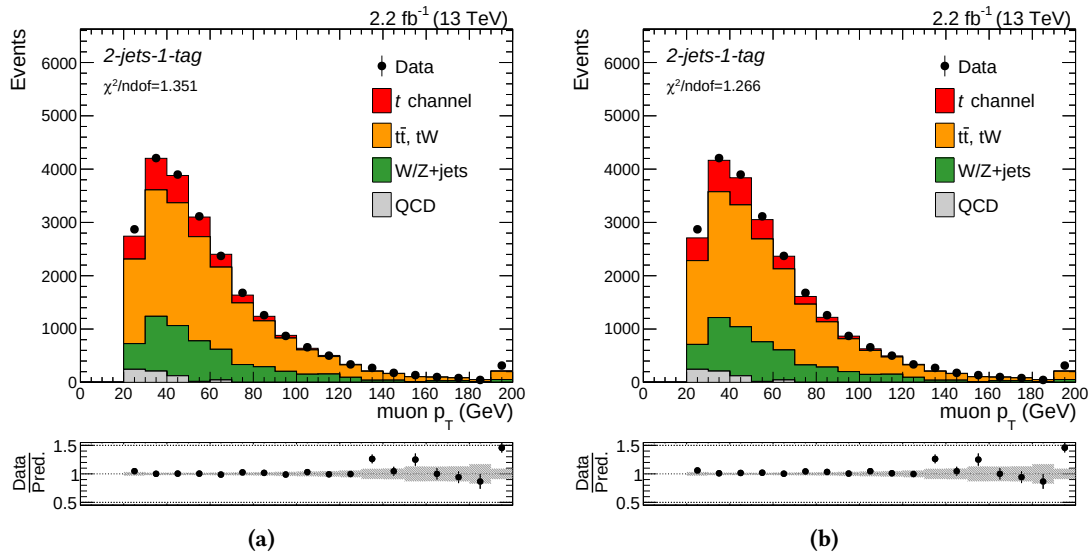


Figure 6.5.: Effect of correcting the simulation for muon ID, isolation and trigger efficiencies, exemplarily shown for the transverse momentum of the muon (a). A minor improvement to the modeling of the transverse momentum of the selected muon is observed when corresponding weights are assigned (b). Both distributions are scaled to the integrated luminosity.

provided by the b Tag and Vertexing Physics Object Group (BTV POG) [252] for simulations of the 2015 data-taking period [253]. The tagging efficiencies are evaluated independently for each simulated process to account for variations due to different topologies. The final event weight w is then obtained by the ratio of the two quantities above:

$$w = \frac{P_{\text{Data}}}{P_{\text{MC}}}. \tag{6.7}$$

The effect of the reweighting is shown in Fig. 6.6 for the transverse momentum of the b-tagged jet in the signal region.

6.5. Background Estimation for QCD Multijet Events

Because of the high cross section and low selection efficiency of QCD multijet production, this background component is difficult to estimate and model with MC simulation samples, as a correct prediction can only be achieved with an extremely high number of simulated events. Therefore, the contribution from QCD multijet production is estimated with a data-driven technique to ensure an adequate prediction of this background. This is achieved by defining a sideband region enriched in QCD multijet events. The sideband region is defined by inverting the selection criterion for the muon isolation, i. e., muons are required to have a relative isolation higher than 0.12, instead of lower than 0.06 for the signal and control regions. The modeling of other kinematic variables of QCD multijet production is independent of the selection into the signal region or this specific sideband region, thus, the modeling of variables

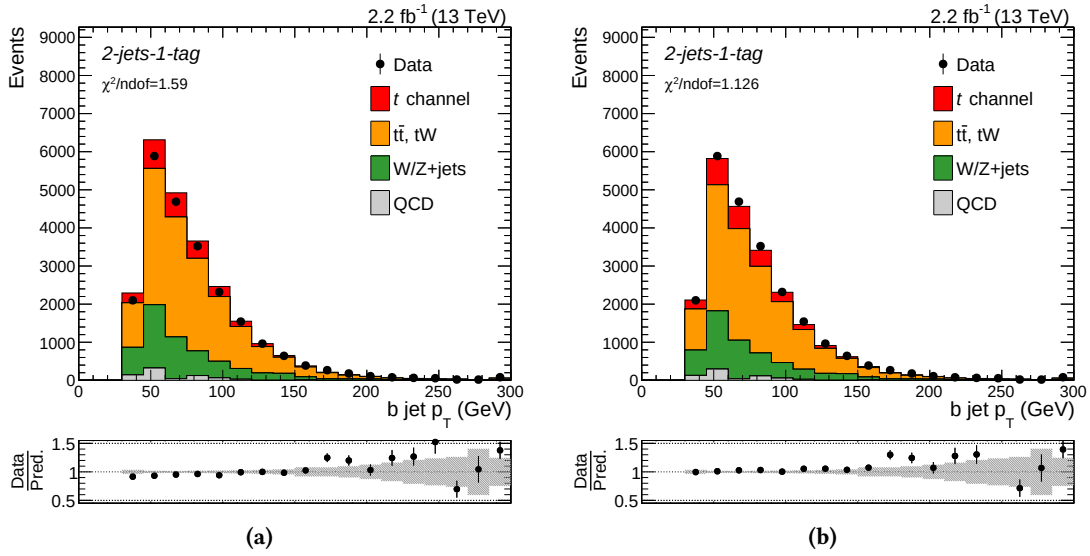
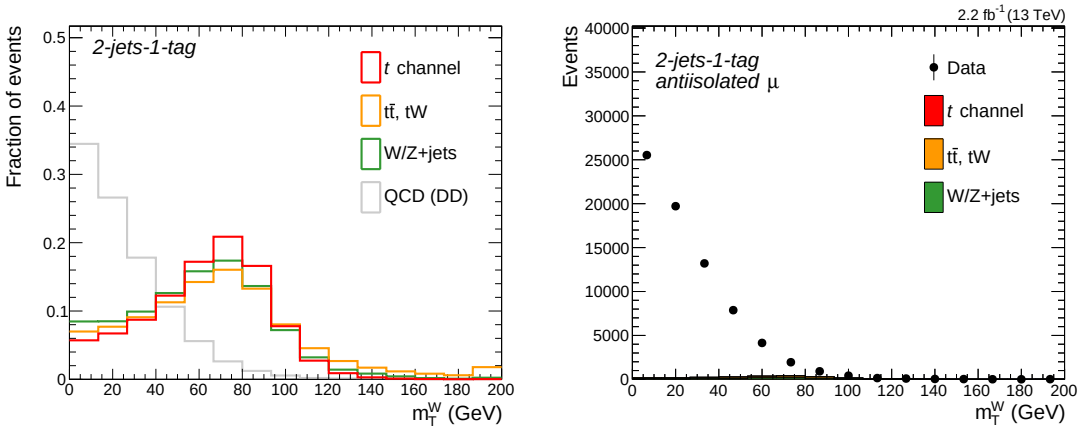


Figure 6.6.: Effect of correcting for different b tagging efficiencies in simulation and data on the transverse momentum of the jet assigned to the bottom quark in the 2-jets–1-tag signal region before (a) and after the correction (b). Both distributions are scaled to the integrated luminosity.

in this sideband can be used for the signal region. To estimate the correct normalization of the QCD multijet production background, a maximum-likelihood fit is performed to the distribution of the transverse W boson mass $m_{\mathrm{T}}^{\mathrm{W}}$. This single variable can help to distinguish between QCD multijet events and all other processes involving top quarks and electroweak bosons in the standard event selection. This is attributed to the fact that the distribution of the transverse W boson mass has a maximum around the W boson mass for processes with a real W boson, while the distribution is steeply falling for processes without real W bosons, such as QCD multijet events. The distribution of the transverse W boson mass in the 2-jets–1-tag signal region is illustrated for different processes in Fig. 6.7(a). The maximum-likelihood fit is performed with two templates, one for QCD multijet events and one for the sum of non-QCD processes. The template for the QCD multijet events is obtained in the sideband region with the selection with inverted isolation criterion (antiisolated) by removing all non-QCD contributions from the distribution of the transverse W boson mass. The distribution in the sideband region is shown in Fig. 6.7(b). The maximum-likelihood fit is then performed on data in the signal region without the selection criterion for the transverse W boson mass with the following formula for the fit function F:

$$F(m_{\mathrm{T}}^{\mathrm{W}}) = N_{\mathrm{QCD}} \cdot F_1(m_{\mathrm{T}}^{\mathrm{W}}) + N_{\text{non-QCD}} \cdot F_2(m_{\mathrm{T}}^{\mathrm{W}}), \tag{6.8}$$

where F_1 is the template obtained from the sideband region and F_2 the template of non-QCD processes in the signal region, derived from simulations. The numbers N_{QCD} and $N_{\text{non-QCD}}$ are the fit parameters and represent the number of QCD multijet and non-QCD events, respectively. The number of QCD multijet events for the full event selection, i. e., including the cut on the transverse W boson mass of 50 GeV, is obtained by calculating the integral of the fitted QCD



(a) **(b)**

Figure 6.7.: As the distribution of the transverse W boson can distinguish between QCD multijet events and non-QCD processes (a), this variable is well suited to estimate the QCD multijet background contribution. As the size of the simulation sample for the QCD multijet background is not sufficient, a data-driven approach in an antiisolated, QCD-enriched sideband is used (b), with only minor contribution from non-QCD processes.

multijet template $F_1(m_{\mathrm{T}}^{\mathrm{W}}) \cdot N_{\mathrm{QCD}}$ from 50 GeV to infinity.
The data-driven QCD estimation is first tested and validated in the 2-jets–0-tag control region, which mainly consists of W+jets and QCD multijet events. The procedure is then applied to the 2-jets–1-tag signal region and the 3-jets–1-tag control region. The contribution of the QCD multijet event background to the 3-jets–2-tags control region is found to be negligible, therefore the procedure is not applied in this region.

6.6. Evidence for Single Top Quark Production at $\sqrt{s} = 13$ TeV with the First Data in 2015

The primary goal of the single top quark t-channel analysis in 2015 was to rediscover single top quark production at a center-of-mass energy of $\sqrt{s} = 13$ TeV. The t-channel process is the best choice to accomplish this goal, as not only the cross section is the largest among all single top quark production modes, but also the signature of the final state is special enough to distinguish the signal from the background processes without any multivariate approach. Because of this, the t-channel single top quark production cross section is extracted in this first measurement by performing a maximum-likelihood fit to the distribution of the absolute pseudorapidity of the light-flavored jet in the 2-jets–1-tag signal region. The qualitative shape of this distribution is shown in Fig. 6.8 for the signal and all relevant background processes. The measurement is performed on the data set recorded in summer 2015 with a bunch spacing of 50 ns. Although the original predictions for the integrated luminosity were around 1 fb^{-1}, the data set during this

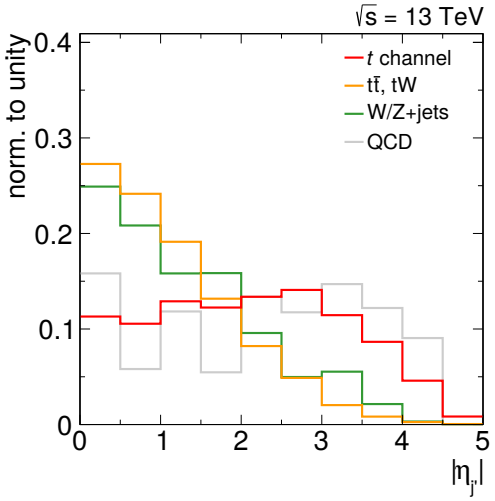


Figure 6.8.: Shape distribution of the absolute pseudorapidity of the light-flavored jet in the 2-jets–1-tag signal region for the signal process and the relevant background processes. This particular variable is well-suited to distinguish the signal against the background processes and is therefore used to extract the t-channel single top quark production cross section in this early measurement.

period only corresponds to an integrated luminosity of 42 pb^{-1} due to technical problems with the LHC and the cooling system of the CMS experiment. Nevertheless, the accumulated data is still sufficient to perform an early measurement of the t-channel single top quark production cross section at $\sqrt{s} = 13$ TeV.

Compared to the event selection described in Section 6.3, the 2-jets–1-tag region is further divided into a signal region (SR) and a sideband (SB) region, where the SR contains all events inside the reconstructed top quark mass window of 130 GeV $< m_{\text{top}} <$ 225 GeV and the SB region contains all events outside of this window. This additional selection aims at rejecting additional W+jets background events. The predicted and observed event yields for the SR and the SB are listed in Table 6.2. The estimation of the QCD multijet background is performed independently for the SR and the SB. The results are presented in Fig. 6.9. The distributions of the absolute pseudorapidity in the 2-jets–1-tag SB and the 3-jets–2-tags region, for the latter without the requirement of $m_{\text{T}}^{\text{W}} > 50$ GeV, are shown in Fig. 6.10. The cross section of single top quark production in the t channel is then extracted with a binned maximum-likelihood fit in the 2-jets–1-tag SR and the 3-jets–2-tags region simultaneously. Adding the additional 3-jets–2-tags control region to the fit provides additional signal contribution, but the main motivation is to constrain the top quark pair production background and to decorrelate this contribution from the W+jets background, as the $|\eta|$ distribution of both processes is the same in the 2-jets–1-tag SR, but slightly different in the 3-jets–2-tags region. The result of the maximum-likelihood fit is presented in Fig. 6.11, where each simulation template is scaled to the result of the fit. The distribution of the reconstructed top quark mass for the fit results is shown in Fig. 6.12 for the entire 2-jets–1-tag SR and a signal-enriched subset, where the light-flavored jet is required to have an absolute pseudorapidity of at least 2.5. The result of the fit yields a single top quark

Table 6.2.: Expected and observed event yields for the 2-jets–1-tag SR and the SB. A 50% uncertainty is assigned to the expected number of QCD multijet events due to the estimation of this background. For the other processes, the quoted uncertainties refer to the limited size of simulation samples.

	Signal region (SR)	Sideband (SB)
Top quark background ($t\bar{t}$ and tW)	157 ± 1	71.7 ± 0.4
W/Z+jets	40 ± 4	47 ± 4
QCD multijet events	10 ± 5	2 ± 1
t-channel single top quark	33 ± 1	7.2 ± 0.3
Total expected	240 ± 6	128 ± 4
Data	252	127

production cross section in the t channel of

$$\sigma_{t\text{-ch.}} = 274 \pm 98\ (\text{stat}) \pm 52\ (\text{syst}) \pm 33\ (\text{lumi})\ \text{pb}, \tag{6.9}$$

where the terms stat, syst and lumi refer to the statistical uncertainty, the systematic uncertainties and the uncertainty on the measured integrated luminosity, respectively. The dominant uncertainty of this measurement is the limited number of data events, represented by the statistical uncertainty given above. The observed significance of the signal corresponds to 3.5 standard deviations, whereas only 2.7 standard deviations were expected for the given integrated luminosity. Thus, evidence for single top quark production in the t channel can already be claimed, despite the very small data set. A breakdown of systematic uncertainties is presented in Table 6.3. A detailed explanation of each individual source of systematic uncertainty is provided in Section 6.10. The extracted single top quark cross section is in agreement with the SM prediction within the uncertainties of the measurement. In Fig 6.13, the result is compared to the theory prediction and to previous measurements of the CMS Collaboration at center-of-mass energies of 7 and 8 TeV, as well as to the combined result of experiments at the proton-antiproton Tevatron collider. From the measured cross section, the V_{tb} matrix element is extracted:

$$|f_{\text{LV}} \cdot V_{\text{tb}}| = 1.12 \pm 0.24\ (\text{exp}) \pm 0.02\ (\text{theo}), \tag{6.10}$$

with the combined uncertainties of the measurement (exp) and the theoretical uncertainties (theo). An event display of a typical t-channel single top quark candidate is provided in Fig. 6.14. The analysis of the first data in 2015 has been published by the CMS Collaboration as a Physics Analysis Summary [255] and was the first measurement of the single top quark cross section at a center-of-mass-energy of 13 TeV at the LHC.

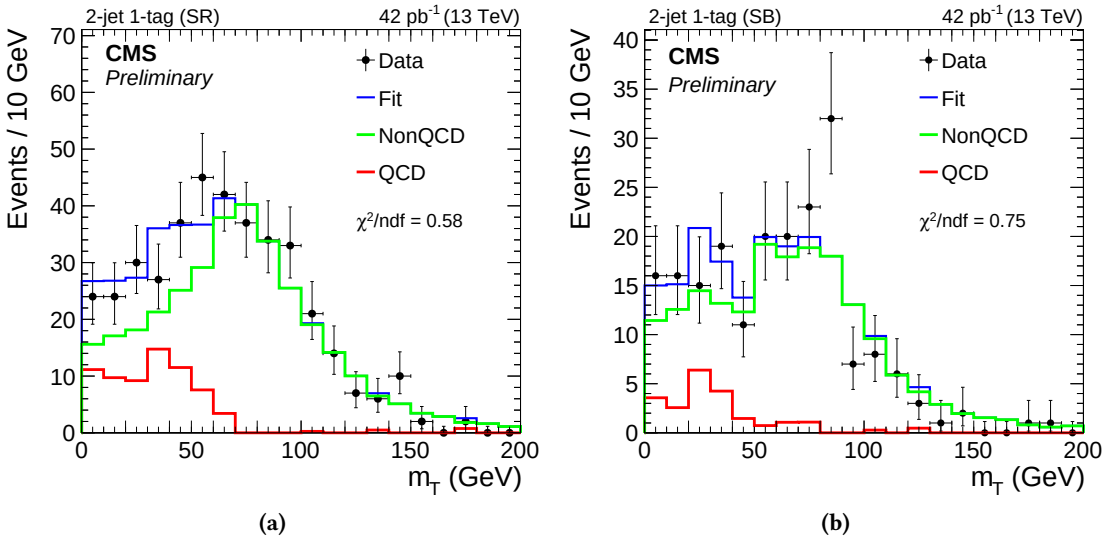


Figure 6.9.: Result of the QCD multijet background estimation in the 2-jets–1-tag region, following the description in Section 6.5, performed independently for the signal region (a) and the sideband (b). The templates are scaled to the result of the fit.

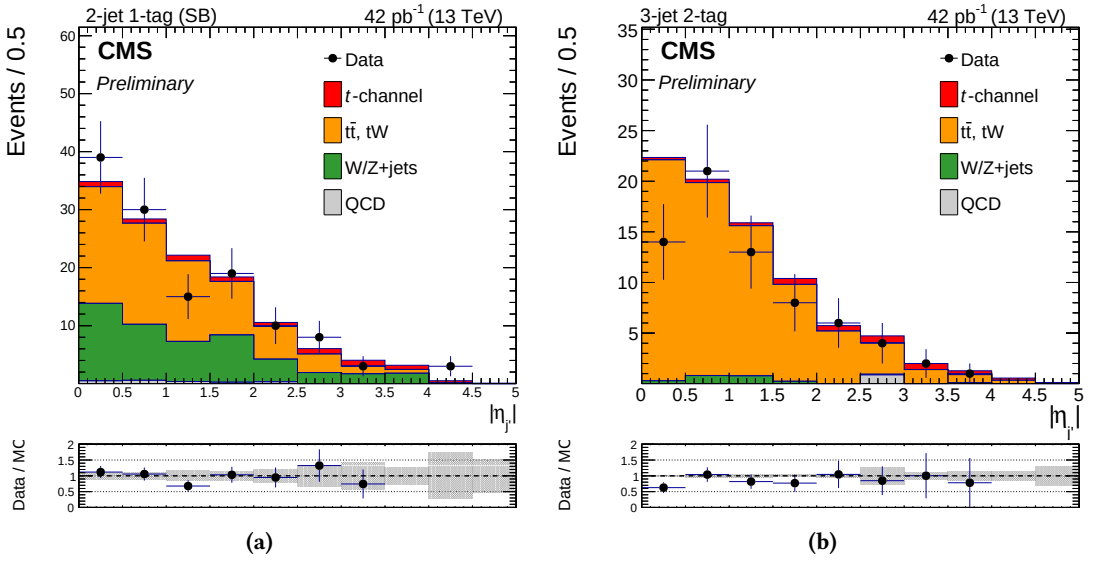


Figure 6.10.: The distribution of the absolute pseudorapidity of the light-flavored jet in the 2-jets–1-tag sideband region (a) and the 3–jets-2-tags control region (b), in which the requirement $m_T^W > 50$ GeV is removed. Sufficient agreement of simulation and data is observed for the given uncertainties. Both distributions are scaled to the integrated luminosity.

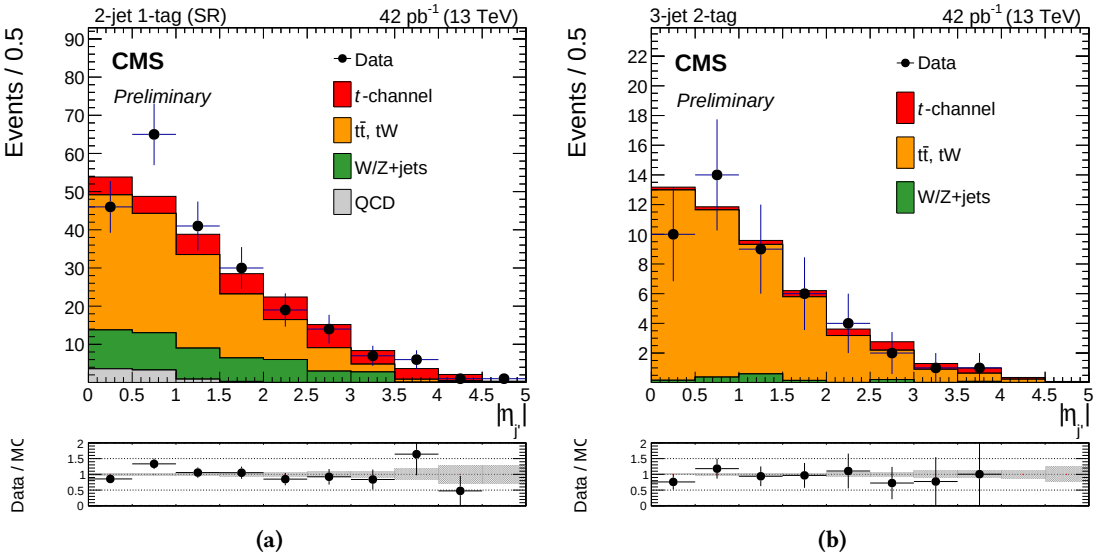


Figure 6.11.: Resulting distributions of the absolute pseudorapidity of the light-flavored jet in the 2-jets–1-tag SR (a) and the 3-jets–2-tags control region (b) after the combined fit for $m_{\mathrm{T}}^{\mathrm{W}} > 50$ GeV. Good agreement of simulation and data is observed for the given uncertainties. Both distributions are scaled to the result of the fit.

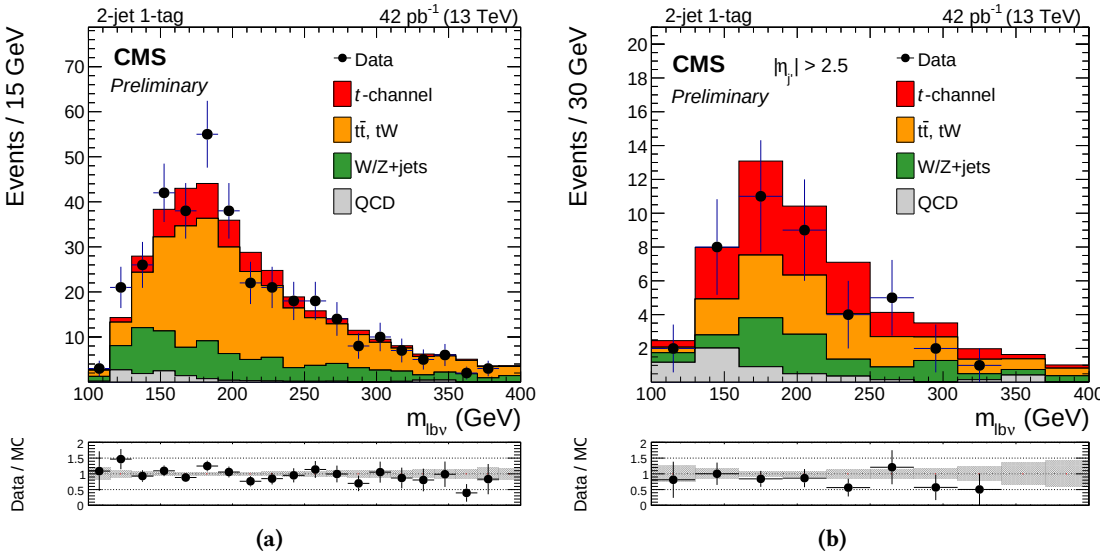


Figure 6.12.: Distributions of the reconstructed top quark mass in the 2-jets–1-tag SR for the entire region (a) and for events where the light-flavored jet is required to have an absolute pseudorapidity of at least 2.5 (b). The latter clearly demonstrates that the data is only accurately described around the top quark mass of 172.5 GeV by the simulation if t-channel single top quark production is realized in nature. Both distributions are scaled to the result of the fit.

Table 6.3.: Impact of the statistical and each of the systematic uncertainties on the measured cross section of t-channel single top quark production. The most dominant systematic uncertainties are the jet energy scale, b tagging and the choice of the PDF. Additional large uncertainties arise from the measured value of the integrated luminosity. Although all the previously mentioned uncertainties are relatively large, the most dominant uncertainty in this measurement is the statistical uncertainty due to the small data set.

Uncertainty source	$\Delta\sigma_{t\text{-ch.},\,t+\bar{t}}/\sigma^{\text{obs}}_{t\text{-ch.},\,t+\bar{t}}$
Jet energy scale	17%
Jet energy resolution	1.1%
b tagging	5.6%
Muon trigger/reconstruction	3.4%
QCD extraction	1.1%
Signal generator	1.9%
Factorization and renormalization scales	3.3%
PDF	4.5%
Missing transverse momentum	1.2%
Pileup	1.4%
Total systematic uncertainty	19%
Statistical uncertainty	36%
Luminosity	12%
Total uncertainty	42%

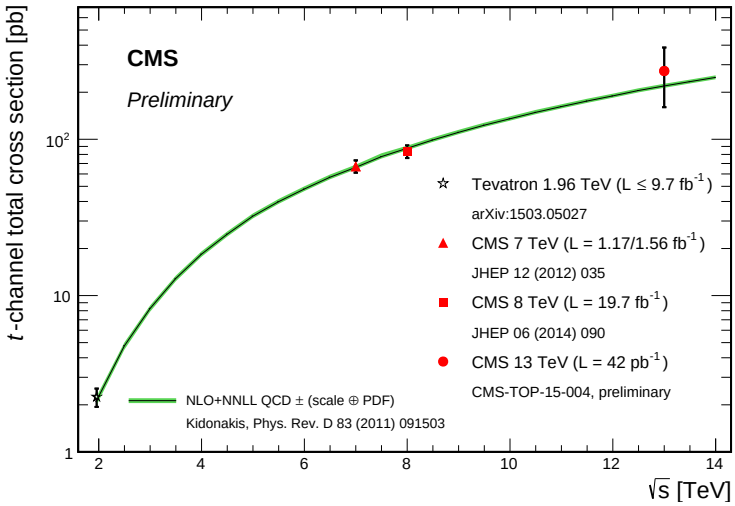


Figure 6.13.: Comparison of the measured t-channel single top quark production cross section at 13 TeV with the theory prediction. In addition, measured values at 7 and 8 TeV from the CMS Collaboration and the combined result of the Tevatron are shown. All measured cross sections are in agreement with the prediction from theory.

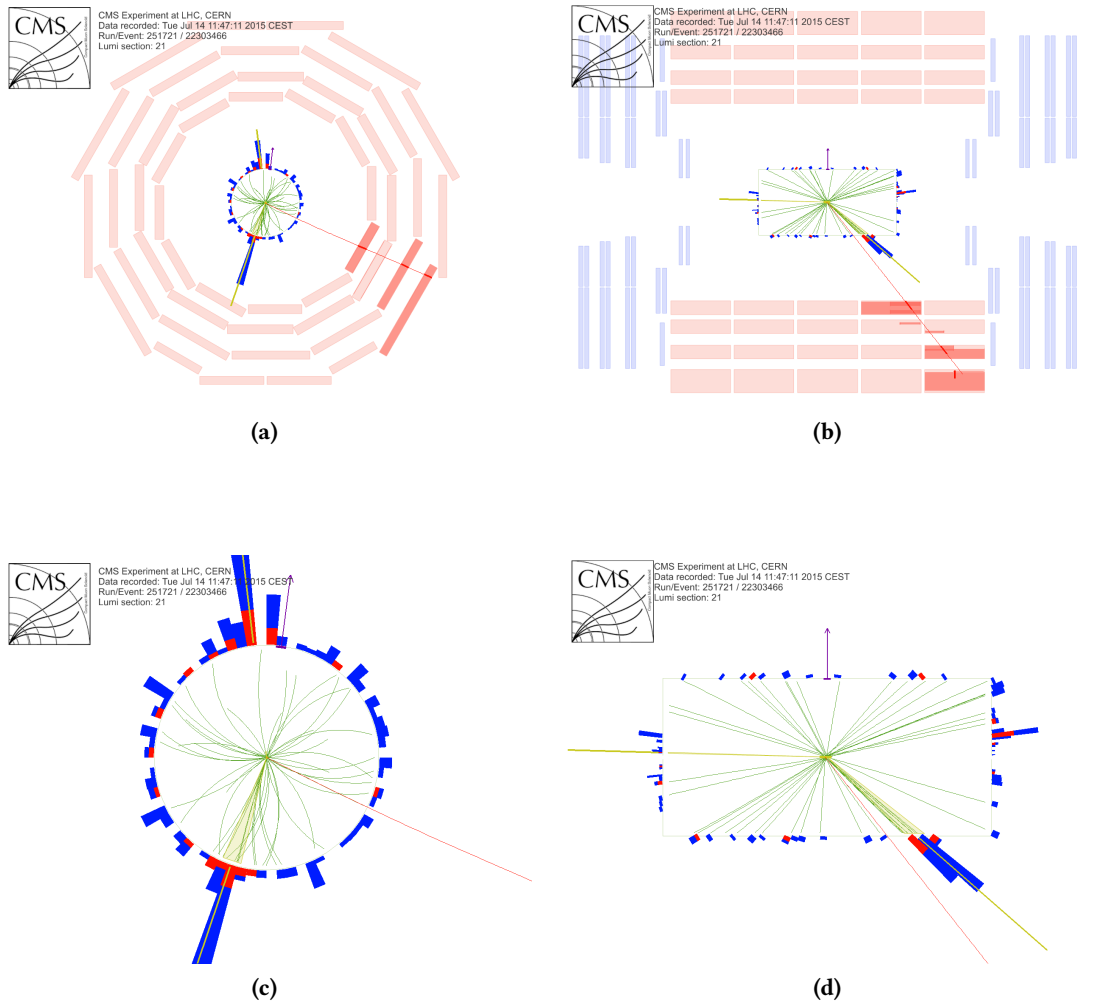


Figure 6.14.: An event display of a specific proton-proton collision recorded with the CMS experiment during the 50 ns bunch spacing run of the LHC in July 2015. The displays on the left ((a) and (c)) show the event in the x-y plane and the ones on the right ((b) and (d)) along the z-axis. This event shows all characteristics of a t-channel single top quark event: one b-tagged jet and one untagged jet, an isolated muon and missing transverse momentum. The untagged jet has a pseudorapidity of -4.3 and is interpreted as the light quark recoiling against the top quark. Combining the b-tagged jet, the muon and the missing transverse momentum results in a reconstructed top quark mass of 177 GeV [254].

Table 6.4.: Predicted and observed event yields for the 2-jets–1-tag signal region in the full 2015 analysis (2.2 fb^{-1}), split by the charge of the selected muon. The uncertainties are due to the size of the simulation samples, except for the QCD multijet events, where an uncertainty of 50% is assigned to the values obtained from the estimation in Section 6.7.1.

Process	μ^+	μ^+
Top quark background ($t\bar{t}$ and tW)	6837 ± 13	6844 ± 13
W/Z+jets	2752 ± 82	2487 ± 76
QCD multijet events	308 ± 154	266 ± 133
Single top quark t-channel	1493 ± 13	948 ± 10
Total expected	11 390 ± 175	10 545 ± 154
Data	11 877	11 017

6.7. Analysis with the Full Data Set of 2015

Compared to the early data set with an integrated luminosity of 42 pb^{-1}, the full data set of 2015 contains more than 50 times more events. Therefore, the statistical uncertainty is expected to be reduced drastically for a measurement on the full data set with respect to the early data set. The increased amount of data allows to improve the analysis with a multivariate approach and to measure the single top quark production in the t channel independently for top quarks and top antiquarks, as outlined in Section 6.1, to compare the measured ratio of up- and down-type quarks in the proton with the predictions of different PDF sets. The differentiation of the 2-jets–1-tag region into a SR and a SB based on the reconstructed top quark mass is no longer required, since the W+jets background process can be separated from the signal process with the help of the multivariate analysis. The event yields obtained for the selection of the 2-jets–1-tag region are presented in Table 6.4, split by the charge of the muon.

6.7.1. QCD Multijet Background Estimation

The estimation of the QCD multijet background is performed as described in Section 6.5. The maximum-likelihood fit is performed independently for a selection of μ^+ and μ^-. As a cross check, also the inclusive fit, i. e., with the combined selection, is performed separately. The method is validated in the 2-jets–0-tag region, which contains roughly ten times more events compared to the 2-jets–1-tag signal region. The result of the fit in the 2-jets–0-tag control region is presented in Fig. 6.15. Good agreement between the result of the fit and the data is observed. The procedure yields a QCD multijet event contribution to the 2-jets–1-tag signal region, i. e. for $m_{\text{T}}^{\text{W}} > 50\,\text{GeV}$, of 308 events for the μ^+ category and 266 events for the μ^+ category. To cover any possible systematic effects specific to this method, a conservative 50% uncertainty is assigned to the obtained yields. The result from the charge-split estimation is in agreement with the inclusive estimation, which yields 589 events. The results of the fit are shown in Fig. 6.16.

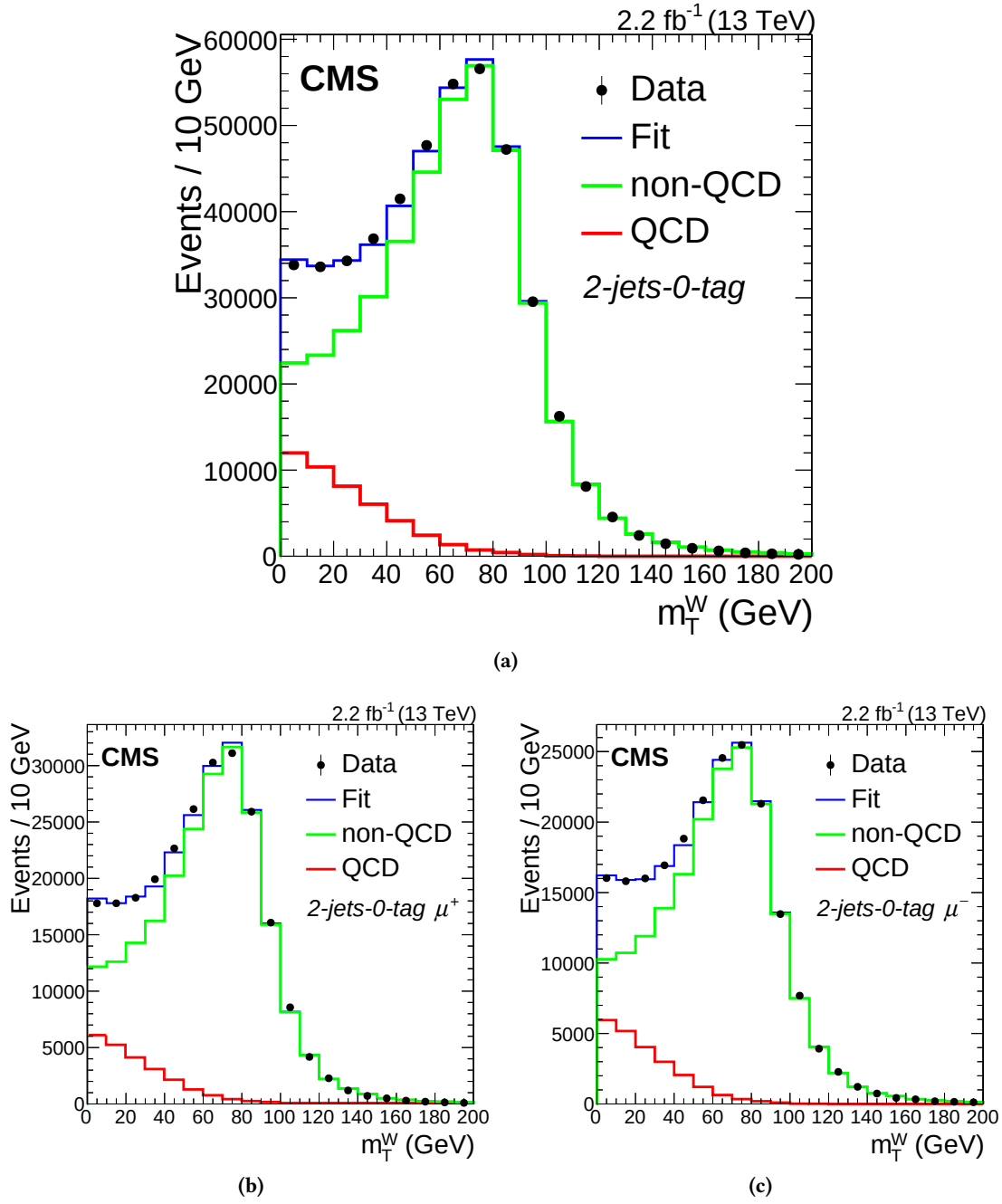
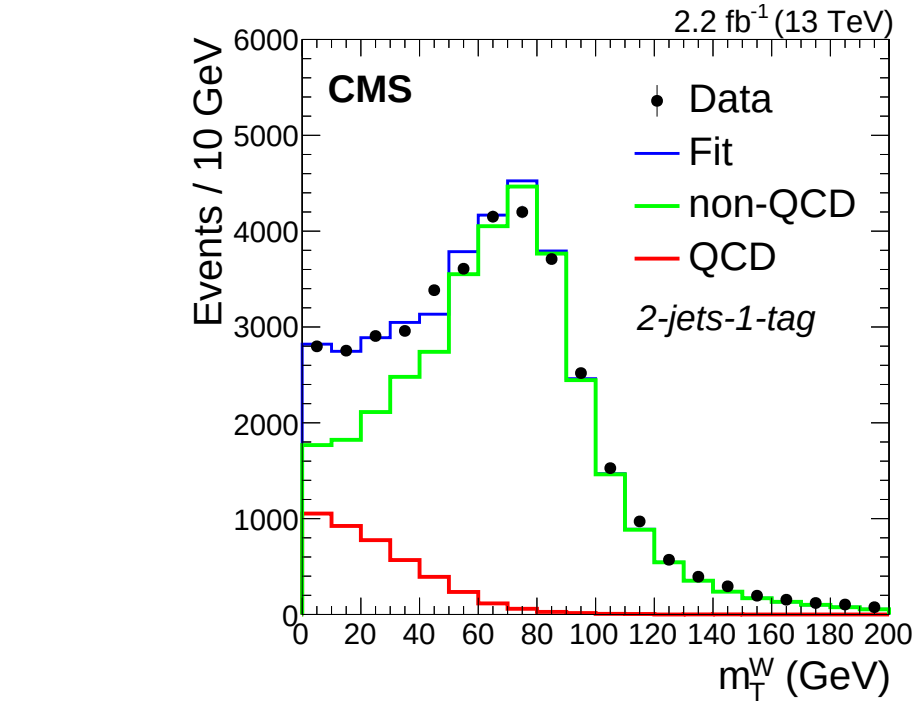
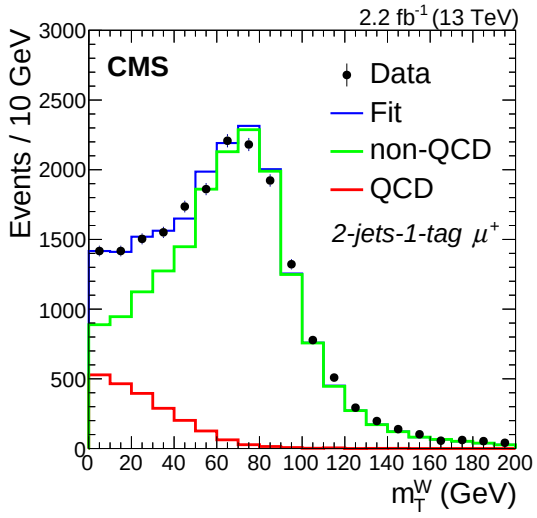


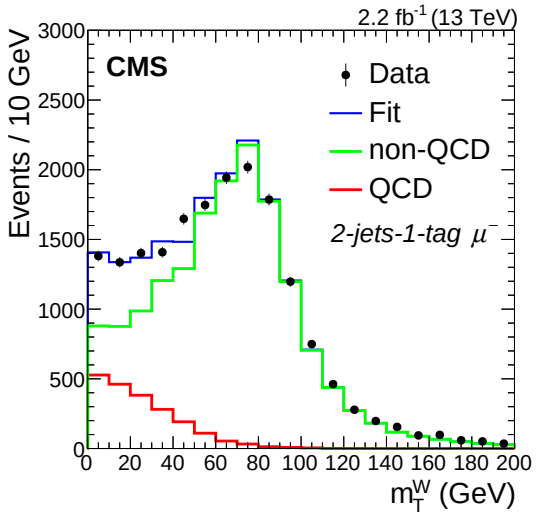
Figure 6.15.: Estimation of the contribution from QCD multijet events in the 2-jets–0-tag control region for the inclusive selection (a), only μ^+ events (b) and only μ^- events (c). The templates are scaled to the result of the fit. Good agreement between simulation and data for all three categories is observed.



(a)



(b)



(c)

Figure 6.16.: Estimation of the contribution from QCD multijet events in the 2-jets–1-tag signal region for the inclusive selection (a), only μ^+ events (b) and only μ^- events (c). The templates are scaled to the result of the fit. Good agreement between simulation and data for all three categories is observed.

Table 6.5.: Ranking of all relevant input variables in the training of the neural network. The ranking is determined by loss in significance if the corresponding variable is removed from the variable set. The missing transverse momentum was also tested as a variable in the training, but the variable was removed because of being insignificant to the overall result.

Rank	Variable	Description
1	Light-flavored jet $\|\eta\|$	Absolute value of the pseudorapidity of the light-flavored jet
2	Top quark mass	Invariant mass of the top quark reconstructed from the muon, the missing transverse momentum and the b-tagged jet
3	Dijet mass	Invariant mass of the two selected jets
4	Transverse W boson mass	Transverse mass of the reconstructed W boson
5	Jet p_{T} sum	Scalar sum of the transverse momenta of the two jets
6	$\cos\theta^*$	Cosine of the angle between the muon and the light-flavored jet in the rest frame of the top quark
7	Hardest jet mass	Invariant mass of the jet with the largest transverse momentum
8	ΔR (light-flavored jet, b-tagged jet)	ΔR between the momentum vectors of the light-flavored jet and the b-tagged jet.
9	Light-flavored jet p_{T}	Transverse momentum of the light-flavored jet
10	Light-flavored jet mass	Invariant mass of the light-flavored jet
11	W boson $\|\eta\|$	Absolute value of the pseudorapidity of the reconstructed W boson

6.8. Multivariate Analysis

For the analysis on the full data set of 2015, the maximum-likelihood fit is not performed on the distribution of the absolute pseudorapidity of the light-flavored jet, but instead on the output distribution of an artificial neural network. This allows to combine the separation power of multiple variables, including the previously utilized pseudorapidity distribution of the light-flavored jet, to create a new distribution, which is optimized to separate signal from background events. For this analysis, the NeuroBayes package [256], interfaced with TMVA [86], is used to create, train and evaluate the neural network. The NeuroBayes package provides a powerful tool for the preprocessing of variables and automatic removal of insignificant variables. All relevant input variables are listed in Table 6.5, ranked by their importance in the training of the network. The importance ranking is defined by the loss of separation power of the remaining set of variables if the corresponding variable is removed from the variable list. The missing

transverse momentum variable was automatically removed from the list of input variables, as the additional significance is below a threshold. The most important input variable is the absolute pseudorapidity of the light-flavored jet, followed by the reconstructed top quark mass and the invariant mass of the dijet system. All three variables are also expected to provide good separation power as they represent characteristic features of the signal and background processes. The neural network is trained in the inclusive 2-jets–1-tag region, as an individual training of two different networks in the μ^+ and μ^- region did not improve the results. All relevant input variables are checked for a decent description in simulation and data. The three most important input variables in the 2-jets–1-tag signal region are shown in Fig. 6.17. The distributions for all remaining input variables can be found in Appendix A.2. In principle, the charge of the muon provides strong separation power of signal and background events, but it is not used in the training to prevent any possible bias in the charged-separated measurement. The signal events used in the training of the neural network are t-channel single top quark events obtained from simulation, while as background events top quark pair production and W+jets events from simulation are used. In addition, the background in the training also consists of QCD multijet events obtained from the antiisolated sideband ($I_\mu > 0.12$). All signal and background events are weighted according to their relative contribution, which ensures that the background rejection is not driven by the size of the training data of each individual background. The training is based on a subset of all available signal and background events and these events are discarded from the rest of the analysis to prevent any possible bias of the network. The result of the training for the classifier output is shown in Fig. 6.18. The correlation matrix of all input variables is presented in Fig. 6.19. In Fig. 6.20, the structure of the trained neural network is illustrated with the input layer, hidden layer and output layer, as well as the weights of each interconnection. The neural network training is checked against overtraining by comparing the obtained result on a statistically independent subset of events. The result is shown in Fig. 6.21, where the distributions of signal and background events both behave the same for events of the training and testing subset. Therefore, no overtraining has occurred.
The neural network is then applied to the remaining simulated events and data in the 2-jets–1-tag, the 3-jets–1-tag and 3-jets–2-tags region. In the regions with three jets, the dijet mass variable corresponds to the invariant mass of the two hardest jets. The distribution of all input variables in the 3-jets–1-tag and 3-jets–2-tags region can be found in Appendices A.3 and A.4, respectively.

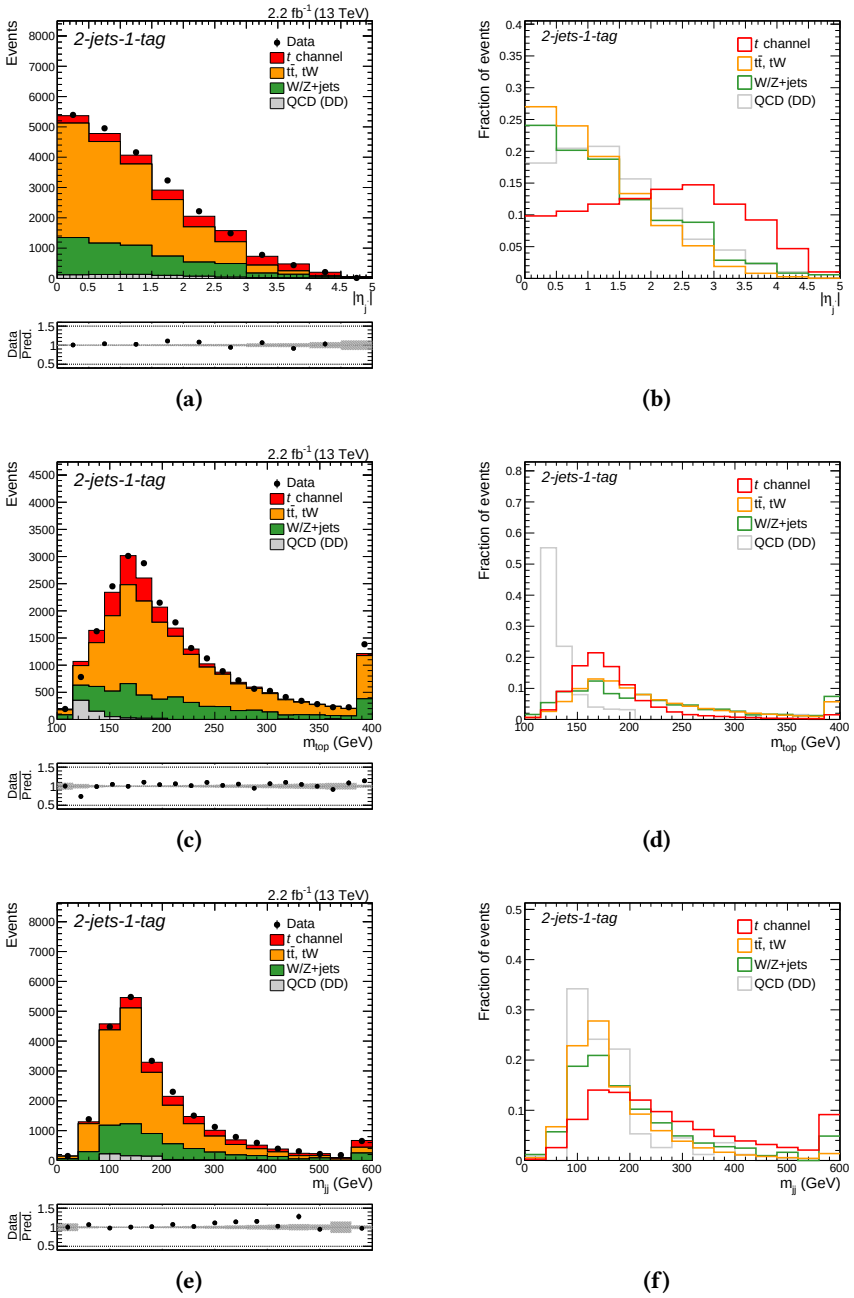


Figure 6.17.: Distributions of the three most important input variables in the 2-jets–1-tag signal region for the training of the neural network. These are the absolute pseudorapidity of the light-flavored jet ((a) and (b)), the reconstructed top quark mass ((c) and (d)) and the invariant mass of the two selected jets ((e) and (f)). All three variables are shown in comparison with data to illustrate proper modeling, where the distributions are scaled to the integrated luminosity, and as a shape comparison between different templates to show their discriminating power. In the latter case, each distribution is scaled to an integral of one. The gray band indicates only the uncertainty due to the limited size of simulation samples.

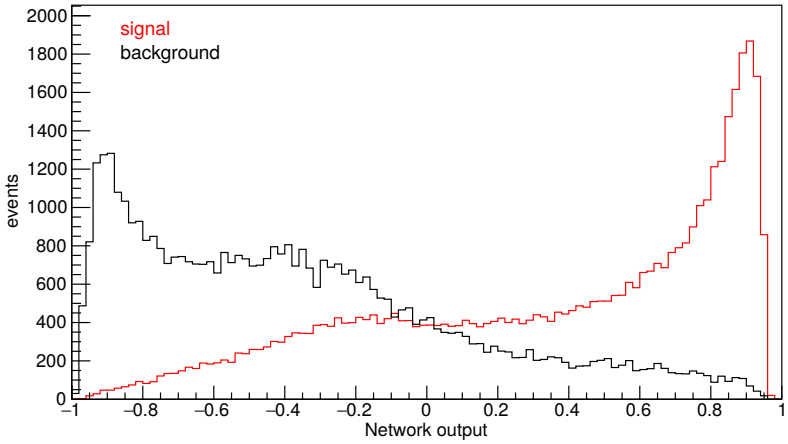


Figure 6.18.: The resulting neural network output distribution obtained from the training of the network. The network has clearly learned to separate the signal and background categories from the characteristics of the t-channel single top quark signal process (red) and the background processes from top quark pair production, W+jets production and QCD multijet events (black).

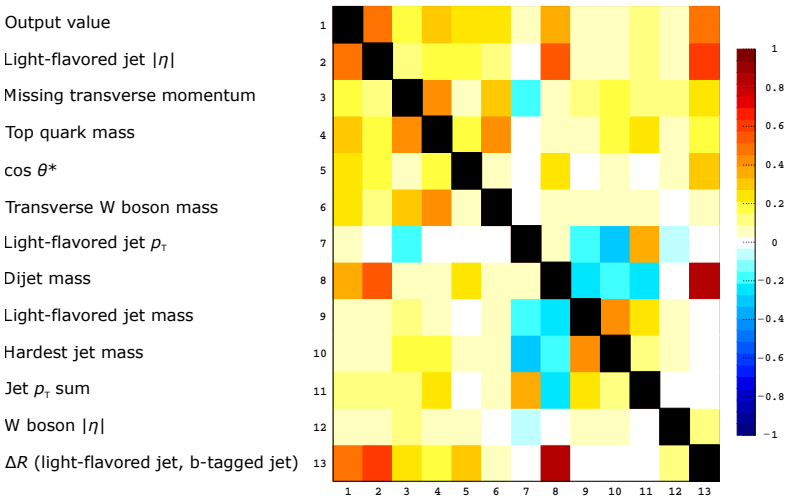


Figure 6.19.: Correlation matrix of input variables used in the training of the neural network. Since there is no strong correlation between the output variable and any single input variable, the performance of the neural network does not rely on any single variable alone.

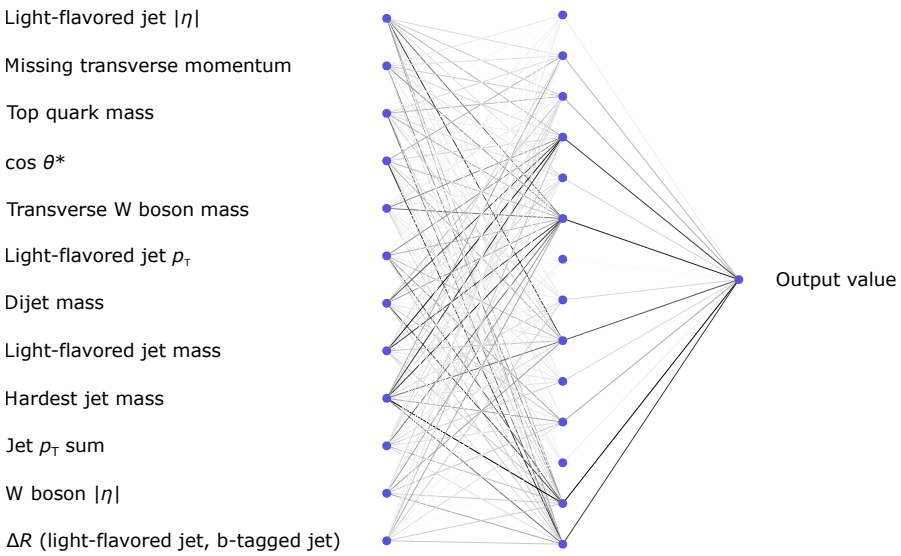


Figure 6.20.: Illustration of the trained neural network. The input nodes are shown with a label for the corresponding input variable, as well as the hidden nodes and the output node. The weights of each intersection, determined from the training, are illustrated with a grayscale, where a darker gray refers to a larger weights and vice versa.

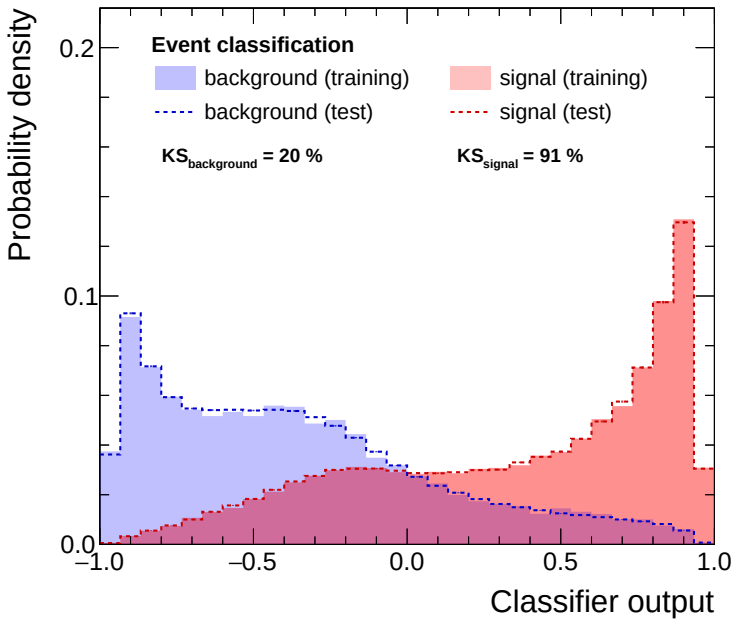


Figure 6.21.: Result of the check performed for possible overtraining of the neural network. The resulting distributions for signal and background events in the training sample are compared to distributions obtained with a statistically independent testing sample. No significant difference in the performance for signal and background events is observed for both samples.

Table 6.6.: Parameters estimated from the simultaneous fit in the 2-jets–1-tag, 3-jets–1-tag and 3-jets–2-tags region. The values refer to the scale factors of each process, or in case of $R_{t\text{-ch.}}$ to the measured value. The uncertainties quoted include the statistical uncertainty and experimental uncertainties, further details are given in Section 6.10.

Process	Scale factor
Signal, t-channel single top quark	1.13 ± 0.08
Top quark background ($t\bar{t}$ and tW)	1.00 ± 0.02
W/Z+jets	1.11 ± 0.09
QCD multijet events	0.86 ± 0.29
$R_{t\text{-ch.}}$	1.81 ± 0.19

6.9. Signal Extraction

The t-channel single top quark production cross section is extracted with a binned maximum-likelihood fit to the output distribution of the neural network classifier. This fit is performed using the THETA framework [102]. The fit is able to scale different templates independently, one template for the signal and three templates for different background categories. The signal here only includes single top quark production, not single top antiquark production. The first background category is the top quark background and consists of associated tW single top quark and top quark pair production. Electroweak background is the second category with W+jets and Z+jets processes. The third background category consists of QCD multijet events. All background templates have an individual nuisance parameter assigned to them, which reflects the uncertainty in the normalization of their prediction. These nuisance parameters have a log-normal distribution with a mean at their initial prediction and a width of 10%, 30% and 50% for the top quark background, the electroweak background and the QCD multijet background, respectively. The fit parameter of each of these four templates is the scale factor of each template with respect to the prediction:

$$S_i = \frac{N_i^{\text{obs}}}{N_i^{\text{pred}}}, \tag{6.11}$$

where N_i^{obs} is the normalization result of the fit and N_i^{pred} is the predicted normalization. An additional fifth fit parameter is the ratio of the scaling factors for the single top quark and single top antiquark production, $R_{t\text{-ch}}$. The fit is performed simultaneously in the 2-jets–1-tag signal region and the two control regions, 3-jets–1-tag and 3-jets–2-tags. Both control regions are able to constrain the normalization of the top quark pair production and help to reduce the uncertainty on the signal normalization. The distributions in each region are split into a μ^+ and a μ^- selection, where the μ^+ region only contains contribution from from single top quark production and the μ^- region only from single top antiquark production. The fit is then performed on these six distributions simultaneously. The results for the fit parameters are shown in Table 6.6. This rather complicated fit setup ensures that for the ratio the correct correlations between single top quark and single top antiquark production are considered, which reduces the impact of systematic uncertainties on the measured ratio. Since the signal scale factor

only corresponds to single top quark production, the measured values for the cross sections of single top antiquark production and the combined single top quark and single top antiquark production are obtained with the scale factors of the signal and the ratio, while taking the correlation matrix of the fit into account. The result of the fit on the distribution of the neural network discriminator in the 2-jets–1-tag region is shown in Fig. 6.22, where the simulation is scaled to the result of the fit. The same distributions for the 3-jets–1-tag and 3-jets–2-tags regions are shown in Fig. 6.23.

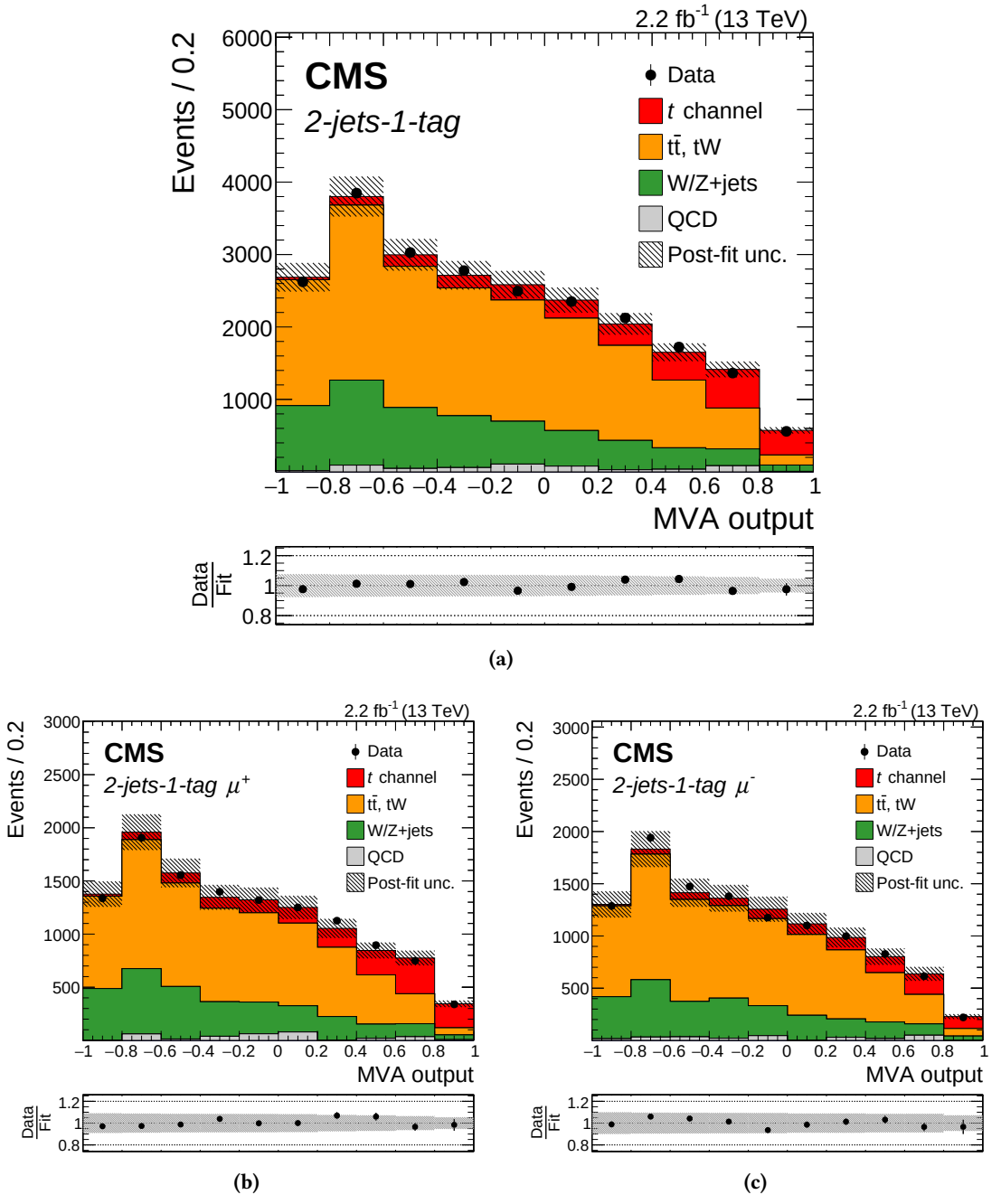


Figure 6.22.: Distributions of the neural network output in the 2-jets–1-tag signal region for the inclusive selection (a), only μ^+ events (b) and only μ^- events (c). The prediction is scaled to the result of the fit. Good agreement between data and simulation is observed.

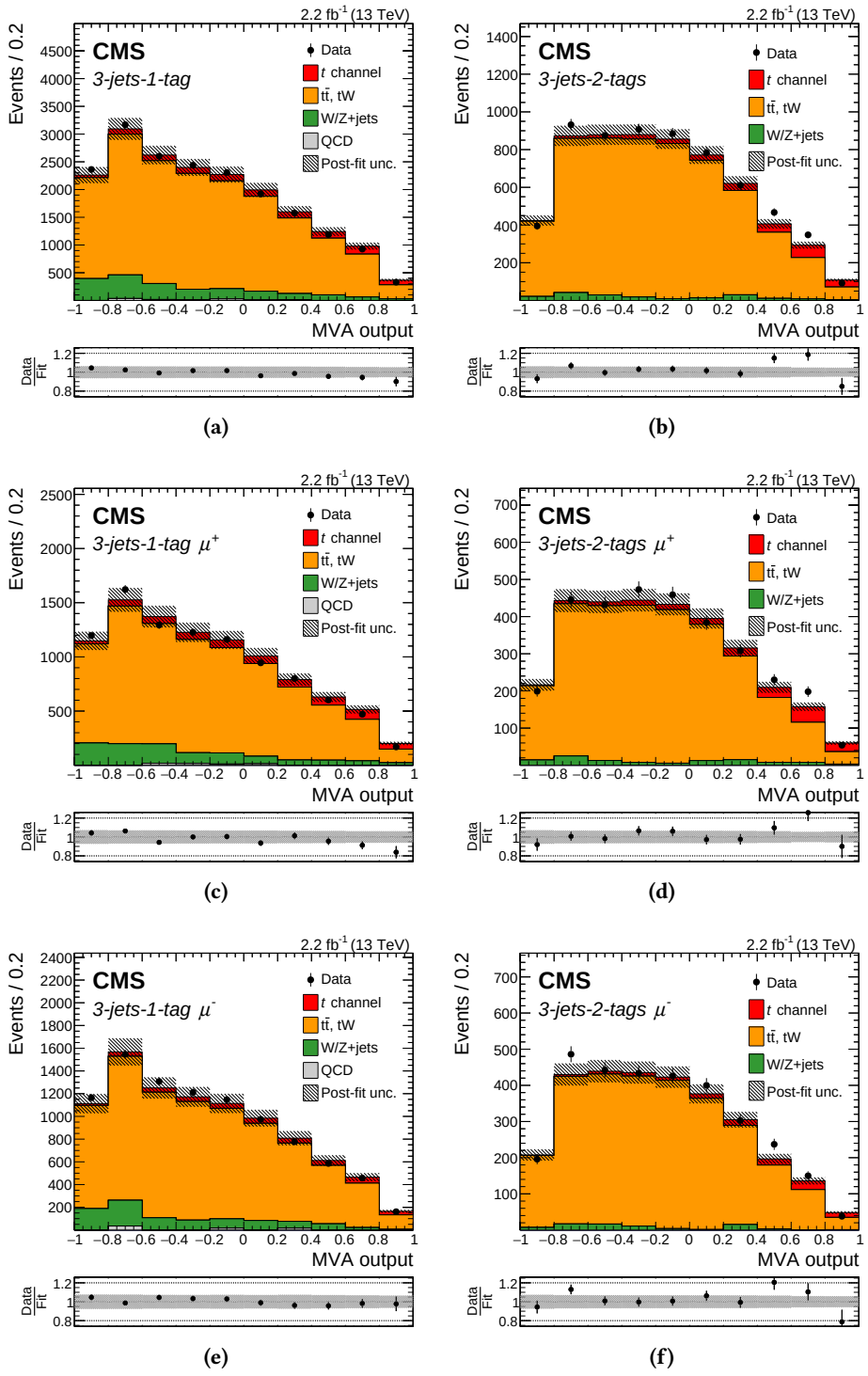


Figure 6.23.: Distributions of the neural network output in the 3-jets–1-tag and 3-jets–2-tags control regions, which have been also used in the fit, for the inclusive selection ((a) and (b)), only μ^+ events ((c) and (d)) and only μ^- events ((e) and (f)). The prediction is scaled to the result of the fit. Good agreement between data and simulation is observed.

6.10. Systematic Uncertainties

Various different sources of systematic uncertainties affect the measurement of the t-channel single top quark production cross section presented in this chapter. These different sources of uncertainty can be grouped into experimental and theoretical uncertainties. Depending on which category an uncertainty source is assigned to, the estimation of their impact to the measurement is treated differently.

6.10.1. Experimental Uncertainties

Each source of the following experimental uncertainty is included in the fit as an additional nuisance parameter:

- **Jet energy scale (JES):** The effect of the jet energy corrections described in Section 4.2.6 is varied according to p_{T}- and η-dependent uncertainties [215,216]. These variations refer to ± 1 standard deviations of the jet energy correction and change the entire jet four-momentum. The effect is also propagated to the calculation of the missing transverse momentum. To also take possible migrations of simulated events into different categories into account, separate simulation samples with the shifted jet energy scale variations are created and the analysis is repeated on these simulations.

- **Jet energy resolution (JER):** The effect of the jet energy resolution correction [216] is also evaluated with the given uncertainties [232] and propagated through the whole analysis chain to take migration effects into account.

- **b tagging:** Corrections applied to the simulation, which account for inefficiencies of the b tagging algorithm (see Section 6.4.3), are affected by various systematic effects. These effects are taken into account by shifting the scale factors according to their uncertainties [253,257]. The b tagging uncertainty is divided into a b tagging efficiency uncertainty and an uncertainty on the mistagging probability.

- **Muon trigger and reconstruction:** The efficiency corrections for the muon trigger and reconstruction described in Section 6.4.2 are varied in a similar way as the b tagging uncertainties to take into account possible systematic effects in their estimation. An additional conservative uncertainty is added on top depending on the jet multiplicity of the corresponding region to reflect the denser environment in events with higher jet multiplicity.

- **Limited size of MC simulation samples:** The fit is performed following the Barlow–Beeston method to take into account the finite size of the simulation samples [258]. The method introduces an additional nuisance parameter for each bin in every region (the so-called "lite" approach), which gets constrained by the number of entries from which the given bin content was created.

The following systematic uncertainty is not included directly in the fit:

- **Luminosity:** The measured value of the integrated luminosity is known with an uncertainty of ±2.3% [242]. Since variations of the overall normalization of the simulation affect the analysis globally, this uncertainty is not included directly in the fit and the effect is quoted independently from the rest of experimental uncertainties. A possible impact on the data-driven QCD estimation is covered by the conservative 50% uncertainty on the result.

The following source of uncertainty is found to be negligible and is not investigated further:

- **Pileup:** The corrections made in Section 6.4.1 to account for a different number of pileup interaction in data and simulations depend on the choice of the minimum-bias cross section. The nominal value of 69 mb is varied by ±5% and the so-obtained shifted corrections are used to estimate this uncertainty. As already mentioned, this uncertainty is found to be negligible and is not considered for the overall uncertainty of the measurement.

6.10.2. Theoretical Uncertainties

Theoretical uncertainties are not included as nuisance parameters in the fit, as this can underestimate their effect. This is especially important at the beginning of Run II since many theoretical uncertainties are new or have been implemented differently compared to Run I. Instead, pseudo-experiments are used to determine their impact on the measurement. The following sources of theoretical uncertainties are considered in the analysis:

- **Modeling of the signal process:** Effects on the modeling of the t-channel single top quark process are estimated by changing the default simulation sample, obtained with MadGraph5_aMC@NLO, with an alternative simulation provided by event generator POWHEG [259]. In addition, modeling effects based on the parton shower and hadronization simulation choice are evaluated by comparing the default PYTHIA setup with the alternative Herwig++, both interfaced with MadGraph5_aMC@NLO.

- **Modeling of the top quark pair background:** The same treatment is applied for estimating the modeling uncertainties of the top quark pair background. The nominal simulation of the POWHEG event generator is compared to an alternative simulation sample generated with MadGraph5_aMC@NLO. Effects on the parton shower modeling are estimated by comparing the nominal choice of POWHEG and PYTHIA with a different simulation from POWHEG interfaced with Herwig++.

- **Modeling of the W+jets background:** During past measurements of the t-channel single top quark production cross section, the W+jets background was underestimated by LO simulation samples, especially for W+c and W+b events, which are the largest fraction of the W+jets background in the signal region. Although this analysis uses NLO simulation samples for the W+jets background and observes no significant mismodeling, the templates for the W+jets background are split into W+b, W+c and W+udsg components based on the parton flavor. An uncertainty of 30% on the initial normalization is assigned independently to the W+b and W+c components to estimate possible effects of heavy-flavor mismodeling in the W+jets background.

- **Factorization and renormalization scales:** The effect on the choice of the factorization and renormalization scales in simulation is estimated either by reweighting events [260] for the single top t-channel signal process, the top quark pair production and electroweak production background, or with dedicated simulation samples for the associated tW single top quark background. Both scales are varied independently with double and half of the nominal value and the envelope of all possible combinations is chosen as the resulting uncertainty. The scale uncertainty is evaluated for each process independently.
- **PDF and α_s:** The uncertainty in the choice of the PDF is evaluated by reweighting the templates of the fit with 100 different eigenvectors of the NNPDF set [40, 41], which is used as a default in the analysis. In addition, two different templates with an alternative choice of α_s are used. The final uncertainty is obtained by creating the envelope of the 102 different variations.
- **Modeling of the top quark p_T:** Measurements of the differential top quark pair production cross section observe a significant difference in the description of the top quark transverse momentum, where the observed spectrum is shifted to lower values of the transverse momentum compared to the prediction [261]. This effect can be corrected by reweighting events based on the top quark transverse momentum. The correction is not applied by default, but considered as an additional source of systematic uncertainty when the corrections are applied. Only simulated events of top quark pair production are reweighted with this procedure.

6.10.3. Impact of Systematic Uncertainties

Depending on the categorization of each uncertainty source, the effect is evaluated differently. As the experimental uncertainties are all included in the fit as nuisance parameters, their individual effect is obtained by repeating the fit while fixing the corresponding nuisance parameter to its best fit value. In case of the uncertainty source for the limited size of the simulation samples, the impact is evaluated by comparing the result to a result without considering this effect at all. The complete list of evaluated experimental uncertainties for the production of single top quarks, single top antiquarks, the combined single top quark and single top antiquark production and the ratio of single top quark and single top antiquark production are shown in Table 6.7. For the theoretical uncertainties, the impact is estimated by generating pseudo-data for the nominal templates and performing the fit with the systematically shifted templates on the generated pseudo-data. The estimated theoretical uncertainties are listed in Table 6.8, together with the combined experimental uncertainties, the statistical uncertainty and the total uncertainty. The statistical uncertainty, i. e., the uncertainty only due to the size of the data set on which the fit is performed, is obtained by fixing all nuisance parameters in the fit to their best fit values. The summed impact of profiled experimental uncertainties is calculated by subtracting the statistical uncertainty quadratically from the complete uncertainty of the fit. Asymmetric theoretical uncertainties are added together by treating them each as standard deviations of a Gaussian function [262].
For the experimental uncertainties, the largest impacts on the t-channel single top quark production cross section stem from the limited size of simulation samples and the jet energy scale.

Table 6.7.: Impact of experimental uncertainties on the measurement. The respective source of systematic uncertainty is either evaluated by fixing the respective nuisance parameter in the fit to its best fit value or, in case of the MC sample size, by comparing the result to the result obtained by neglecting this effect.

Uncertainty source	$\Delta\sigma_{t\text{-ch.,t}+\bar{\text{t}}}/\sigma^{\text{obs}}_{t\text{-ch.,t}+\bar{\text{t}}}$	$\Delta\sigma_{t\text{-ch.,t}}/\sigma^{\text{obs}}_{t\text{-ch.,t}}$	$\Delta\sigma_{t\text{-ch.,}\bar{\text{t}}}/\sigma^{\text{obs}}_{t\text{-ch.,}\bar{\text{t}}}$	$\Delta R_{t\text{-ch.}}/R_{t\text{-ch.}}$
JES	±4.1%	±4.7%	±3.5%	±2.1%
JER	±1.7%	±1.2%	±2.4%	±0.6%
b tagging efficiency	±1.9%	±2.0%	±1.8%	±1.4%
Mistag probability	±0.9%	±0.6%	±0.8%	±0.5%
Muon reco./trigger	±2.0%	±2.3%	±1.9%	±1.8%
MC samples size	±3.4%	±4.1%	±3.8%	±3.2%

The effect of the latter source can be significantly reduced for the ratio measurement. Larger impacts are estimated for theoretical uncertainties, in particular the signal process modeling, the choice of the factorization and renormalization scales of the signal process and the modeling of the top quark pair production background.

6.11. Results

The result of the fit yields a cross section of single top quark production in the t channel of

$$\begin{aligned}\sigma_{t-\text{ch.,t}} &= 154 \ \pm 8 \ (\text{stat}) \ \pm 9 \ (\text{exp}) \ \pm 19 \ (\text{theo}) \ \pm 4 \ (\text{lumi}) \ \text{pb} \\ &= 154 \ \pm 22 \ \text{pb}, \end{aligned} \tag{6.12}$$

where the terms stat, exp, theo and lumi correspond to the statistical, experimental, theoretical and luminosity uncertainty, respectively. The ratio of single top quark and antiquark production is measured to be

$$R_{t-\text{ch.}} = 1.81 \ \pm 0.18 \ (\text{stat}) \ \pm 0.15 \ (\text{syst}). \tag{6.13}$$

A comparison of the measured ratio with the prediction of different PDF sets is provided in Fig. 6.24. From both measured values, the single top antiquark cross section is calculated utilizing the covariance matrix of the fit:

$$\begin{aligned}\sigma_{t-\text{ch.,}\bar{\text{t}}} &= 85 \ \pm 10 \ (\text{stat}) \ \pm 4 \ (\text{exp}) \ \pm 11 \ (\text{theo}) \ \pm 2 \ (\text{lumi}) \ \text{pb} \\ &= 85 \ \pm 16 \ \text{pb}. \end{aligned} \tag{6.14}$$

Similarly, the inclusive cross section of single top quark and antiquark production is found to be:

$$\begin{aligned}\sigma_{t-\text{ch.,t}+\bar{\text{t}}} &= 238 \ \pm 13 \ (\text{stat}) \ \pm 12 \ (\text{exp}) \ \pm 26 \ (\text{theo}) \ \pm 5 \ (\text{lumi}) \ \text{pb} \\ &= 238 \ \pm 32 \ \text{pb}. \end{aligned} \tag{6.15}$$

Table 6.8.: Summary of all uncertainties affecting the measurement of the t-channel single top quark production. The systematic uncertainties are grouped into different categories. The different sources of theoretical uncertainty are evaluated independently and added following the proposal in Ref. [262] for asymmetric uncertainties.

Uncertainty source	$\Delta\sigma_{t\text{-ch.,t+}\bar{\text{t}}}/\sigma^{\text{obs}}_{t\text{-ch.,t+}\bar{\text{t}}}$	$\Delta\sigma_{t\text{-ch.,t}}/\sigma^{\text{obs}}_{t\text{-ch.,t}}$	$\Delta\sigma_{t\text{-ch.,}\bar{\text{t}}}/\sigma^{\text{obs}}_{t\text{-ch.,}\bar{\text{t}}}$	$\Delta R_{t\text{-ch.}}/R_{t\text{-ch.}}$
Statistical uncertainty	±5.5%	±5.3%	±11.5%	±9.7%
Profiled exp. uncertainty	±5.2%	±5.7%	±4.9%	±3.3%
Total fit uncertainty	±7.6%	±7.8%	±12.5%	±10.3%
Integrated luminosity	±2.3%	±2.3%	±2.3%	—
Signal modeling	±6.9%	±8.2%	±8.5%	±5.3%
$\text{t}\bar{\text{t}}$ modeling	±3.9%	±4.3%	±4.5%	±4.0%
W+jets modeling	−1.8/+2.1%	−1.6/+2.3%	−2.5/+2.3%	−1.7/+2.0%
$\mu_{\text{R}}/\mu_{\text{F}}$ scale t-channel	−4.6/+6.1%	−5.7/+5.2%	−7.2/+5.1%	−0.7/+1.2%
$\mu_{\text{R}}/\mu_{\text{F}}$ scale $\text{t}\bar{\text{t}}$	−3.5/+2.9%	−3.5/+4.1%	−4.7/+3.1%	−1.1/+1.0%
$\mu_{\text{R}}/\mu_{\text{F}}$ scale tW	−0.3/+0.5%	−0.6/+0.8%	−1.1/+0.7%	−0.2/+0.1%
$\mu_{\text{R}}/\mu_{\text{F}}$ scale W+jets	−2.9/+3.7%	−3.5/+3.0%	−4.9/+3.8%	−1.2/+0.9%
PDF uncertainty	−1.5/+1.9%	−2.1/+1.6%	−1.8/+2.1%	−2.2/+2.5%
Top quark p_{T} modeling	±0.1%	±0.2%	±0.2%	±0.1%
Total theory uncertainty	−10.7/+11.1%	−12.2/+12.1%	−13.6/+12.9%	±7.5%
Total uncertainty	−13.4/+13.7%	±14.7%	−18.7/+18.2%	±12.7%

A comparison of the measured inclusive cross section at 13 TeV with the prediction is shown in Fig. 6.25. From this inclusive cross section, the CKM matrix element V_{tb} is determined:

$$|f_{\text{LV}} \cdot V_{\text{tb}}| = 1.05 \ \pm 0.07 \ (\text{exp}) \ \pm 0.02 \ (\text{theo}), \tag{6.16}$$

where the experimental uncertainty refers to the combined uncertainty of the measured inclusive single top quark production cross section and the theoretical uncertainty is the uncertainty on the SM prediction of V_{tb}. The form factor f_{LV} takes possible anomalous Wtb couplings into account [243] and the value is 1 in the SM. All measured observables are in agreement with the predictions of the SM, given the uncertainties of the measurement and prediction.

6.12. Summary and Outlook

In this chapter, the measurement of the single top quark production cross section in the t channel with the CMS experiment has been presented. The analysis utilizes the full data set of 2015, corresponding to an integrated luminosity of 2.2 fb^{-1}. In addition, a preliminary result on the first data of 2015 has been presented, which has been published by the CMS Collaboration as a *Physics Analysis Summary* [255]. The analysis on the full data set of 2015 has been published in *Physics Letters B* [263]. A compatible result with a similar analysis strategy has been obtained

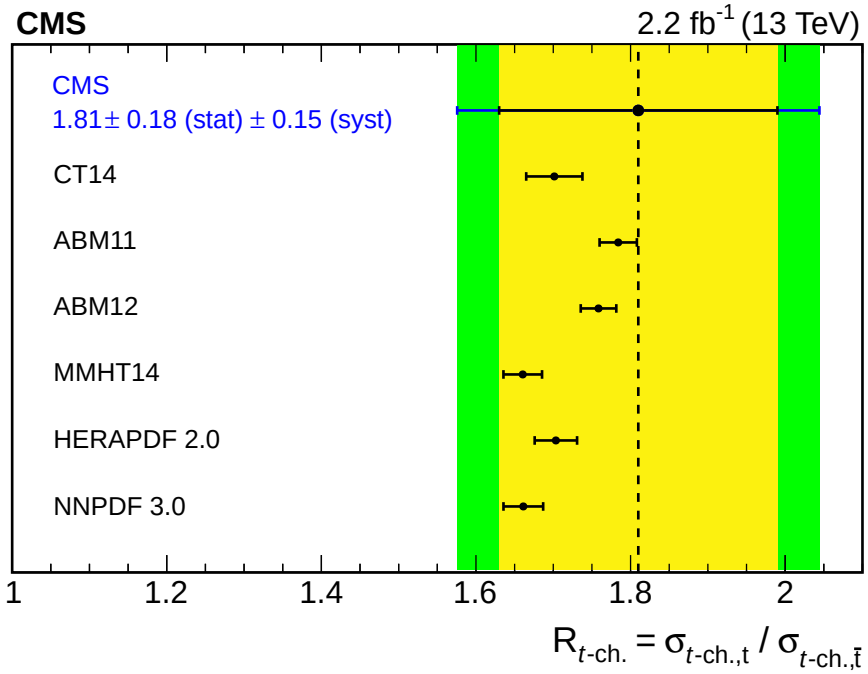


Figure 6.24.: Comparison of the measured ratio for single top quark and single top antiquark production in the t channel with the prediction of different PDF sets. All predictions are in agreement with the measured ratio and the corresponding uncertainty of the measurement. Uncertainties on the prediction include the statistical uncertainty, the uncertainty of the factorization and renormalization scales, as well as an uncertainty on the top quark mass of ±1 GeV.

by the ATLAS Collaboration [264].
The increased amount of proton-proton collision data taken in the years after 2015 allows to reduce the statistical uncertainty of this measurement by a significant amount. This especially helps to reduce the overall uncertainty of the ratio measurement, as there the statistical uncertainty is dominant for the measurement presented in this chapter. A result for the ratio of single top quark and antiquark production in the t channel with the full data set of 2016, corresponding to an integrated luminosity of 35.9 fb^{-1}, is shown in Fig. 6.26, following an analysis strategy similar to the one presented in this thesis [265]. Most PDF sets are still compatible with the measured value of $R_{t-\text{ch.}}$, only the ABMP16 PDF set shows significant deviations from the measurement.
With the data available now, the single top quark production in the t channel can be used to study different properties of the top quark. For instance, single top quark production can be used to measure the top quark mass, which provides an orthogonal measurement to the standard top quark mass measurements from top quark pair production [266]. Single top quarks are also well-suited to probe the Wtb vertex structure to search for anomalous couplings and to measure the polarization of top quarks [243, 267, 268], which has already been done by the ATLAS and CMS Collaborations during Run I of the LHC [269–272]. Due to the missing direct observation of dark matter candidates, associated production of single top quarks with dark

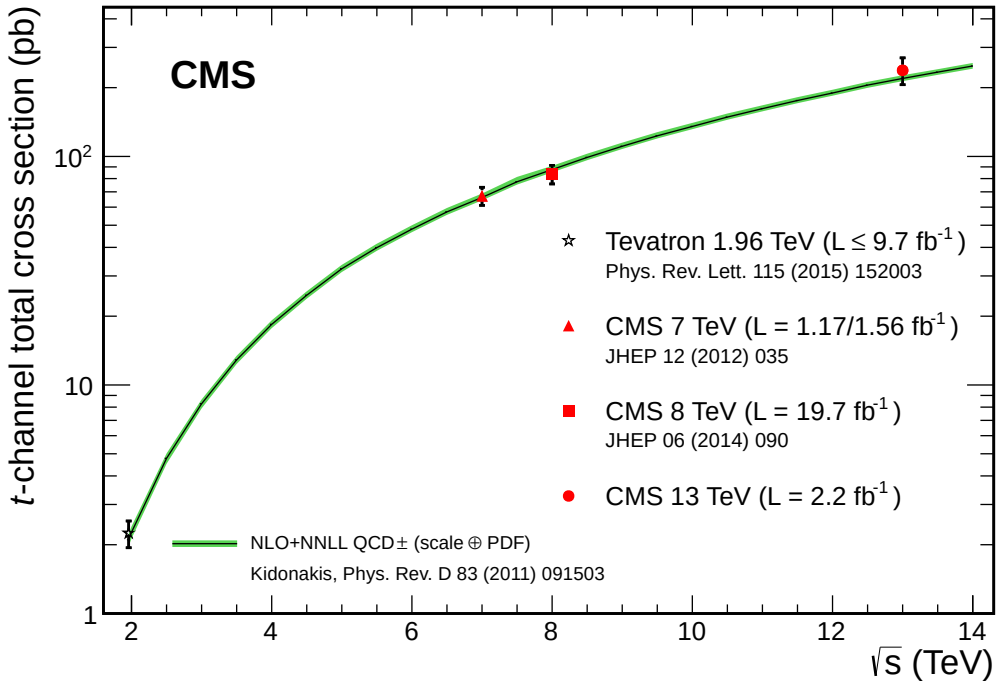


Figure 6.25.: Comparison of the measured t-channel single top quark production cross section with the prediction from the theory. The measurements of the CMS Collaboration at a center-of-mass energy of 7 and 8 TeV are also shown, along with combined result at the Tevatron. All measurements are in agreement with the prediction.

matter candidates has also come into focus in the recent years [273–277].
To improve the uncertainties of the measurement, different approaches can be used. For instance, a measurement of the cross section in a fiducial volume, i. e., close to the kinematic region of the detector, is not affected by extrapolating the measurement to the whole kinematic regime. By this approach, systematic uncertainties because of this extrapolation can be avoided. The largest systematic uncertainties of the measurement in this chapter can be contributed to the simulation and modeling of different processes. Measurements of differential cross sections, where the cross section is provided dependent on kinematic variables, can be used to tune MC event generators to provide more accurate predictions. This is achieved by comparing the generator-level predictions with the differential cross section measurement, which has been unfolded to parton or particle level to correct for detector resolution effects.
Since the LHC not only accelerates protons, but also heavy nuclei, it is also interesting to search for single top quark production in heavy ion collisions [278]. The observation of top quark pair production in proton-lead collisions has been published by the CMS Collaboration recently [279].

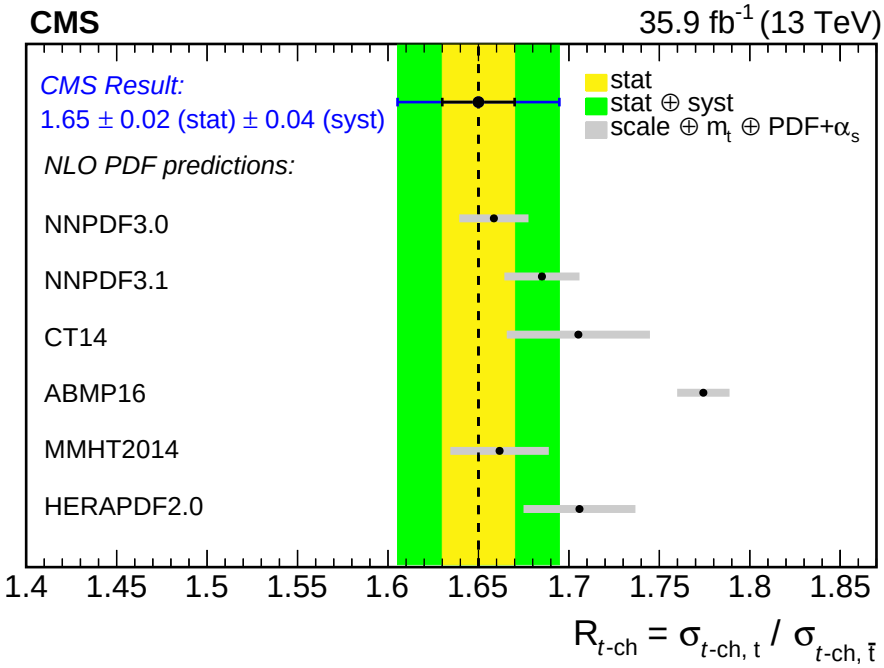


Figure 6.26.: Preliminary result for the ratio of single top quark and antiquark production with the data set of 2016. Compared to the result presented in this chapter with the data set of 2015, the statistical uncertainties can be reduced significantly.

7. Search for s-Channel Single Top Quark Production at $\sqrt{s} = 13$ TeV

After establishing t-channel single top quark production at a center-of-mass energy of 13 TeV in Chapter 6 and the observation of associated tW production by the ATLAS and CMS Collaborations [280–282], the last of the three main production modes missing is the s-channel single top quark production.
The first evidence for single top quark production in the s channel, without including the t channel, was found at the Tevatron Collider by the DØ Collaboration in 2013 [283]. An observation was finally claimed by a combined measurement of the CDF and DØ Collaborations in the following year [284]. At the LHC, the production is less likely compared to the Tevatron due to the missing valence antiquarks from the antiproton. Instead, the necessary antiquark has to be generated from sea quarks. The ATLAS and CMS Collaborations have searched for this production mode in Run I at a center-of-mass energy of 7 and 8 TeV [285, 286], but only evidence has been found so far [287]. The prospects for this channel at Run II with a center-of-mass energy of 13 TeV are difficult, as the production of top quark pair events, the dominant background process, increases faster than the signal process when the center-of-mass energy is increased. Nevertheless, the outstanding performance of the LHC so far at Run II has provided enough proton-proton collision data to yield at least the same sensitivity as in Run I. The analyzed data set in this chapter corresponds to an integrated luminosity of $35.9\,\text{fb}^{-1}$ for data taken in 2016 and $41.5\,\text{fb}^{-1}$ for data taken in 2017.
The search for s-channel single top quark production with the combined 2016 and 2017 data is described in detail in this chapter. The general analysis strategy is outlined in the first section. In the second section, details about the event topology for signal and background processes are presented. The third section is dedicated to the selection of events to increase the signal-to-background ratio. A discussion about the simulation of events is provided in the fourth section, along with several corrections applied to simulated events. A special treatment for the QCD multijet event background is provided in the fifth section. The sixth section explains the reconstruction of the top quark from the various final-state objects. An event classification to separate signal from background events is given in the seventh section. Systematic uncertainties affecting the sensitivity of the search are explained in the eighth section of this chapter. The final results are provided in the ninth section, followed by a summary and an outlook in the last section.

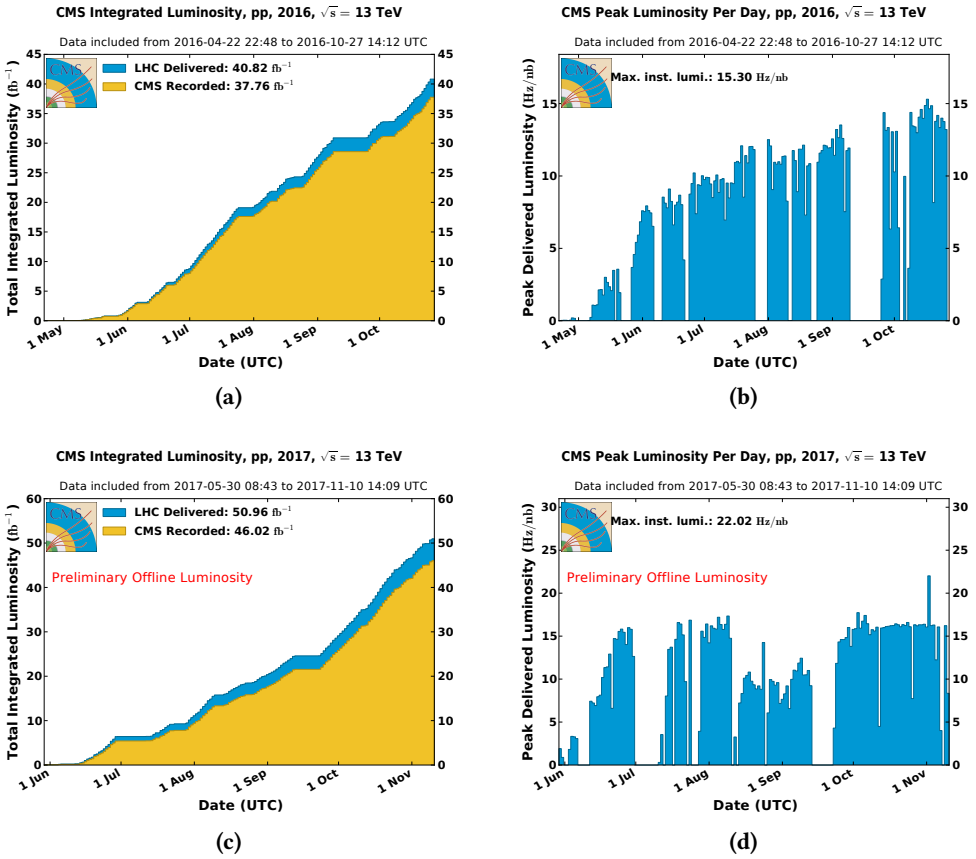


Figure 7.1.: The cumulative integrated luminosity of 2016 (a) and 2017 (c), together with the corresponding instantaneous luminosity of 2016 (b) and 2017 (d), recorded by the CMS experiment [125]. The certified data for this analysis yields an integrated luminosity of 35.9 fb^{-1} and 41.5 fb^{-1} for 2016 and 2017, respectively.

7.1. Analysis Strategy

This analysis aims at observing single top quark production in the s channel, a yet unobserved production mode of single top quarks at the LHC. To achieve this goal, the combined data set of 2016 and 2017 at a center-of-mass energy of 13 TeV is used, corresponding to an integrated luminosity of 35.9 fb^{-1} and 41.5 fb^{-1}, respectively. The integrated luminosities over time, as well as the peak luminosities per day, for both years are shown in Fig. 7.1. Notably, the LHC reached already more than twice its design luminosity of $2 \cdot 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ ($= 20\,\text{Hz}\,\text{nb}^{-1}$) by the end of 2017. Data taken in 2015, which were used for the t-channel single top quark analysis in Chapter 6, is not added to the analyzed data set, as the small integrated luminosity of 2.2 fb^{-1} would only add a small number of additional events, compared to the significantly larger data sets of 2016 and 2017. The analyzed data set therefore includes more than 97% of the certified data at a center-of-mass energy of 13 TeV recorded by the CMS detector.
The method already established in the two previous chapters to define different phase-space

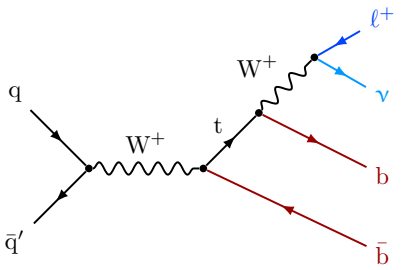


Figure 7.2.: Leading order Feynman diagram of single top quark production in the *s* channel and the subsequent decays of final-state particles, i. e., the top quark decaying into a bottom quark and a W boson, which then further decays into a charged lepton and the corresponding neutrino.

regions dependent on the number of jets and the number of b-tagged jets is also used in the search for *s*-channel single top quark production to increase the signal-to-background ratio. To separate signal and background events in the signal region, multivariate techniques are used, in particular a DNN. The employment of a DNN showed significant improvements in the separation power compared to a BDT for the analysis and the involved processes [288]. The non-negligible background from QCD multijet events is estimated with the same data-driven technique as in Section 6.5. A maximum-likelihood fit of the prediction to the observed data is performed on the resulting distributions of the DNN in different regions to compare the background-only with the signal-plus-background hypothesis and, in case of a significant excess, to measure the cross section of single top quark production in the *s* channel.

7.2. Event Topology

With a predicted production cross section of only 10.32 pb for proton-proton collisions at a center-of-mass energy of 13 TeV, *s*-channel single top quark production is more than 20 times less likely than the *t*-channel production mode, which was the subject of study in Chapter 6. It is therefore of great importance to find a specific phase-space region to reject as many background events as possible, while most signal events fall in this specific region. This decision is made based on the final-state properties of the *s*-channel single top quark signal process, as well as the relevant background processes, which will be discussed in the following.

7.2.1. Signal Process

The signal process, whose observation is the goal of this analysis, is *s*-channel single top quark production. The leading order Feynman diagram of the process is shown in Fig. 7.2, along with the decay of the top quark and subsequent W boson decay. Only leptonically decaying W bosons are considered for this analysis, as the hadronic decay mode would result in a final state that only consists of multiple jets, for which the contamination of background from QCD multijet events would be too high. The final state of the signal process contains one bottom quark from the top quark production vertex and the corresponding charge-conjugated bottom quark from the top quark decay vertex, as well as a charged lepton and a neutrino from the leptonically

decaying W boson. On the detector level, this results in two jets, which can be identified as stemming from bottom quarks, a charged lepton and missing transverse momentum.

7.2.2. Background Processes

The background processes relevant for this analysis are described in the following lines. Exemplary Feynman diagrams for all processes are shown in Fig. 7.3.

t-Channel Single Top Quark Production

The single top quark production mode with the highest cross section, the t channel, is an important background in this analysis. The relevant Feynman diagram is shown in Fig. 7.3(a) for the modeling in the 4FS. Although the t-channel production mode features the additional light-flavored quark in forward direction, the final-state is quite similar. In cases where the jet stemming from the light-flavored quark is not detected, but instead the b jet stemming from the second bottom quark of the initial gluon splitting, the objects in the detector are identical to that of the s-channel process. In cases where all final-state objects can be identified, contributions from t-channel single top quark production can be suppressed by rejecting events with more than two jets in the detector.

Associated tW Single Top Quark Production

Non-negligible background contributions also arise from the third of the three main single top quark production modes, associated tW production (Fig 7.3(b)). Since this production mode has two W bosons in the final state, different combinations of jets, charged leptons and neutrinos can occur, depending on the specific W boson decay. Closest to the final state of s-channel single top quark production with a leptonic W boson decay is the final state of associated tW production where one W boson decays leptonically and one hadronically. In this case, the final-state objects in the detector are three jets, one of which stems from a bottom quark, a charged lepton and missing transverse momentum. If one jet of the hadronically decaying W boson fails the detector acceptance and the other jet is falsely identified as originating from a bottom quark, the same objects will be found as for the s-channel production mode.

Top Quark Pair Production

As in the previous chapter, the most important background contribution arises from top quark pair production because of the high cross section of 832 pb and the two W bosons, which can decay either leptonically or hadronically. The prominent decay mode is the semileptonic top quark pair production, for which a Feynman diagram is shown in Fig. 7.3(c). Here, the final state of the top quark pair production is identical to that of the signal process, but with two additional jets from the hadronically decaying W boson. Vetoing events with more than two jets can help to suppress this background component. For the dileptonic decay mode, the number of jets and b-tagged jets is identical to the signal process. Here, events can be rejected by enforcing exactly one charged lepton to be present in the event. Contributions from fully hadronic top quark pair production can be suppressed entirely by the two selection criteria above.

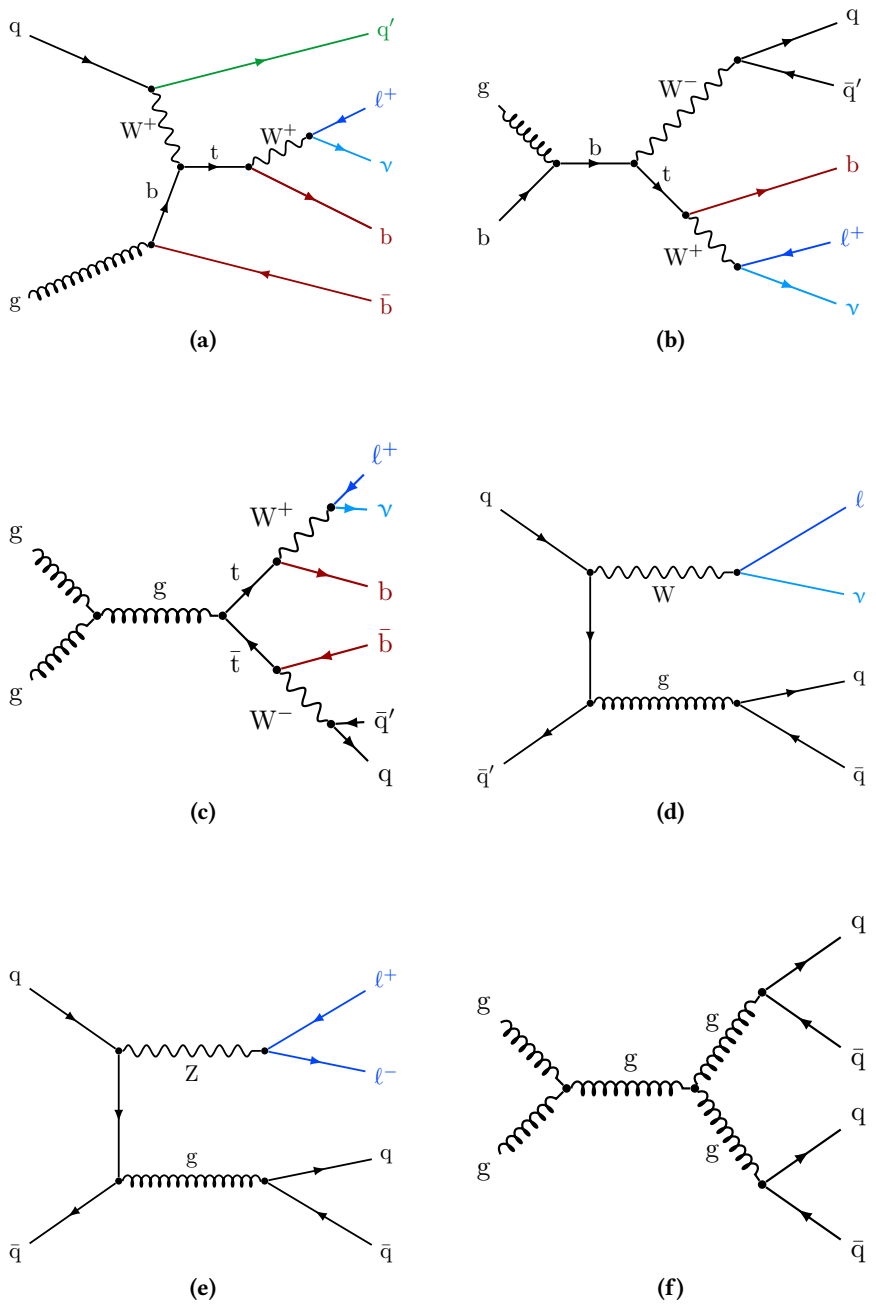


Figure 7.3.: All relevant background processes considered in the search for single top quark production in the *s* channel. Since the *s*-channel production has similar features compared to that of the more likely *t*-channel (a) and associated tW (b) single top quark production modes, both production modes are considered as background. The by far dominating background is top quark pair production (c). Other background contributions arise from the production of W bosons (d) and Z bosons (e) in association with jets. QCD multijet events (f) contribute also to the background composition, as even tiny fractions of misreconstructed events have an impact due to the high cross section.

Production of Electroweak Bosons in Association with Jets

Contributions from the associated production of vector bosons and jets, i. e., W+jets and Z+jets, are also relevant for this analysis because of their high production cross section. An exemplary Feynman diagram for W+jets is provided in Fig. 7.3(d) and for Z+jets production in Fig. 7.3(e). To mimic the final state of the signal process, the vector bosons have to decay leptonically, i. e., into a charged lepton and the corresponding neutrino for W bosons or into two charged leptons for Z bosons. Since the latter can be rejected more easily by vetoing events with additional leptons, the main contribution from this background arises from W+jets production.

QCD Multijet Events

QCD multijet events significantly contribute to the total background composition, despite missing a lepton with sufficient transverse momentum in the final state (Fig. 7.3(f)). However, the cross section of QCD multijet events is several orders of magnitude higher than the cross section of any other process considered in the analysis, thus rare cases of events with misreconstructed leptons or leptonic meson decays inside a jet can lead to events passing the selection criteria.

7.3. Event Selection

The purpose of the event selection is to define a phase-space region in which the signal-to-background ratio can be enhanced with respect to the whole kinematic regime, but also to ensure proper modeling of selected events in simulation. Based on this selection process, a main signal region is defined and several control regions, which provide additional constraints for certain background processes.
To be considered in this analysis, every event must pass at least one HLT decision for muons or electrons. Each HLT path has a threshold for the transverse momentum of the corresponding lepton, which is kept as low as possible to ensure that a high efficiency for the selection of leptons can be achieved. The individual thresholds are driven by the rate of the given trigger and therefore by the instantaneous luminosity, resulting in differences for the chosen HLT paths for the 2016 and 2017 data analysis. A list of the employed HLT paths in this analysis is provided in Table 7.1. To consider turn-on effects of the chosen HLT paths, the corresponding reconstructed lepton in each event must pass a slightly higher criterion on the transverse momentum. For muons, this threshold is 27 GeV for data taken in 2016 and 30 GeV for data taken in 2017. In case of electrons, the transverse momentum must be at least 35 GeV and the absolute pseudorapidity at most 2.1, independent of the year. To select only events with well-reconstructed leptons, the muon or electron must satisfy certain quality criteria (tight ID), as defined in Section 5.1.2 and Section 5.1.3, respectively. In addition to passing the tight ID, each muon must have an isolation of at most 0.06. Each electron must pass the selection criteria of the tight ID, including the hereby defined electron isolation criterion. Reconstructed electrons with a corresponding supercluster located in the transition region between the ECAL barrel and endcap ($1.442 < |\eta_{\mathrm{sc}}| < 1.566$) are excluded from the analysis. With the help of these selection criteria, it is already possible to reject most contributions from QCD multijet events, which mostly contain nonprompt leptons. Events with more than one lepton are rejected if

Table 7.1.: Complete list of high-level-trigger paths used in the analysis for different lepton flavors and years of data taking. Because of the higher instantaneous luminosity in 2017 compared to 2016, the threshold for the transverse momentum of the lepton was raised. In the electron channel, this effect is countered by adding the requirement of a central jet with a transverse momentum of at least 35 GeV.

Lepton flavor	Year	Trigger
μ	2016	HLT_IsoMu24_v*
		HLT_IsoTkMu24_v*
	2017	HLT_IsoMu27_v*
e	2016	HLT_Ele32_eta2p1_WPTight_Gsf_v*
	2017	HLT_Ele30_eta2p1_WPTight_Gsf_CentralPFJet35_EleCleaned_v*

at least one of the additional leptons passes the loose ID of the corresponding lepton flavor. This veto significantly reduces the contribution from background processes with leptonically decaying vector bosons, such as Z+jets and diboson production, from which the latter can be even neglected entirely. A jet must have at least a transverse momentum of 40 GeV to be considered for the analysis. In addition, a jet is not considered if it has a distance of $\Delta R < 0.4$ in the η-ϕ plane to the selected lepton in the event or if the jet fails the PF jet ID, as defined in Section 5.1.4. Jets considered for b tagging must have an absolute pseudorapidity of at most 2.4, otherwise values up to 4.7 are allowed. To decide whether a reconstructed jet is stemming from a bottom quark in the final state, the DeepCSV b tagging algorithm is employed for data taken in 2016 and 2017 and the medium working point is used. Compared to the tight working point, which has a lower mistagging rate, but also a lower efficiency for b jets, the medium working point results in a higher significance of the signal process. The fraction of QCD multijet events can be further reduced by imposing a requirement on each event to contain a reconstructed transverse W boson mass of at least 50 GeV or a missing transverse momentum of at least 30 GeV for events with muons or electrons, respectively.

The definition of the main signal region and additional control regions is entirely based on the number of jets and the number of b-tagged jets in an event. The selection criteria on the HLT paths and leptons are shared among all regions. As the final state of the *s*-channel single top quark production has exactly two jets, both of which are originating from a bottom quark, the main signal region of the analysis is the 2-jets–2-tags region. As the background from top quark pair production is the main background contribution in the signal region, three additional control regions are defined to validate the modeling and to constrain this background process. These three regions are the 3-jets–1-tag, 3-jets–2-tags and 2-jets–1-tag regions. The 2-jets–1-tag region also contains a higher number of *t*-channel single top quark background events compared to the signal region, which also helps to better constrain this background component. To validate the QCD multijet event estimation and modeling (see Section 7.5) in the former regions, a 2-jets–0-tag region is defined whose sole purpose is to test the procedure in a phase-space region with an increased number of events. The selection criteria for each of the aforementioned regions are summarized in Table 7.2. The number of selected events in the four main regions of interest, i. e.,

Table 7.2.: Complete list of event selection criteria for all defined categories. The main signal region of the analysis is the 2-jets–2-tags region. All other defined regions serve the purpose of validating the modeling of variables or constraining different background components.

	2-jets–0-tag (2j0t)	2-jets–1-tag (2j1t)	2-jets–2-tags (2j2t)	3-jets–1-tag (3j1t)	3-jets–2-tags (3j2t)
Trigger match	see Table 7.1	see Table 7.1	see Table 7.1	see Table 7.1	see Table 7.1
Number of tight leptons	1	1	1	1	1
Number of additional loose leptons	0	0	0	0	0
Number of selected jets	2	2	2	3	3
Number of b-tagged jets	0	1	2	1	2
Transverse W boson mass (only μ)	> 50 GeV	> 50 GeV	> 50 GeV	> 50 GeV	> 50 GeV
Missing transverse momentum (only e)	> 30 GeV	> 30 GeV	> 30 GeV	> 30 GeV	> 30 GeV

the 2–jets-2-tags signal region and three additional control regions, are listed in Table 7.3 and Table 7.4, respectively. In the following, two improvements of the event selection procedure are outlined, which improve the selection with respect to the analysis described in the previous chapter.

7.3.1. Electron Cross Trigger

The aforementioned high-level-trigger in the electron channel for the 2017 data-taking period, `HLT_Ele30_eta2p1_WPTight_Gsf_CentralPFJet35_EleCleaned_v*`, was specifically designed for analyses of top quark production in general [289]. When comparing the instantaneous luminosities of 2016 and 2017 (Figs. 7.1(b) and 7.1(d)), the collision rate in 2017 is higher and requires a higher threshold for the transverse momentum of the electron used in the trigger compared to the threshold in 2016. Higher thresholds not only result in fever recorded events overall, but will also reduce the signal-to-background ratio for single top quark analyses. In order to retain a similar number of recorded events and signal-to-background ratio in 2017 with respect to 2016, the already existing single-electron trigger in 2016 was extended with a requirement for a central PF jet with a transverse momentum of at least 35 GeV. With the help of this so-called electron cross trigger, the threshold for the transverse momentum of the electron was even lowered down to 30 GeV.

7.3.2. b Tagging Improvements

Compared to the measurement of t-channel single top quark production cross section in the previous chapter and the search for tH production in the following chapter, the analysis in this chapter utilizes an improved b tagging algorithm, the DeepCSV algorithm (see also Section 5.1.5). Although the data recorded in 2016 is also used in the search for tH production, the DeepCSV

Table 7.3.: Predicted and observed event yields for the 2-jets–2-tags signal region, divided by the flavor of the lepton and the year. The uncertainties quoted refer to all statistical and systematic uncertainties included in the analysis.

2-jets–2-tags region	2016 (μ)	2016 (e)	2017 (μ)	2017 (e)
t channel	4827 ± 773	3278 ± 510	5390 ± 1173	4469 ± 900
associated tW	1976 ± 231	1591 ± 213	2198 ± 319	2043 ± 299
Top quark pairs	54 811 ± 8369	41 668 ± 5899	61 467 ± 11 427	53 713 ± 9346
W+jets	6141 ± 828	3928 ± 575	6116 ± 7944	4619 ± 7557
Z+jets	1101 ± 144	753 ± 270	1004 ± 135	1011 ± 153
QCD multijet events	1263 ± 626	2848 ± 1411	2475 ± 1226	2148 ± 1066
Sum of backgrounds	70 119 ± 8473	54 065 ± 6123	78 649 ± 14 024	68 001 ± 12 105
s channel	1001 ± 41	643 ± 29	1189 ± 1182	916 ± 62
Data	76 807	57 550	83 892	67 552

algorithm has only become available recently. The new algorithm provides higher efficiency for b jets at the same working point, i. e., for the same mistagging efficiency, compared to the CSVv2 algorithm. The effect is illustrated in Fig. 7.4 for simulated events of s-channel single top quark production and top quark pair production, which is further divided into semileptonic and dileptonic top quark pair decays. The b tag efficiency for b jets is increased by at least 5% across the entire transverse momentum spectrum for simulations of the 2016 data-taking period. This is of great importance for the event selection process, as the population of events in specific regions strongly depends on the number of correctly identified b jets. The same effect can be seen in simulations of the 2017 data-taking period. In addition, the b tag efficiency in 2017 is higher overall compared to 2016 because of the upgrade of the pixel detector during the shutdown of the LHC between 2016 and 2017.

7.4. Simulations of Events and Corrections

The simulation of the s-channel single top quark production process is performed with the MadGraph5_aMC@NLO MC event generator, version 2.2.2, using the 4FS description [259]. The other single top quark production processes, t channel and associated tW production, are simulated with Powheg [176,177] version 2.0 in the 4FS [259] and 1.0 in the 5FS [244], respectively. Top quark pair production is also modeled using Powheg version 2.0. For each simulation involving top quarks, a top quark mass of 172.5 GeV is used. A prediction of W/Z+jets production is obtained by employing MadGraph5_aMC@NLO with the help of the FxFx merging technique to combine different multiplicities of additional partons in the matrix element. For all simulations, the parton shower and the hadronization process are performed with Pythia version 8.2 [180]. Underlying event contribution for each simulated event is modeled with the tune CUETP8M1 [181,182], except for the top quark pair production, where the tune CUETP8M2T4 is used. For simulations of 2017, the tune CP5 is used for all processes. The

Table 7.4.: Predicted and observed event yields for the 2-jets–1-tag, 3-jets–1-tag and 3-jets–2-tags control regions, divided by the flavor of the lepton and the year. The uncertainties quoted refer to all statistical and systematic uncertainties included in the analysis.

2-jets–1-tag region	2016 (μ)	2016 (e)	2017 (μ)	2017 (e)
t channel	58 303 ± 8797	37 276 ± 5707	61 981 ± 10 100	48 134 ± 7886
associated tW	27 483 ± 3188	21 581 ± 2441	30 649 ± 53 109	27 666 ± 3191
Top quark pairs	224 187 ± 36 352	180 835 ± 26 890	241 416 ± 49 315	223 198 ± 39 488
W+jets	219 092 ± 27 482	128 766 ± 16 605	228 810 ± 257 522	169 767 ± 198 812
Z+jets	16 816 ± 2047	17 364 ± 2796	18 457 ± 2807	21 251 ± 4219
QCD multijet events	26 628 ± 13 226	55 558 ± 27 574	50 472 ± 25 032	38 309 ± 19 019
Sum of backgrounds	572 509 ± 48 409	441 380 ± 42 491	631 785 ± 268 899	528 326 ± 203 808
s channel	1869 ± 66	1169 ± 50	1909 ± 106	1463 ± 84
Data	605 852	456 507	666 227	543 878

3-jets–1-tag region	2016 (μ)	2016 (e)	2017 (μ)	2017 (e)
t channel	16 226 ± 3015	11 677 ± 2228	17 008 ± 4693	14 150 ± 4213
associated tW	18 005 ± 2558	15 717 ± 2218	20 779 ± 2528	20 629 ± 2568
Top quark pairs	247 004 ± 35 362	208 498 ± 28 130	263 837 ± 35 155	255 289 ± 33 147
W+jets	69 736 ± 11 767	48 725 ± 8560	68 809 ± 107 589	56 257 ± 90 688
Z+jets	6103 ± 1023	6953 ± 1479	6506 ± 1551	8703 ± 2595
QCD multijet events	7418 ± 3684	22 524 ± 11 196	12 957 ± 6430	10 397 ± 5157
Sum of backgrounds	364 491 ± 37 672	314 094 ± 31 654	389 897 ± 113 505	365 427 ± 96 854
s channel	599 ± 46	431 ± 33	551 ± 98	477 ± 81
Data	375 175	314 254	426 697	381 108

3-jets–2-tags region	2016 (μ)	2016 (e)	2017 (μ)	2017 (e)
t channel	7703 ± 1252	5584 ± 927	8376 ± 1452	7076 ± 1268
associated tW	3512 ± 543	3127 ± 441	4028 ± 452	4089 ± 455
Top quark pairs	118 112 ± 17 342	94 941 ± 13 180	135 173 ± 19 094	125 036 ± 16 584
W+jets	4978 ± 847	3390 ± 564	4246 ± 6735	3833 ± 5946
Z+jets	819 ± 159	768 ± 154	681 ± 128	754 ± 184
Sum of backgrounds	135 124 ± 17 417	107 809 ± 13 233	152 504 ± 20 305	140 789 ± 17 670
s channel	379 ± 24	274 ± 18	395 ± 50	344 ± 46
Data	140 920	112 363	159 244	135 069

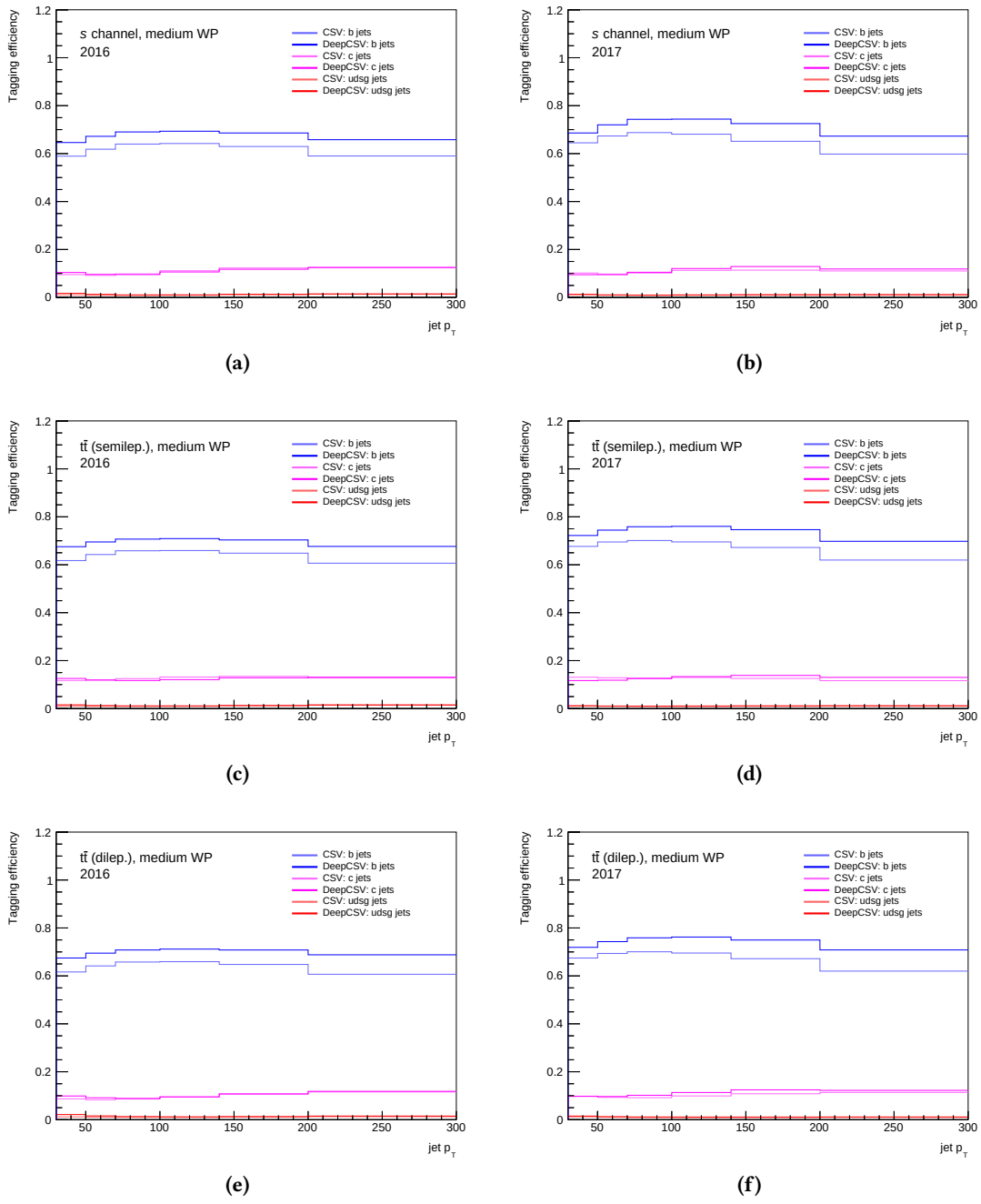


Figure 7.4.: Comparison of the CSVv2 and DeepCSV b tagging efficiencies for *s*-channel single top quark production ((a) and (b)), top quark pair production with semileptonic decays ((c) and (d)) and top quark pair production with dileptonic decays ((e) and (f)). The efficiencies are shown for the same working point, which is defined by a mistagging rate of about 1%. The higher efficiency in 2017 can be attributed to the new pixel detector, which has been installed between 2016 and 2017.

probability for each parton in the initial state to be present in the proton with a given fraction of the proton momentum is obtained by the NNPDF30_NLO_AS_0118 PDF [159] for simulations of 2016 and the NNPDF31_NNLO_HESSIAN_PDFAS PDF [160] in 2017. Experimental conditions, such as pileup interactions, are added for each event to simulate the data-taking conditions. A full list of simulation samples is provided in Appendix B.1.

7.4.1. Number of Pileup Interactions

As already explained in Section 6.4.1 of the previous chapter, the pileup profile for simulated events is estimated before the actual data-taking. To correct this number in order to resemble the actual conditions and to describe the measured data properly, the distribution of the number of pileup interactions in simulation is reweighted based on the distribution obtained from minimum-bias events, measured in data [245, 246]. The effect of this reweighting on the distribution of pileup interactions per event is shown in Appendix B.2.1, separately for data taken in 2016 and 2017.

7.4.2. Lepton Efficiencies

The reconstruction and selection efficiencies of leptons depend on the kinematic properties of the lepton and differences in simulation and data can occur. To correct for such differences, dedicated scale factors are applied to simulated events to take this effect into account.
For muons, the overall efficiency is given by:

$$\epsilon = \epsilon_{\text{Tracking}} \cdot \epsilon_{\text{ID|Tracking}} \cdot \epsilon_{\text{Iso|ID}} \cdot \epsilon_{\text{Trigger|Iso}}, \tag{7.1}$$

with efficiencies for tracking, ID, isolation and trigger. Individual efficiency corrections are used for the data taken in 2016 [290] and 2017 [291]. These corrections are provided by the MUO POG and have been calculated with a tag-and-probe method [248–250]. Since an isolation of 0.06 is chosen for the selection of muons and no scale factors are centrally available for this value, the efficiency corrections for muon isolation and trigger are derived specifically for this analysis with a tag-and-probe method.
Similarly to muons, efficiency corrections are applied for all selected electrons. These scale factors include corrections for the simulation, accounting for electron reconstruction, identification and trigger. The electron trigger efficiencies for the HLT in 2016, `HLT_Ele32_eta2p1_WPTight_Gsf_v*`, have been measured in the scope of an updated version of the analysis described in Chapter 6 with the data of 2016 [292, 293]. For the `HLT_Ele30_eta2p1_WPTight_Gsf_CentralPFJet35_EleCleaned_v*` trigger in 2017, the trigger efficiencies have been measured exclusively for this analysis.
The result of these lepton efficiency corrections for muons and electrons are shown in Appendix B.2.2 for the data taken in 2016 and 2017, respectively.

7.4.3. b Tagging Efficiencies

The efficiency for the employed b tagging algorithm is different for jets in simulation and data. Therefore, the simulation is adjusted to resemble the same number of events in simulation

with a certain number of b-tagged jets as in data. To achieve this, events in simulation are reweighted with scale factors dependent on the kinematic properties of each jet in the event and the b tagging efficiency in simulation [251]. These efficiencies are derived for each process independently and year of data-taking. The jet-based scale factors are provided by the BTV POG of the CMS Collaboration for 2016 [294] and 2017 [295]. These scale factors are calculated following the same procedure as outlined in Section 6.4.3. The effect of applying this correction is shown in Appendix B.2.3.

7.5. Background Estimation for QCD Multijet Events

Although the contribution from QCD multijet events can be significantly reduced by the event selection, this process still contributes to the background processes in the selected phase-space regions because of its high cross section. Since a sufficient modeling of this process would require extremely large simulation samples, a data-driven approach is used to model the kinematic properties and to estimate the normalization of this process. Similarly to Section 6.5, an orthogonal sideband region is defined. This sideband region is almost entirely populated with QCD multijet events and the selected data in the sideband region can be used to model the QCD multijet events in the signal region, as the kinematic properties are independent of the sideband-defining variables. The normalization of QCD multijet contribution in the signal region is extracted with a two-parameter binned maximum-likelihood fit to the distribution of a certain variable x, which is capable of distinguishing QCD multijet events from events of non-QCD processes. The fit function of this estimation is defined as

$$F(x) = N_{\text{QCD}} \cdot F_1(x) + N_{\text{non-QCD}} \cdot F_2(x), \tag{7.2}$$

with the template F_1, obtained from data in the sideband region, where the remaining non-QCD contributions have been subtracted from the data, the template F_2 of non-QCD processes in the signal region obtained from simulation and the two fit parameters N_{QCD} and $N_{\text{non-QCD}}$ for the number of QCD and non-QCD events, respectively. The variable x depends on the flavor of the selected lepton and is discussed in the following subsections. This estimation is done independently for the 2-jets–1-tag, 2-jets–2-tags and 3-jets–1-tag regions, while the QCD contribution in the 3-jets–2-tags region can be neglected.

7.5.1. Muon Channel

For selected events with muons, the sideband region is defined by inverting the muon isolation requirement from lower than 0.06 to higher than 0.2. The variable in the maximum-likelihood fit to extract the QCD normalization is the transverse W boson mass. This variable provides excellent separation power between QCD multijet events and non-QCD processes, as all non-QCD processes involve the decay of a real W boson, featuring a peak around the W boson mass of about 80 GeV in the transverse W boson mass distribution. In contrast, the distribution of the transverse W boson mass is steadily falling from 0 GeV for QCD multijet events because of a virtual W boson decay. For the yields estimation, the selection criterion for the transverse W boson mass is relaxed and the yields for $m_{\text{T}}^{\text{W}} > 50\,\text{GeV}$ are given by integrating the fitted

Table 7.5.: Results of the QCD estimation for the muon channel in 2016 and 2017 after applying a requirement of $m_{\mathrm{T}}^{\mathrm{W}} > 50$ GeV. The uncertainties quoted include only the statistical uncertainty of the maximum-likelihood fit.

Muon channel	2016	2017
2-jets–2-tags	1263 ± 36	2475 ± 58
2-jets–1-tag	26 628 ± 125	50 472 ± 192
3-jets–1-tag	7418 ± 66	12 957 ± 93

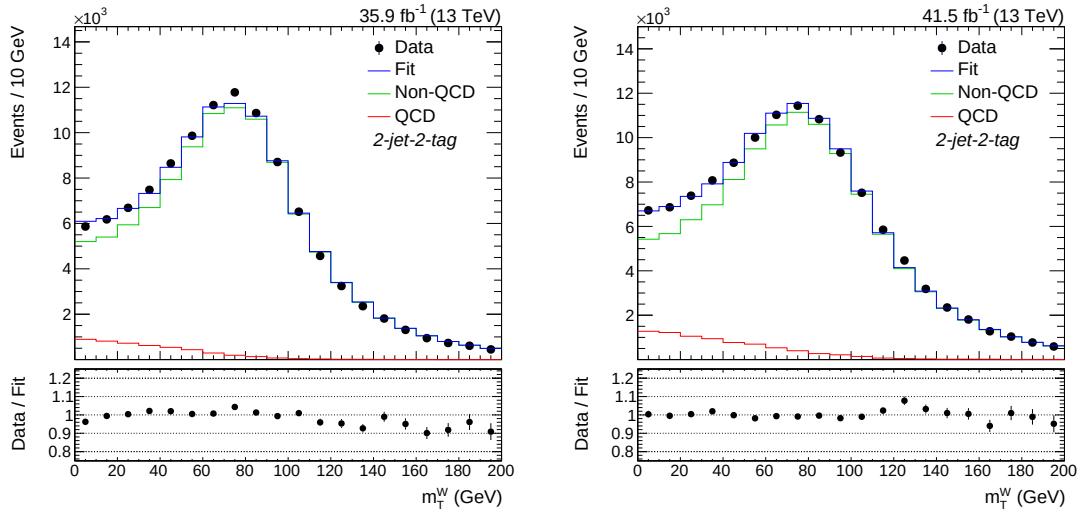


Figure 7.5.: The result of the QCD estimation from the fit to the transverse W boson mass distribution in the muon channel in 2016 (a) and 2017 (b). Good agreement between data and the fitted templates is observed for both years.

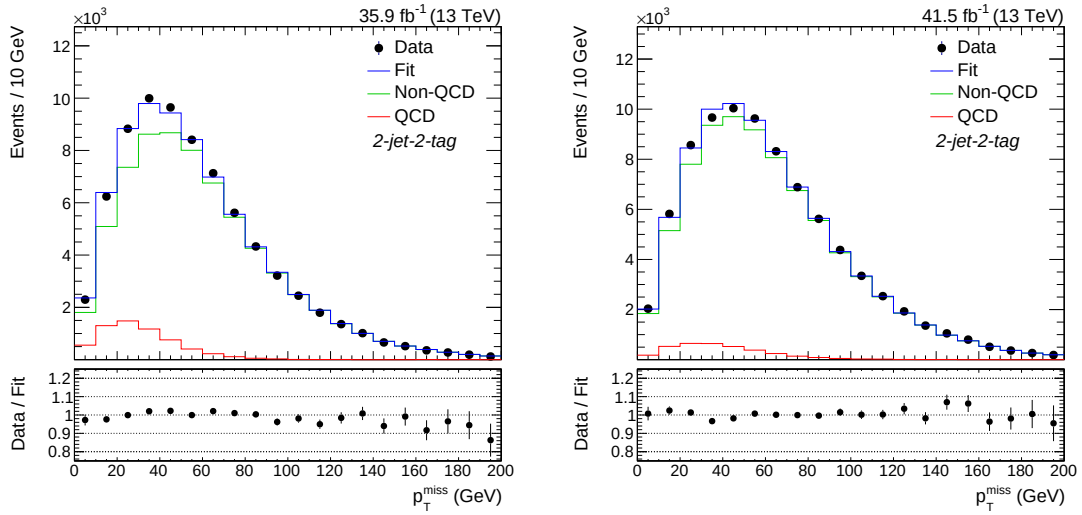
distribution of the whole spectrum from 50 GeV to infinity. The result for the yields in each region are summarized in Table 7.5. The fitted distributions for the QCD estimation in 2016 and 2017 are shown in Fig. 7.5 for the 2-jets–2-tags signal region. The remaining distributions can be found in Appendix B.3 for the 2-jets–1-tag region and in Appendix B.4 for the 3-jets–1-tag control region.

7.5.2. Electron Channel

Since the isolation is included in the ID requirements for electrons (see Section 5.1.3), the sideband region for selected events with electrons is defined by explicitly failing the veto ID, instead of passing the tight ID. The missing transverse momentum is chosen as the discriminating variable in the maximum-likelihood fit, since the transverse W boson mass is not modeled well enough for events with electrons. High missing transverse momentum is caused by neutrinos from real W boson decays, which results in a peak at a value of around half the W boson mass.

Table 7.6.: Results of the QCD estimation for the electron channel in 2016 and 2017 after applying the requirement $p_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$. The uncertainties quoted include only the statistical uncertainty of the maximum-likelihood fit.

Electron channel	2016		2017	
2-jets–2-tags	2848	± 65	2148	± 73
2-jets–1-tag	55 558	± 282	38 309	± 268
3-jets–1-tag	22 524	± 212	10 397	± 138



(a) **(b)**

Figure 7.6.: The result of the QCD estimation from the fit to the missing transverse momentum distribution in the electron channel in 2016 (a) and 2017 (b). Good agreement between data and the fitted templates is observed for both years.

For QCD processes with virtual W boson decays, the peak of the distribution is shifted to lower values. Similarly to the muon channel, the fit is performed without the requirement on the missing transverse momentum in each event and the resulting yields are determined by the integral of the fitted distribution from $p_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$ to infinity. The obtained yields for all fitted regions are shown in Table 7.6. The fitted distributions in the 2-jets–2-tags region are given in Fig. 7.6, both for the estimation of the 2016 and 2017 data sets. Distributions for the two control regions, 2-jets–1-tag and 3-jets–1-tag, are given in Appendices B.3 and B.4, respectively.

7.6. Top Quark Reconstruction

To reconstruct the top quark from its decay products, i. e., the b jet, the charged lepton and the neutrino, the four-momenta of the associated objects have to be added together. For the reconstruction of the leptonically decaying W boson, the procedure has already been outlined

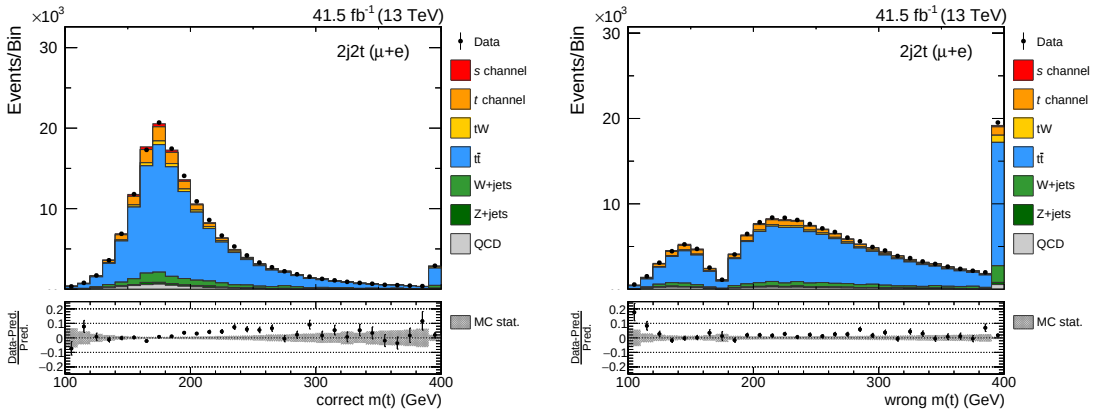


Figure 7.7.: Distribution of the reconstructed top quark mass in the 2-jets–2-tags signal region for different reconstruction hypotheses. The b-tagged jet that results in a reconstructed top quark mass closer to the reference value of 172.5 GeV in each events is chosen for the reconstruction of the top quark (a) and the other b-tagged jet is assumed to be stemming from the weak top quark production vertex. The simulation is scaled to the expected number of events.

in Section 5.2. The assignment of the b-tagged jet to the b jet from the top quark decay is trivial in regions with only one selected b-tagged jet (2-jets–1-tag and 3-jets–1-tag regions). For the 2-jets–2-tags signal region and the 3-jets–2-tags control region, this assignment is ambiguous because of the two b-tagged jets in each event. In this case, both reconstruction hypotheses are evaluated and the one hypothesis is chosen that results in a reconstructed top quark mass closer to the reference value of 172.5 GeV. This is motivated by the fact that the spectrum of the correct assignments should result in a peak of the reconstructed top quark mass distribution at 172.5 GeV, while wrong assignments result in a continuous spectrum. This is exemplarily shown for simulation and data of 2017 in Fig. 7.7.

7.7. Event Classification

Although the signal region is defined to enhance the signal-to-background ratio, the 2-jets–2-tags region is mostly populated with top quark pair production events. It is therefore necessary to have a good separation between the signal process and the dominant background processes, as this will improve the statistical significance of the search for the signal process. To achieve such a separation, the analysis employs an event classification with a DNN, implemented within the self-developed MIST framework. Compared to the simpler and easier classification task for the t-channel single top quark production measurement in Chapter 6, this neural network has multiple hidden layers and more input variables to achieve a similar separation power. This approach is necessary, as s-channel single top quark production cannot be so easily separated from top quark pair production. A separate training is performed for the 2016 and 2017 data sets to avoid any possible mismodeling caused by different data-taking conditions. The training

is done exclusively in the 2-jets–2-tags region on a combined data set of events with muons and electrons. The same input variables are used for the training of the 2016 and 2017 data sets. A full list of input variables is provided in Table 7.7, where the input variables are ranked by their average separation power in the two trainings. The three most important input variables in descending order are the zeroth-order Fox-Wolfram moment [296, 297], the difference in ϕ between the b-tagged jet with the higher transverse momentum (leading) and the top quark, reconstructed from the b-tagged jet with the lower transverse momentum (subleading), and difference in ϕ for the opposite case. All input variables are checked against mismodeling by comparing the simulation with the data. This comparison is shown in Fig. 7.8 for the three input variables with the most separation power in 2016 and in Fig 7.9 for 2017. The corresponding distributions for the remaining input variables can be found in Appendix B.5. The processes used in the training for the DNN are the s-channel single top quark production as signal process and top quark pair production with semileptonic and dileptonic decays as background processes. The contributions of signal and background processes are weighted according to their respective cross section and selection efficiency. The training of the DNN is performed on a special subset of simulated events that are only used for the training step and are discarded afterwards to prevent a bias when applying the network and performing the final maximum-likelihood fit. To check for possible overtraining during the determination of the weights between different network nodes, the result of the DNN obtained with the training data set is compared to the result obtained with the remaining data set. The results for the 2016 and 2017 training are shown in Fig. 7.10. No overtraining has occurred, since no differences in the performance for the training and testing data set are observed.
After the training, the DNN is applied to simulated events of all processes and to data in the 2-jets–2-tags signal region, as well as the 2-jets–1-tag, 3-jets–1-tag and 3-jets–2-tags control regions. For regions with only one b-tagged jet, the leading b jet always refers to the selected b-tagged jet. The assignment of the subleading b jet is then either the untagged jet (2-jets–1-tag region) or the untagged jet with the higher transverse momentum (3-jets–1-tag region). The distributions of the applied DNN in all four regions are shown in Fig. 7.11 and Fig. 7.12 for 2016 and 2017, respectively. The distributions of some input variables and the final DNN output show slight disagreement between simulation and data. In 2016, the effect is mainly visible in a normalization offset. For the prediction of 2017, the differences can be attributed to preliminary calibrations and corrections for the simulation. However, all discrepancies between simulation and data are covered by systematic uncertainties.

Table 7.7.: Complete list of input variables for the training of the DNN. The ensemble of variables include global event properties, as well as variables that depend on the top quark reconstruction, which follows the description of Section 7.6.

Variable	Description
Fox Wolfram #0	Zeroth-order Fox-Wolfram moment of the event
$\Delta\phi(\mathrm{t_s}, \mathrm{b_l})$	$\Delta\phi$ between the top quark reconstructed from the subleading b jet and the leading b jet
$\Delta\phi(\mathrm{t_l}, \mathrm{b_s})$	$\Delta\phi$ between the top quark reconstructed from the leading b jet and the subleading b jet
$m_\mathrm{T}(\mathrm{W})$	Transverse mass of the reconstructed W boson
$m(\mathrm{t})$	Invariant mass of the reconstructed top quark
$\lvert p_\mathrm{T}(\mathrm{t}) - p_\mathrm{T}(\mathrm{b_W})\rvert$	Absolute difference of transverse momenta of the reconstructed top quark and the b jet from the time-like W boson
$\Delta R(\mathrm{t}, \mathrm{b_W})$	ΔR between the reconstructed top quark and the b jet from the time-like W boson
$\Delta\phi(\mathrm{t}, \mathrm{b_l})$	$\Delta\phi$ between the reconstructed top quark and the leading b jet
$p_\mathrm{T}^\mathrm{miss}$	Missing transverse momentum
$m(\ell, \mathrm{b_s})$	Invariant mass of lepton and subleading b jet
p_T^bb	Vectorial sum of transverse momenta of the two b jets of the event
H_T	Scalar p_T sum of all b jets, the isolated lepton and $p_\mathrm{T}^\mathrm{miss}$ in the event
$\Delta\phi(\mathrm{b_l}, \mathrm{b_s})$	$\Delta\phi$ between the two b jets of the event
$p_\mathrm{T}(\ell)$	Transverse momentum of the charged lepton
$\lvert\eta(\ell) - \eta(\mathrm{b_l})\rvert$	Absolute difference of the pseudorapidity between the charged lepton and the leading b jet
$\lvert\eta(\mathrm{t_s}) - \eta(\mathrm{b_l})\rvert$	Absolute difference of the pseudorapidity between the top quark reconstructed from the subleading b jet and the leading b jet
$\lvert\eta(\mathrm{t}) - \eta(\mathrm{b_W})\rvert$	Absolute difference of the pseudorapidity between the reconstructed top quark and the b jet from the time-like W boson
$\cos\theta^*$	Cosine of the angle between the lepton from the top quark decay and the b jet from the time-like W boson
Fox Wolfram #3	Third-order Fox-Wolfram moment of the event

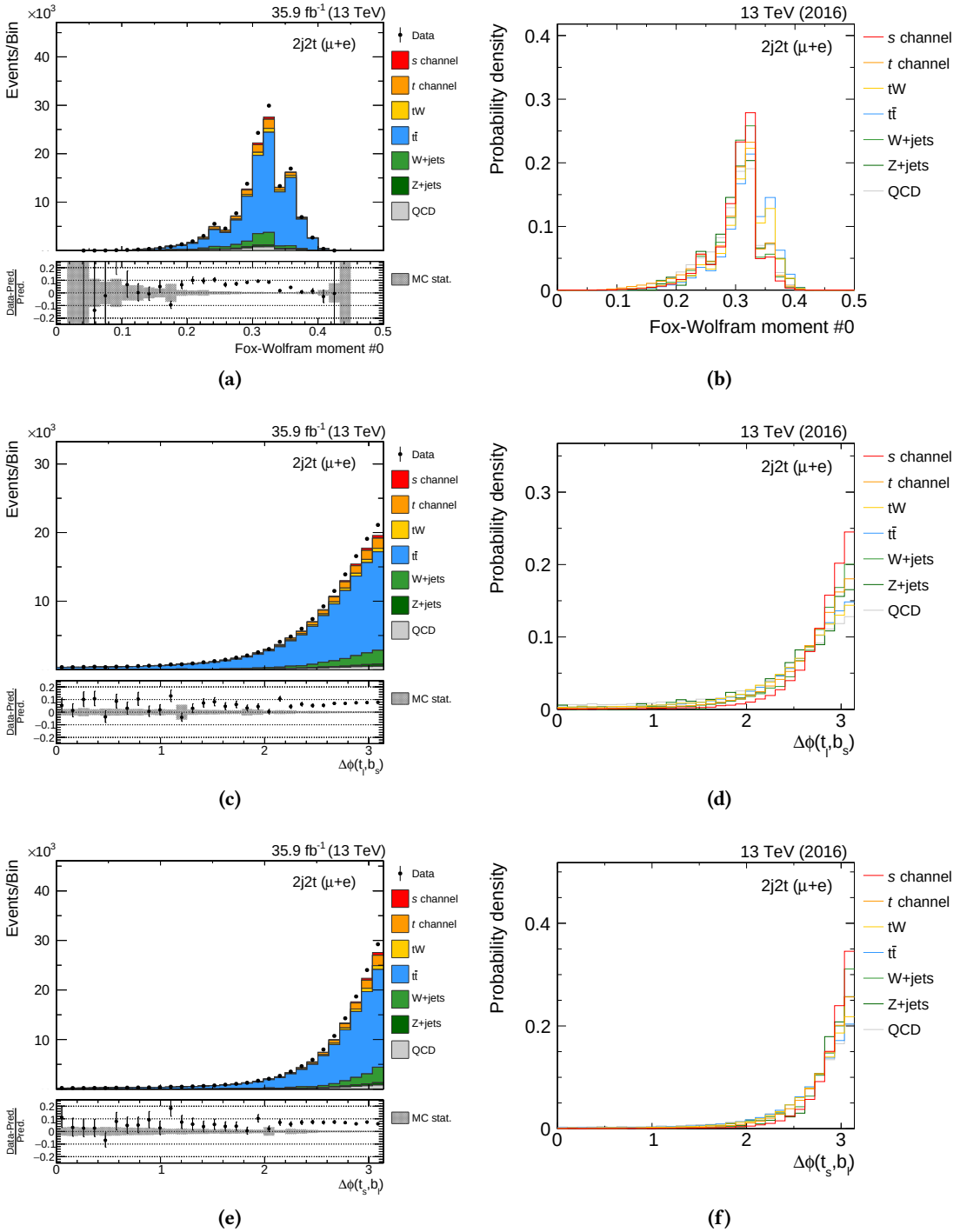


Figure 7.8.: Distributions of the three most significant input variables in the 2-jets–2-tags signal region for the training of the DNN in 2016. These are the zeroth-order Fox-Wolfram moment ((a) and (b)), the difference in ϕ between the b-tagged jet with the highest transverse momentum and the top quark, reconstructed from the other b-tagged jet ((c) and (d)), and the same for the opposite case ((e) and (f)). All variables are shown in comparison with data, where the simulation is scaled to the integrated luminosity, and as a shape comparison between the different processes.

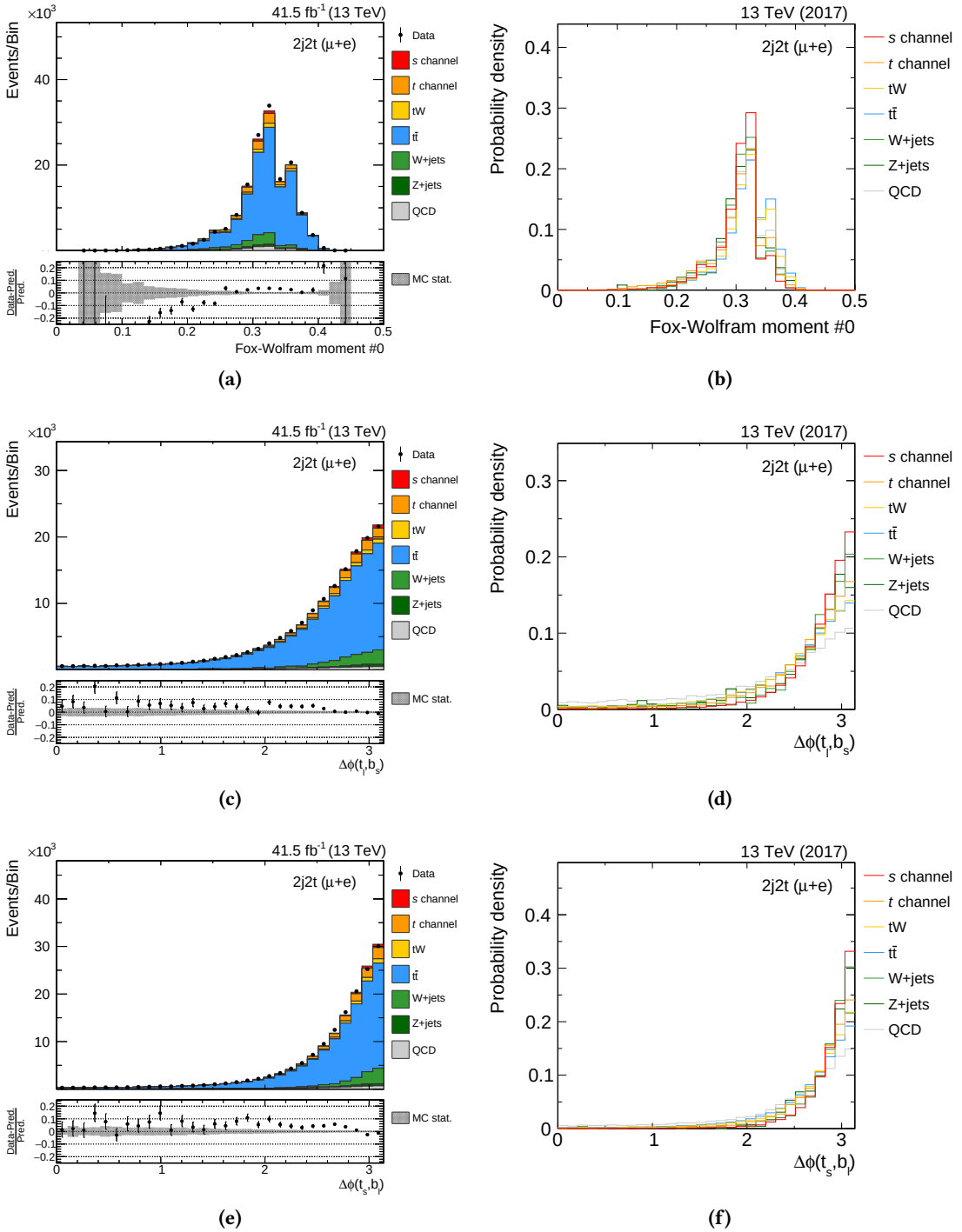
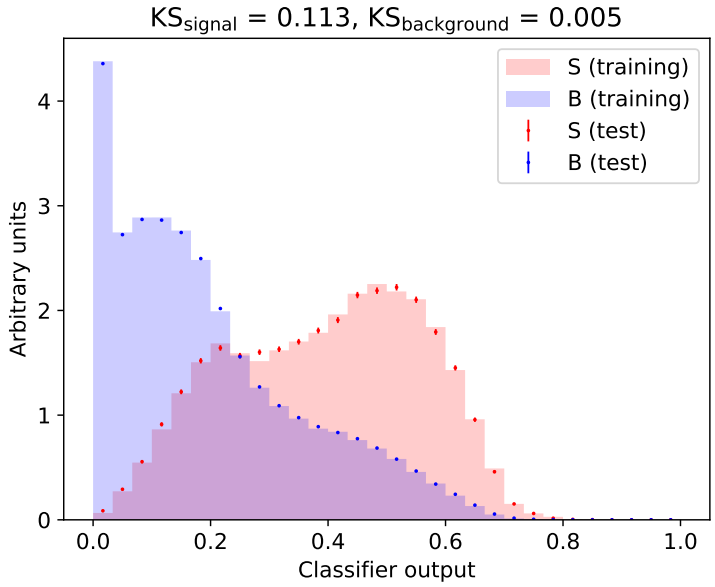
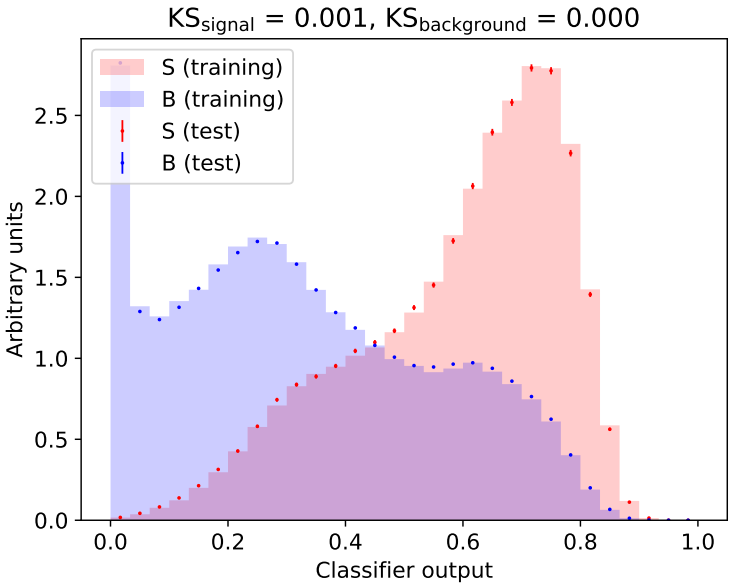


Figure 7.9.: Distributions of the three most significant input variables in the 2-jets–2-tags signal region for the training of the DNN in 2017. These are the zeroth-order Fox-Wolfram moment ((a) and (b)), the difference in ϕ between the b-tagged jet with the highest transverse momentum and the top quark, reconstructed from the other b-tagged jet ((c) and (d)), and the same for the opposite case ((e) and (f)). All variables are shown in comparison with data, where the simulation is scaled to the integrated luminosity, and as a shape comparison between the different processes.



(a)



(b)

Figure 7.10.: Result of the training of the DNN in 2016 (a) and 2017 (a). No overtraining is found when comparing the distributions of the training data set and the data set used in the analysis.

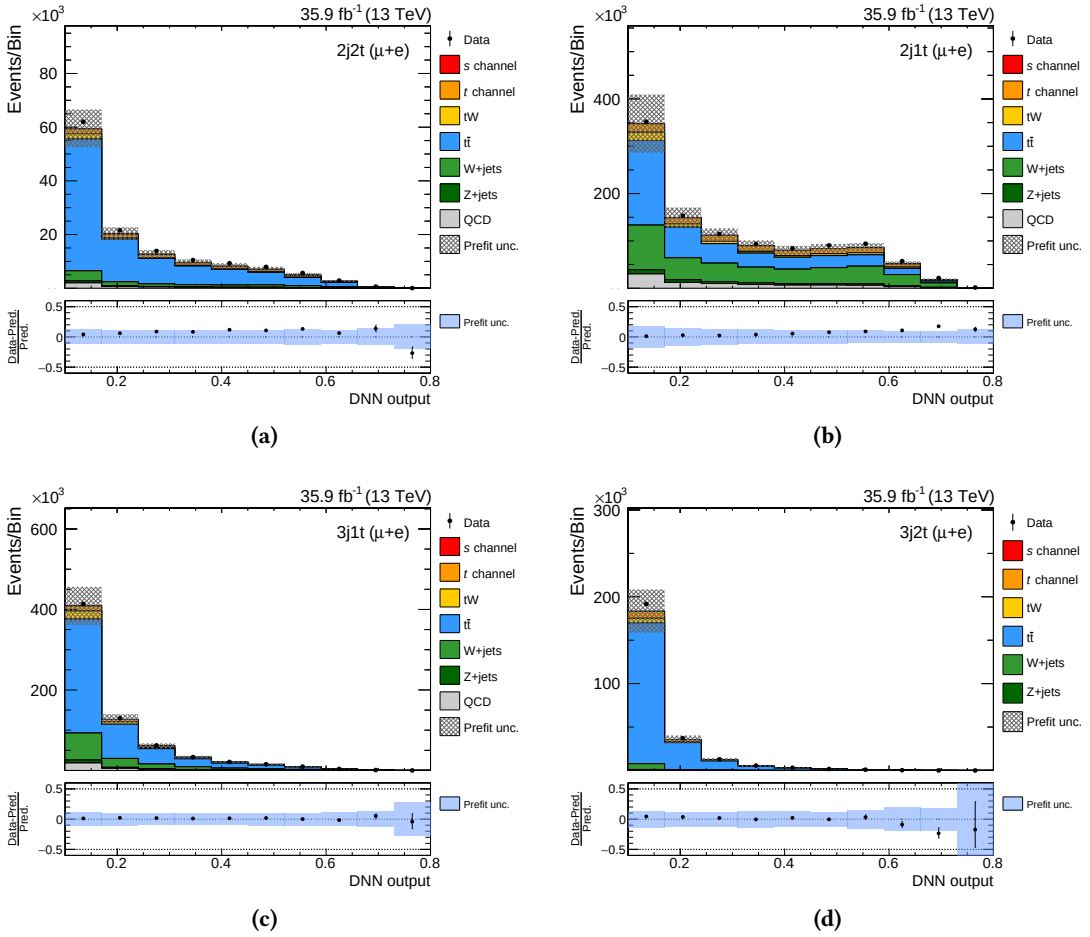


Figure 7.11.: Resulting distributions for the DNN classifier output in 2016 for the 2-jets–2-tags signal region (a), as well as the 2-jets–1-tag (b), 3-jets–1-tag (c) and 3-jets–2-tags (d) control regions. The simulation is scaled to the expected number of events.

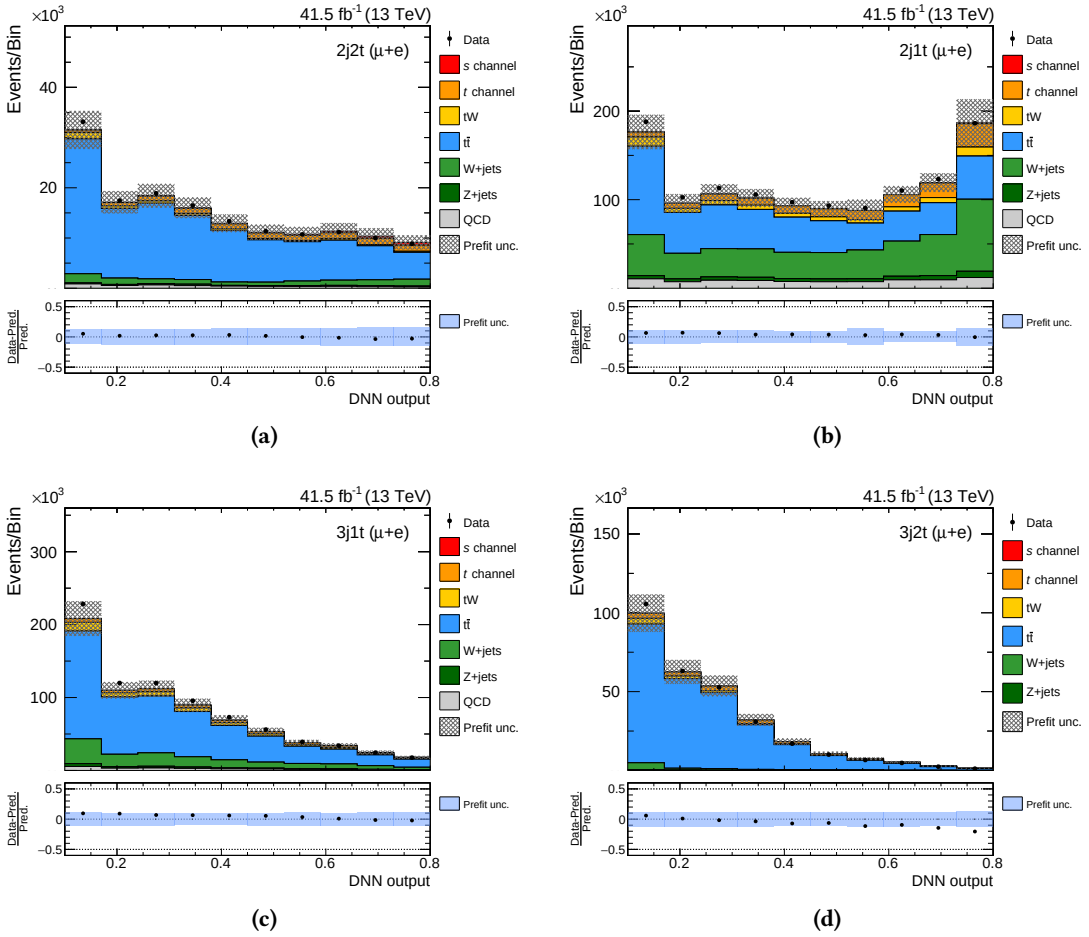


Figure 7.12.: Resulting distributions for the DNN classifier output in 2017 for the 2-jets–2-tags signal region (a), as well as the 2-jets–1-tag (b), 3-jets–1-tag (c) and 3-jets–2-tags (d) control regions. The simulation is scaled to the expected number of events.

7.8. Systematic Uncertainties

The search for s-channel single top quark production is not only limited by the amount of data analyzed, but also affected by multiple different sources of systematic uncertainties. The following section provides a complete list of them. The different systematic uncertainties are grouped into experimental and theoretical uncertainties, although the groups are not treated differently in contrast to their treatment in the analysis described in the previous chapter. Although many systematic uncertainties have already been introduced in the previous chapter, the complete list is provided in this section as the specific treatment for some of the uncertainties has changed or additional systematic effects are considered.

7.8.1. Experimental Uncertainties

- **Jet energy scale (JES):** Uncertainties in the determination of the jet energy corrections applied to simulation (see Section 4.2.6) are applied to all simulated samples to derive an up- and down-varied template [215, 216] for each process. Since these shifted corrections change the four-momenta of all jets and therefore also event-based kinematic variables, the whole analysis is repeated on the corresponding shifted samples to also take migration effects between different regions and different selection efficiencies in general into account.

- **Jet energy resolution (JER):** Corrections applied to jets in simulation to smear the resolution in order to resemble the resolution observed in data, are varied according to their given uncertainties [232]. This uncertainty includes statistical uncertainties in the determination of the corrections, but also systematic effects. The systematic effect due to the JER is also propagated through the whole analysis chain with dedicated simulation samples.

- **Unclustered energy:** A systematic uncertainty is assigned to the determination of the missing transverse momentum to take into account different energy resolutions of different objects that contribute to the calculation of the missing transverse momentum [221]. Similarly to the two previous uncertainty sources, the effect is implemented in dedicated simulation samples on which the nominal analysis steps are repeated to derive shifted templates for each process.

- **b tagging:** The b tagging efficiency corrections applied to simulated events give also rise to an additional systematic uncertainty that has to be considered in the analysis [257]. Each centrally provided jet-dependent scale factor has an associated uncertainty, from which an up- and down-shifted event-based weight is calculated [294, 295]. These shifted weights are used instead of the nominal b tagging correction weights to derive shifted templates for each process.

- **Lepton efficiencies:** The different corrections applied to simulated events accounting for the different efficiencies of leptons in simulation and data, are also subject to systematic effects. Not only the statistical uncertainties in the determination of the nominal scale

factors are taken into account as source of uncertainty, but also additional systematic uncertainty sources because of different event topologies are considered. These include variations of tracking, trigger, identification, isolation and reconstruction efficiencies for electrons and muons.

- **Pileup:** The corrections applied to simulated events to recover the observed distribution of primary vertices depend on the chosen value for the minimum-bias cross section in the reweighting procedure [245, 246]. The nominal value of 69.2 mb is varied by ±4.6% to obtain systematically-shifted corrections that are used to derive the impact of the uncertainty.

- **Luminosity:** The normalization of the simulated MC samples depends on the number of accumulated proton-proton collision data, expressed as integrated luminosity. Since the integrated luminosity cannot be measured with infinite precision, an uncertainty on the overall normalization for all processes is assigned. The integrated luminosity of 2016, 35.9 fb^{-1}, is only known with an accuracy of 2.5% [298]. For 2017 data, the integrated luminosity of 41.5 fb^{-1} has been measured with a precision of 2.3% [299].

- **Limited size of MC simulation samples:** Except for the QCD multijet background, all simulation templates are extracted from MC simulation samples. Since the templates obtained from these simulation samples are derived from a finite number of simulated events, the resulting distributions are affected by the number of events in each sample. The uncertainty due to these limitations are taken into account by a modified Barlow–Beeston method [258, 300]. In general, the method introduces an additional nuisance parameter per bin of each distribution and per process with a prior based on the number of events from which the corresponding distributions of each process had been derived in the specific bin. Instead, a simplified approach is used within the employed COMBINE tool that adds such a nuisance parameter for each bin, but for all processes combined.

7.8.2. Theoretical Uncertainties

- **Background normalization:** The normalization of each background template depends on the corresponding theoretical cross section when the background contribution is derived from simulation. For the QCD multijet background, the normalization is directly taken from the fit result. A rate uncertainty is assigned to each background process with a magnitude reflecting the uncertainty of recent measurements of the given process [263, 282, 301, 302]. In case of the QCD multijet background, a conservative uncertainty of 50% is chosen to account for possible unknown systematic effects.

- **Factorization and renormalization scales:** The uncertainty due to the choice of factorization and renormalization scales in simulated events is taken into account by deriving systematically shifted templates from a reweighting procedure based on LHE weights [260]. These weights correspond to half or double the value of the nominal factorization and renormalization scale. The envelope of all different combinations is used for the final uncertainty, omitting combinations where the two different scales are varied

in opposite directions. Since this method is not available for the simulation of associated tW single top quark production, dedicated samples with different scales are used for the simulation of 2016. For the tW simulation in 2017, a rate uncertainty of 3%, motivated by the theory prediction, is chosen due to the lack of dedicated systematically-shifted simulation samples.

- **PDF and α_s:** Uncertainties caused by the choice of the PDF and the value of α_s are taken into account by deriving shifted templates following the PDF4LHC recommendations [40, 41]. In total, 100 different variations of the nominal PDF and two variations of α_s are considered to construct 102 systematically-shifted templates for each process. The final up and down variations are then extracted by calculating the root mean square from each variation with respect to the nominal template.

- **Modeling of the top quark p_T:** The modeling of the top quark transverse momentum in top quark pair production events, obtained from NLO event generators, predicts a distribution that is shifted to higher values with respect to the measured distribution [261]. In principle, the observed mismodeling can be corrected by a two-parameter scale factor for the simulation of top quark pair production events, based on the transverse momentum of the top quark and antiquark. However, it is not recommended by the Top Quark Physics Analysis Group of the CMS Collaboration [303] to apply this reweighting by default. The application of this method is instead used as an additional systematic uncertainty, where a shifted template of top quark pair production events is obtained by applying the corresponding scale factors.

- **Initial- and final-state radiation:** The chosen value of α_s in the parton shower influences the probability of additional gluon radiation in the initial or final state of simulated processes. Dedicated templates are derived with half and double the probability for initial- and final-state radiation [304]. For the simulation of 2016, the effect is considered by dedicated simulation samples for the top quark pair production background. In the simulation of the 2017 data-taking period, the effect is implemented in dedicated event weights provided by Pythia, which are available for *s*-channel and associated tW single top quark production, as well as for top quark pair production.

- **Matching scale:** The scale at which the matrix-element calculation from Powheg and the parton shower from Pythia are matched is controlled by the h_{damp} parameter of Powheg [304]. In addition, the h_{damp} parameter controls the radiation of partons with high transverse momentum, which can lead to additional jets passing the selection criteria of the analysis. Dedicated simulation samples are used for the top quark pair production background that contain variations of h_{damp} $(^{+0.66}_{-0.59})$ with respect to the nominal value of $1.58 \cdot m_{top}$ [305].

- **Underlying event:** The generator tune for the simulation of top quark pair production in 2016 (CUETP8M2T4) and 2017 (CP5) is responsible for the modeling of the underlying event and multi-parton interactions. Dedicated simulation samples for the top quark pair production in both years are used to take specific choices of generator settings into

account [304]. In particular, settings for multi-parton interaction and color reconnection are varied [305].

7.9. Results

To observe s-channel single top quark production as a possible signal in data, a maximum-likelihood fit is performed to the DNN output distribution of the 2-jets–2-tags signal region and the 2-jets–1-tag, 3-jets–1-tag and 3-jets–2-tags control regions, both for the muon and electron channel. This results in a total of eight distributions that are fitted simultaneously for each year. Results are first obtained for each year separately and the significance of the signal is determined with respect to the background-only hypothesis using a profile-likelihood approach.
Afterwards, a combined maximum-likelihood fit is used including both the distributions of 2016 and 2017. A combined significance for the signal process is calculated including both years of data-taking. In case of an excess above the background-only hypothesis for s-channel single top quark production in the combined data set of 2016 and 2017, i. e., a significance of at least three standard deviations above the background-only hypothesis, the measured cross section is extracted by scaling the predicted cross section of 10.32 pb with the signal strength μ obtained from the maximum-likelihood fit.

7.9.1. 2016 Data

The maximum-likelihood fit only to 2016 data returns an observed significance of 3.45 standard deviations for the signal-plus-background hypothesis with respect to the background-only hypothesis. A significance of 2.66 standard deviations for the signal is expected from simulation. The resulting postfit distributions are shown in Appendix B.6. The accumulated data of 2016 is not sufficient to claim an observation, but evidence for s-channel single top quark production has been found.

7.9.2. 2017 Data

When fitting to the 2017 data, an observed significance of 1.88 standard deviations is obtained, while a significance of 3.02 standard deviations is expected when assuming the SM cross section of s-channel single top quark production. The postfit distributions are shown in Appendix B.7. Similarly to the fit to 2016 data, the result obtained with only 2017 data is not sufficient to claim an observation for the signal process. Furthermore, the excess above the background-only hypothesis is not significant enough to claim evidence as for the 2016 data.

7.9.3. Combination of 2016 and 2017 Data

Since the two data sets of 2016 and 2017 are of roughly similar size, more sensitivity can be gained by combining both data sets in a combined maximum-likelihood fit. The different processes share the same nuisance parameters in both years, i. e., adjusting the simulation in one year also affects the same simulation in the other year. Systematic uncertainties of the same origin, e. g., the jet energy scale uncertainty, are correlated between both years, since the determination of

each systematic uncertainty has not changed between both years and the statistical component in the determination of each uncertainty is smaller compared to the systematic component. The only exception is the systematic uncertainty due to the limited size of MC simulation samples. Here, the uncertainties are not correlated between both years since the systematic uncertainty refers to the statistical uncertainty of the simulation samples and different samples are used in each year. The pulls of all nuisance parameters can be found in Appendix B.8.
The result of the combined maximum-likelihood fit yields an observed significance of 4.63 standard deviations, while 4.45 standard deviations are expected from single top quark production in the *s*-channel based on the SM. Evidence for *s*-channel single top quark production can be claimed. The cross section is derived from the signal strength μ of the maximum-likelihood fit and measured to be:

$$\sigma_{s\ \mathrm{ch.}} = 11.66\ ^{+3.50}_{-3.38}\ (\mathrm{syst} + \mathrm{stat})\ \mathrm{pb}, \tag{7.3}$$

which is in agreement with the standard model prediction. The corresponding postfit distributions for the combined fit are shown in Fig 7.13 and Fig 7.14 for data of 2016 and 2017, respectively. The combination of all distributions can be found in Fig. 7.15, where each bin of the fitted distributions is shown according to their respective signal-to-background ratio. Here, the most sensitive bins show a significant excess above the background-only hypothesis. The data can only be accurately described by considering single top quark *s*-channel production.

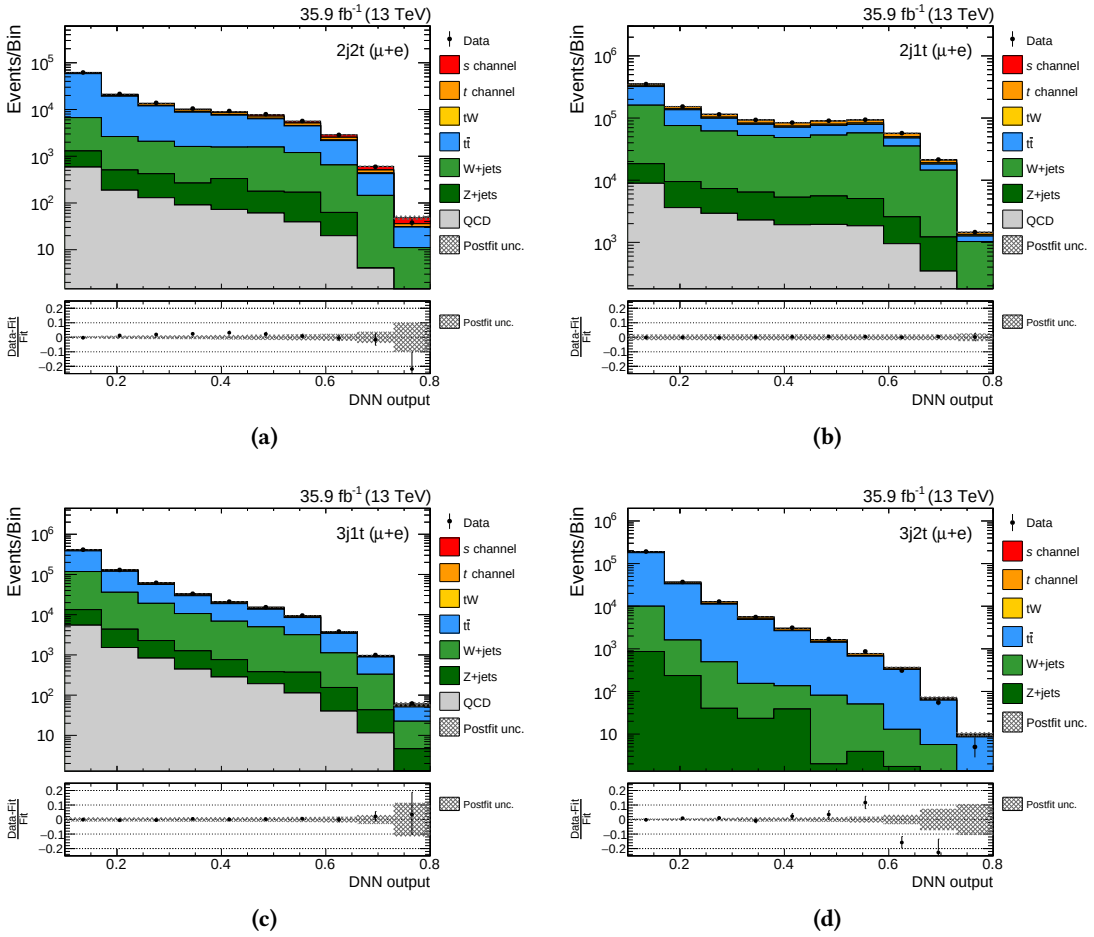


Figure 7.13.: Distributions for the DNN classifier output in 2016 for the 2-jets–2-tags signal region (a), as well as the 2-jets–1-tag (b), 3-jets–1-tag (c) and 3-jets–2-tags (d) control regions, after the maximum-likelihood fit to the combined 2016 and 2017 data. The simulation is scaled to the result of the fit. Good agreement is observed between simulation and data across all regions.

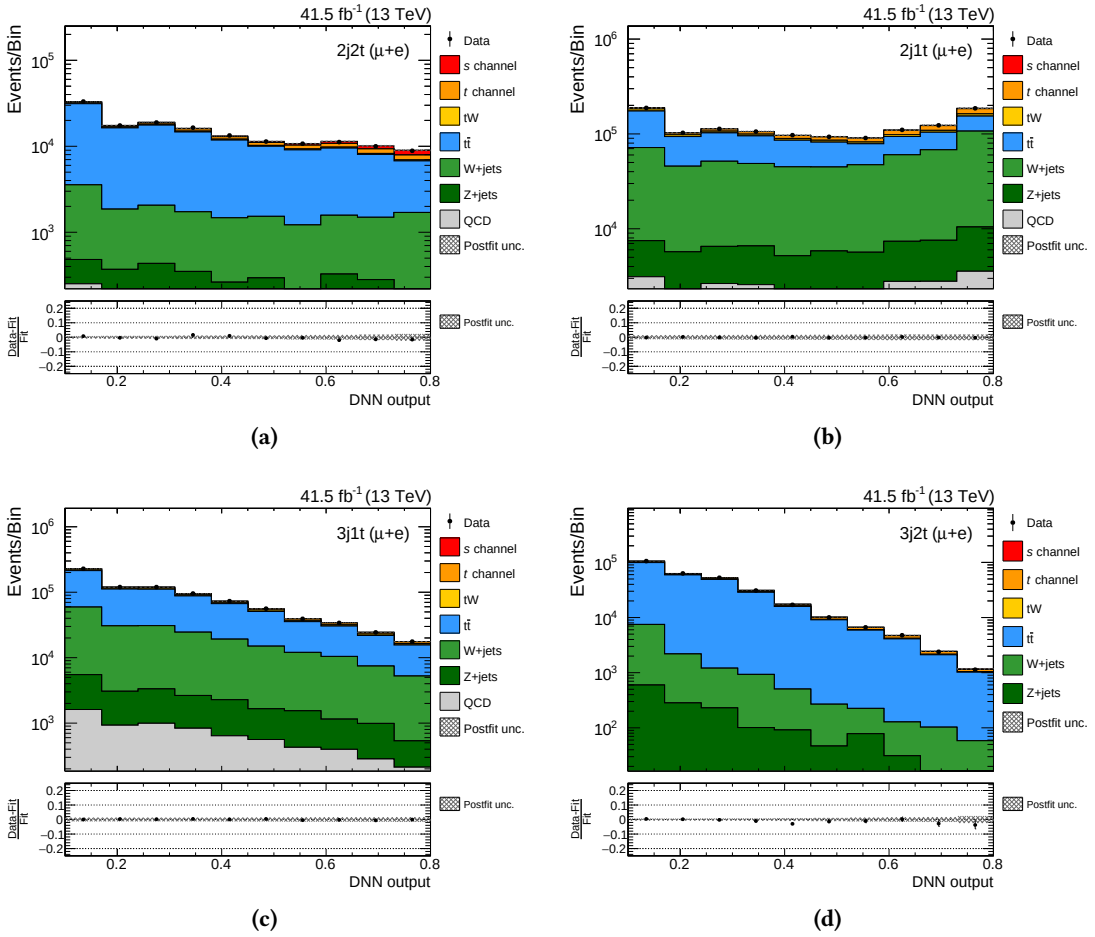


Figure 7.14.: Distributions for the DNN classifier output in 2017 for the 2-jets–2-tags signal region (a), as well as the 2-jets–1-tag (b), 3-jets–1-tag (c) and 3-jets–2-tags (d) control regions, after the maximum-likelihood fit to the combined 2016 and 2017 data. The simulation is scaled to the result of the fit. Good agreement is observed between simulation and data across all regions.

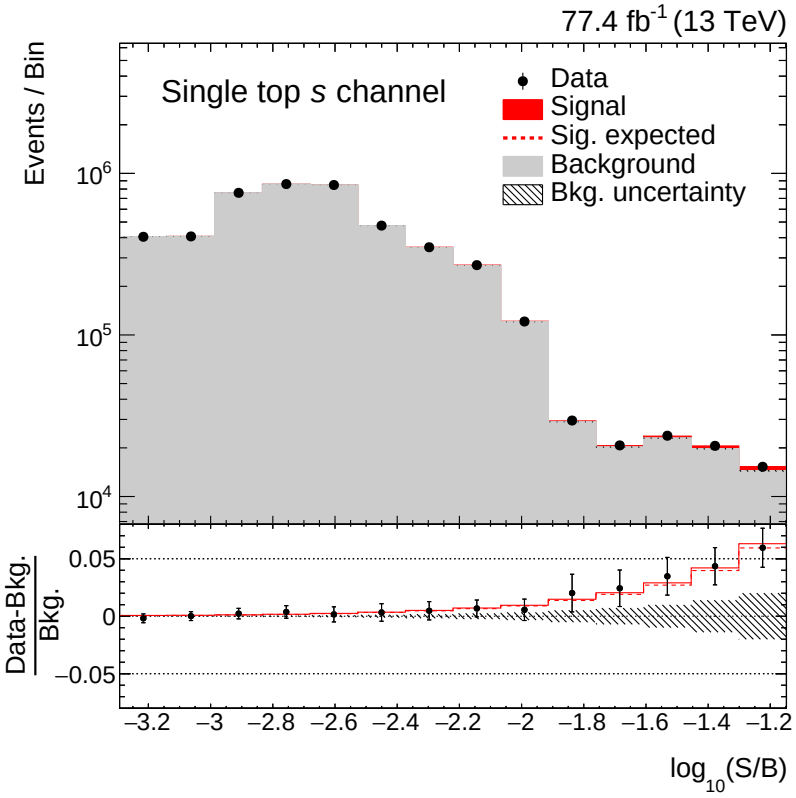


Figure 7.15.: Combined 2016 and 2017 comparison of prediction and data. All bins of the fitted distributions are rearranged according to their signal-to-background ratio such that an excess above the background-only hypothesis would be visible to the right. A significant excess above the background-only hypothesis is observed, which can be described by SM single top quark production in the *s* channel.

7.10. Summary and Outlook

In this section, a search for the at the LHC yet unobserved *s*-channel single top quark production with the CMS experiment has been presented. The search includes the whole data sets of 2016 and 2017, corresponding to an integrated luminosity of 35.9 fb^{-1} and 41.5 fb^{-1}, respectively. Compared to the strategy of this analysis, which is based on a DNN, the ATLAS Collaboration follows a different approach for which they report an increased sensitivity using a matrix element method (MEM) compared to a BDT [286, 287]. With the MEM, a probability to observe an event from measured objects in the detector can be calculated. This incorporates theoretical assumptions, such as matrix element calculations and PDFs, directly into the analysis. However, the MEM requires an enormous amount of computation power to integrate over the phase space of the different partons, which limits the number of partons included in the calculation and the accuracy to LO at the moment. In addition, the transfer function, responsible for linking the final-state partons to the objects in the detector, needs to be well known.
To increase the sensitivity of future analyses, additional proton-proton collision data at 13 TeV will be beneficial. Before going into a two-year shutdown at the end of 2018, the LHC is expected to deliver additional 60 fb^{-1} data to analyze. Other limitations to the presented analysis are the size of available simulation samples, in particular for the W+jets background. The most dominating background of the analysis, the top quark pair production, is slowly becoming a new standard candle in high energy particle physics. A possible way to improve the modeling uncertainties of this background would be to loosen the dependence on dedicated simulation programs and to derive data-driven templates.
Apart from improving the existing analysis, the final state of *s*-channel single top quark production depends on the mediator, in this case a charged SM gauge boson. The same final state with a top and bottom quark also results from the decay of new charged bosons, predicted by different theories of physics beyond the standard model. For example, this includes heavy charged gauge bosons (W') [306, 307] or additional charged Higgs bosons ($H^{\pm}$) [308–311]. Other possible scenarios in which the single top quark *s*-channel topology can be used to search for physics beyond the standard model are dark matter particles with flavor-changing top quark couplings [312]. Deviations in the measurement of *s*-channel single top quark production cross section could therefore be an indirect hint for the existence of such new charged bosons.
The increased amount of data also makes it possible to probe the top-Yukawa coupling of the Higgs boson to top quarks with single top quark events. This analysis is presented in the next chapter.

8. Search for tH Production with $H \to b\bar{b}$ at $\sqrt{s} = 13$ TeV and Study of Higgs Boson Couplings

After establishing single top quark production at a center-of-mass energy of $\sqrt{s} = 13$ TeV with the measurements presented in the previous chapters and the increased amount of data taken in Run II of the LHC, further studies of top quark properties can be performed. One of these properties is the top-Yukawa coupling, which may play a special role in the SM as its value is close to 1. The top-Yukawa coupling is also well-suited to test if the newly observed boson at the LHC during Run I [1, 2] is the long searched-for Higgs boson of the SM. The best way to directly measure the top-Yukawa coupling is the associated production of a Higgs boson with top quark pairs, as the production cross section of these processes is directly proportional to the magnitude of the top-Yukawa coupling. In addition, the associated production with single top quarks allows not only to determine the magnitude of the coupling, but also its sign due to interference effects, resolving this ambiguity from the associated production with top quark pairs. For this reason, the analysis described in this chapter is specialized in searching for the associated production of a Higgs boson with a single top quark. In particular, the analysis focuses only on one decay channel of the Higgs boson, where the Higgs boson decays to a bottom quark-antiquark pair. Although this final state is difficult to analyze due to the high jet multiplicity, the large number of events makes it possible to utilize multivariate methods in order to counteract these circumstances.

A special case of possible deviations from the SM is a flipped sign of the top-Yukawa coupling, the so-called *inverted top coupling* (ITC) scenario. This scenario was already investigated in Run I of the LHC by the CMS Collaboration with different decay channels of the Higgs boson [64, 313–315] and with a preliminary result at Run II with the data obtained in 2015 [316] and the $H \to b\bar{b}$ final state, but the ITC scenario could neither be confirmed nor excluded yet. Given the amount of data to analyze, which is more than 15 times increased compared to the result of 2015, a significant improvement with respect to this result can be expected.

The search for the associated production of a Higgs boson with a single top quark with the data set of 2016 is presented in this chapter. The first section starts with an outline of the analysis strategy. A discussion of the event topology, including the signal and background processes, is provided in the second section, followed by the event selection in the third section. The fourth section is dedicated to the simulation of events and the necessary corrections to describe the data accurately. In the fifth section, the full event reconstruction is explained, where the jets are assigned to the final-state quarks for different event interpretations. Based on this reconstruction, an event classification is applied, described in the sixth section. In the seventh section, all systematic uncertainties affecting the search are discussed. The results are given

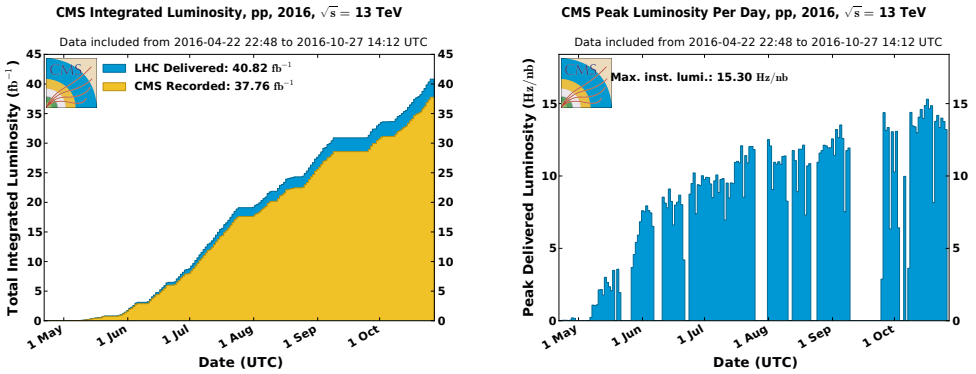


Figure 8.1.: The cumulated integrated luminosity of the data taken by the CMS experiment in 2016 (a) and the instantaneous luminosity (b) [125]. Due to problems with different detector components in different runs, the total integrated luminosity available for physics analysis is 35.9 fb^{-1} instead of 37.8 fb^{-1}.

and interpreted in the eighth section. A combination of the analysis with other Higgs boson decay channels is explained in the ninth section. A summary of the search and an outlook are provided in the tenth and final section.

8.1. Analysis Strategy

The analysis described in this chapter is optimized to search for the associated production of a Higgs boson and a single top quark at a center-of-mass energy of $\sqrt{s} = 13$ TeV. In contrast to the previous chapter, the analysis is performed only on data taken in the year 2016. The integrated and instantaneous luminosity of this year are shown in Fig. 8.1. The data of 2016 corresponds to an integrated luminosity of 35.9 fb^{-1} [298]. Despite this large amount of recorded data, the expected number of signal events is still very low due to the cross section of the two signal processes, i. e., the associated production of a Higgs boson with a single top quark in the t channel (tHq) and associated tW single top quark plus Higgs boson production (tHW), assuming SM couplings. The cross sections in that case are $\sigma_{\text{tHq}} = 71.0\,^{+2.9}_{-4.8}$ (scale) ± 0.1 (PDF) fb and $\sigma_{\text{tHW}} = 15.6\,^{+0.7}_{-1.0}$ (scale) ± 0.4 (PDF) fb for LO accuracy obtained with the MadGraph5_aMC@NLO event generator [317]. As discussed in Section 1.6, the cross sections of both signal processes depend on the coupling parameters κ_t and κ_W. For this analysis, the coupling of the Higgs boson to the different vector bosons is assumed to be equal ($\kappa_V = \kappa_W$). An illustration of the enhanced production cross section for both processes dependent on coupling modifications of κ_t and κ_V is provided in Fig. 8.2. As already pointed out, the most interesting scenario to study is the ITC scenario, where the top-Yukawa coupling is modified with a minus sign, i. e., $\kappa_t = -1$. This will enhance the production cross sections to $\sigma^{\text{ITC}}_{\text{tHq}} = 792.7\,^{+21.4}_{-30.9}$ (scale) ± 7.1 (PDF) fb and $\sigma^{\text{ITC}}_{\text{tHW}} = 147.2\,^{+2.9}_{-2.6}$ (scale) ± 5.7 (PDF) fb, which is roughly one order of magnitude higher for both processes compared to the SM coupling scenario. The associated production of a Higgs boson with top quark pairs ($\text{t}\bar{\text{t}}\text{H}$) is only sensitive

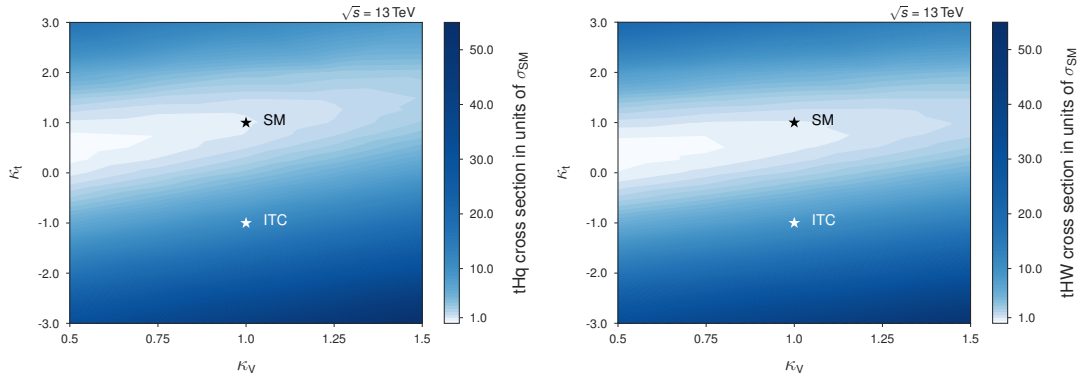


Figure 8.2.: Cross sections of the tHq (a) and tHW (b) signal processes as a function of the coupling modifiers κ_t and κ_V with respect to the SM cross sections. The inverted top coupling (ITC) scenario, where the the top-Yukawa coupling of the Higgs boson to the top quark is modified with a minus sign, i. e., $\kappa_t = -1$, results in a strong enhancement of both cross sections.

to the magnitude of coupling modifications $|\kappa_t|$, therefore, the cross section is the same for the SM and ITC scenarios $\sigma_{t\bar{t}H} = \sigma^{ITC}_{t\bar{t}H} = 506.5\ ^{+29.3}_{-46.6}$ (scale) $^{+20.8}_{-18.2}$ (PDF + α_s) fb. Beyond the direct search for tH production, this analysis also searches for much more pronounced deviations of κ_t and also κ_V. In these cases, the $t\bar{t}H$ process is also considered as a third signal process, since coupling modifications with $|\kappa_t| \neq 1$ also affect the production cross section of the $t\bar{t}H$ process. In addition, the analysis also searches for a possible scenario in which the Higgs boson is composed of CP-even and CP-odd components, parametrized by the mixing angle α, as shown in Eq. 1.47. For a mixing angle of $\alpha = 0°$, the full CP-even SM Higgs boson is recovered, while a value of $\alpha = 180°$ effectively corresponds to the ITC scenario of a flipped sign for the top-Yukawa coupling. Since all other values for the mixing angle introduce a mixture of CP-even and CP-odd components, they cannot be related to a coupling modification based on the κ framework.

In the analysis presented in this chapter, the events are first selected based on the final-state objects of the signal processes. After applying corrections to simulated events in order to improve the description of the data, a dedicated event reconstruction is applied. In this event reconstruction step, different BDTs are used to reconstruct every event under three different hypotheses: tHq signal process, tHW signal process and top quark pair background process. Each reconstruction hypothesis provides distinctive variables, which are used afterwards in an event classification step to separate signal and background processes, also based on BDTs. In parallel to this, another set of BDTs is trained in an orthogonal dileptonic region to distinguish different background components. Exclusion limits are then extracted for tH production in the SM and ITC scenarios, different coupling modifications and CP-mixing scenarios by a simultaneous maximum-likelihood fit to the output distributions of these different classification BDTs. The workflow of the analysis is also illustrated in Fig. 8.3.

In total, there are 51 different combinations of κ_t and κ_V coupling modifications and 21 different

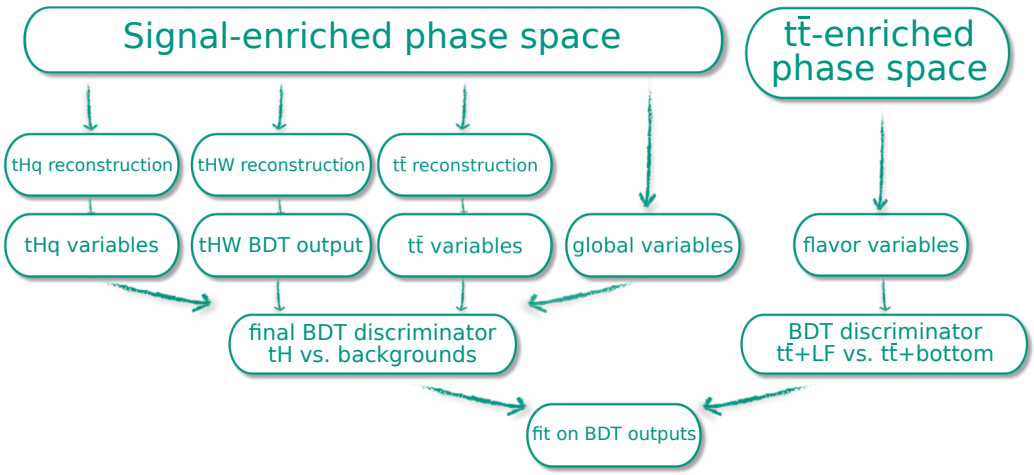


Figure 8.3.: Scheme of the analysis workflow. Events in the signal-enriched phase space are reconstructed under three different event interpretations before being classified as signal- or background-like. A similar classification is performed in a background-dominated sideband. Both distributions are used in a simultaneous maximum-likelihood fit to extract exclusion limits for associated production of a Higgs boson with a single top quark and coupling modifications.

CP-mixing scenarios available in the simulation. For the κ factors, 17 different values of κ_t are evaluated ($\kappa_t = \{-3, -2, -1.5, -1.25, -1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 2, 3\}$) and three different values for κ_V ($\kappa_V = \{0.5, 1, 1.5\}$). Each of these combinations represents a scenario with different production cross sections of the signal processes, but in addition, also the kinematic properties of the final-state objects change with different coupling modifications. However, the change of the kinematic properties of the Higgs boson and the top quark is only caused by different ratios of the coupling modifications κ_t/κ_V and not by the individual values of κ_t and κ_V as for the cross section of the processes. As the exclusion limits in this analysis are given on the cross section times the branching ratio, the difference in the production cross sections for certain κ_t and κ_V scenarios cancels out for they have the same κ_t/κ_V ratio. Therefore, the exclusion limits are only given for the 33 unique κ_t/κ_V ratios of coupling modifications. In case of the CP-mixing scenario, the exclusion limits are also given on the cross section times branching ratio, but in steps of 0.1 for $\cos\alpha$ from -1 to 1.

8.2. Event Topology

The first step of the analysis is to identify a possible selection of events to reject as many events from background processes as possible, while selecting most of the events from the signal processes. To achieve this goal, the final-state topology of signal and background processes provides a first hint for such a selection. The corresponding Feynman diagrams for the signal and background processes are shown in Figs. 8.4, 8.5 and 8.6. This section introduces all relevant signal and background processes for this analysis and their respective final states. All other processes are either completely rejected by the event selection, described in the following section, or their production cross section is insignificant.

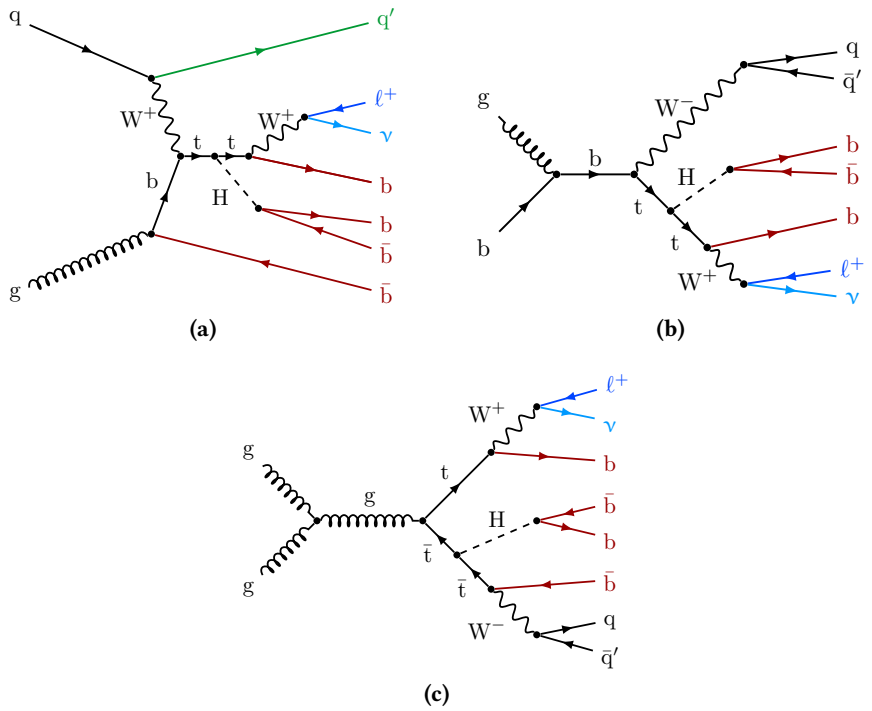


Figure 8.4.: Examples of Feynman diagrams for all three different signal processes in which a Higgs boson is either produced with a single top quark ((a) and (b)) or with top quark pairs (c). The Higgs boson decays to a bottom quark-antiquark pair, while exactly one W boson decays into a charged lepton and the corresponding neutrino.

8.2.1. Signal Processes

The analysis described in this chapter is optimized to search for the associated production of a Higgs boson with a single top quark, i. e., with t-channel single top quark production (tHq) and associated tW single top quark production (tHW). Nevertheless, the associated production of a Higgs boson with top quark pairs ($t\bar{t}H$) is also considered as a signal process for the studies of Higgs boson couplings, since modifications of the coupling modifiers and a possible CP-mixing affect this process in a similar way.
Therefore, the two main signal processes of this analysis are the two single top quark production modes, namely the t-channel production, modeled in the 4FS (Fig. 8.4(a)) and the associated tW production in the 5FS (Fig. 8.4(b)). Similar to the cross section measurement of the t-channel single top quark production in Chapter 6, the tHq process has the advantage of the light-flavored quark recoiling against the top quark, which is a distinct feature of this channel. Only the leptonic decay of the top quark is considered, as the QCD multijet event background in the hadronic decay mode would be too overwhelming compared to the signal processes. Since only Higgs boson decays to a bottom quark-antiquark pair are considered, the final state of the tHq process consists of a charged lepton and missing transverse momentum from the W boson decay, a light-flavored jet from the recoiling quark and either three or four jets stemming from bottom quarks, depending on whether the jet from the initial gluon passes the detector acceptance.

The situation for the tHW process is more difficult, as the final state is similar to the final state of the top quark pair production, missing only one additional bottom quark in the final state. This is also the reason for modeling the tHW process in the 5FS, as modeling the process in the 4FS would result in the same final state as the associated production of Higgs boson with top quark pairs in LO. This introduces interference between the tHW and $t\bar{t}H$ processes, similar to the interference between the associated tW single top quark production and the top quark pair production. Since there are two W bosons in the final state of the tHW process, one is considered to decay leptonically and the other to decay hadronically in order to mimic the final-state topology of the tHq process. In summary, the final state consists of a charged lepton, missing transverse momentum, two light-flavored jets from the hadronically decaying W boson and three jets from bottom quarks, one from the top quark decay and two from the Higgs boson decay.
The final state of the $t\bar{t}H$ process, as shown in Fig. 8.4(c), is similar to the final state of the tHW process if also one W boson decays hadronically and the other W boson decays leptonically, but with one additional jet stemming from a bottom quark. Therefore, the final state of the $t\bar{t}H$ process consists of a charged lepton, missing transverse momentum, two light-flavored jets and four jets stemming from bottom quarks, two from the top quark decays and two from the Higgs boson decay.

8.2.2. Background Processes

No search for rare signal processes can be performed without considering all possible SM background contributions, which will be discussed in the remainder of this section. Exemplary Feynman diagrams for all relevant background processes are provided in Fig. 8.5 and Fig. 8.6.

Top Quark Pair Production

The most important background of this search arises from top quark pair production. In particular relevant for the signal-enriched phase space is the semileptonic decay of top quark pairs, where one W boson decays leptonically and the other W boson decays hadronically. A Feynman diagram of this final-state decay is shown in Fig. 8.5(a). Non-negligible contributions also arise from the dileptonic decay mode, while the fully hadronic decay mode can be rejected by requiring a charged lepton with sufficiently large transverse momentum. Compared to the analyses in the two previous chapters, the top quark pair production can be rejected in principle as the jet multiplicity is lower compared to the single top quark signal processes. Unfortunately, additional initial- and final-state radiation, as well as NLO contributions, give rise to additional jets in top quark pair production events. An example of final-state radiation is given in Fig. 8.5(b), where the additional gluon splits into a bottom quark-antiquark pair. This is an irreducible background, as the corresponding final state is the same as the final state for the $t\bar{t}H$ signal process. To better model this important background, the contribution from top quark pair production is split into different categories, based on the possible additional jets in the events. Of interest here are especially the heavy flavor (HF) components, as they cannot be completely rejected by utilizing b tagging:

- **$t\bar{t}$+$b\bar{b}$:** Events with at least two additional jets, of which each consists of a b hadron.

- **$t\bar{t}$+2b:** Events with only one additional jet containing at least two b hadrons.
- **$t\bar{t}$+b:** All events with exactly one additional jet that could be matched to exactly one b hadron.
- **$t\bar{t}$+c$\bar{c}$/2c/c:** The three former categories combined, but with c hadrons instead of b hadrons.
- **$t\bar{t}$+LF:** Every event not satisfying any of the conditions above, i. e., events without additional jets or with light-flavored (LF) jets.

Single Top Quark Production

Another important background contribution arises from single top quark production without the associated Higgs boson. Similar to top quark pair production, additional radiation can mimic the final-state topology of the signal processes and introduces an irreducible background. The single top quark production modes considered in this analysis are t-channel production (Fig. 8.5(c)) and associated tW production (Fig. 8.5(d)). The modeling of both processes is the same as for the two single-top related signal processes, i. e., the t-channel process is modeled using the 4FS and the associated tW production is modeled with the 5FS. The single top s-channel production mode is not considered as a background, since the applied event selection is able to reject all contributions from this background component in a sufficient way.

Production of Electroweak Bosons in Association with Jets

The production of electroweak bosons in association with jets can also contribute to the composition of background events. For this analysis, only the Z+jets background is found to give significant contribution after the event selection, mainly in the dileptonic region. For the main signal regions, the contribution of Z+jets events can be heavily suppressed by imposing a selection criterion on the missing transverse momentum. The remaining contribution is negligible. An exemplary Feynman diagram of this process is shown in Fig. 8.6(a). The W+jets background is found to be negligible for all regions of interest in this analysis because of the high jet multiplicity in the signal region and the second lepton in the dileptonic region.

Production of Electroweak Bosons in Association with Top Quark Pairs

Background events can also stem from the production of electroweak bosons with top quark pairs, i. e., with W bosons ($t\bar{t}W$) or Z bosons ($t\bar{t}Z$) instead of a Higgs boson ($t\bar{t}H$). Especially the $t\bar{t}Z$ production process, as shown in Fig. 8.6(c), introduces an irreducible background when the Z boson decays into a bottom quark-antiquark pair. But also the $t\bar{t}W$ process (Fig. 8.6(b)) gives rise to additional background contributions due to the total number of three W bosons and two jets from bottom quarks.

Production of Z Bosons in Association with Single Top Quarks

Another source of irreducible background is the production of single top quarks with an additional Z boson, especially in the t channel production mode (tZq). In case of a hadronically

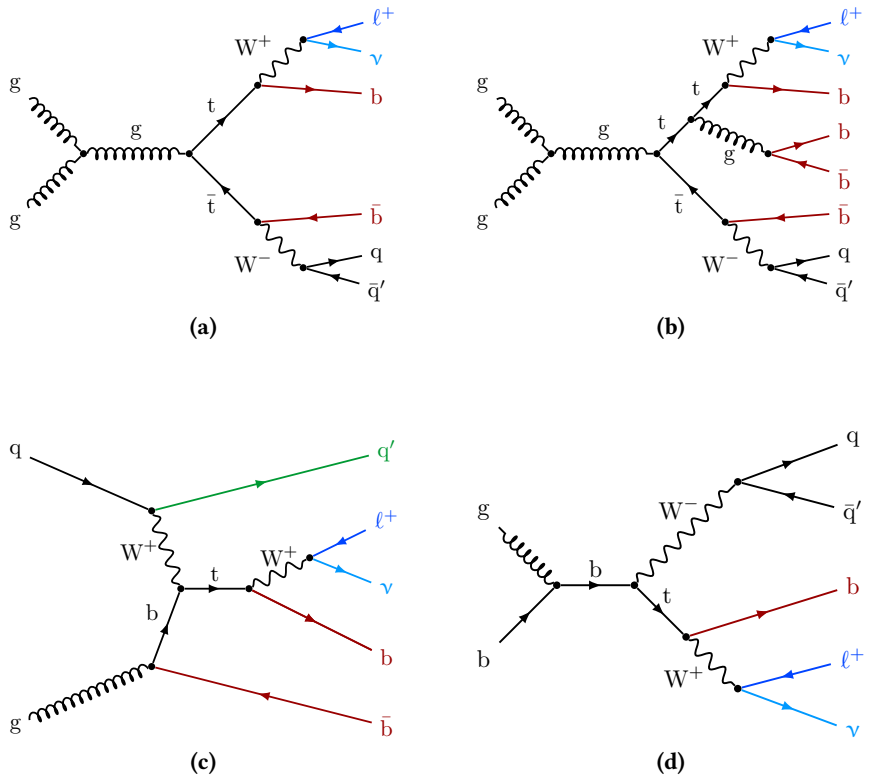


Figure 8.5.: Exemplary Feynman diagrams of the most dominant background processes in the search for tH production. Because of its high cross section, the top quark pair production contributes the most to the background composition, especially in the semileptonic decay mode (a). Furthermore, additional radiation of gluons resembles the final state of the signal processes even more closely (b). For the same reason, background contributions from single top quark production in the t channel (c) and associated tW production (d) are also non-negligible.

decaying Z boson, the final state is similar to the one of the tHq signal process. Assuming the SM coupling scenario, the relative contribution of the tZq process is even larger than for the tHq signal process.

8.3. Event Selection

The event selection is designed to reject most of the background events to enhance the signal-to-background ratio by defining a signal-enriched region. As only one decaying W boson is required for the associated production of a Higgs boson with a single top quark, the selection for the signal-enriched region requires exactly one lepton, the so-called single-lepton region. This single-lepton region is further divided by the number of b-tagged jets in each event, where the regions with exactly three and four b-tagged jets are the two main signal regions. A region with exactly two b-tagged jets serves as a control region to validate the modeling of kinematic variables. In addition, a selection optimized for two leptonically decaying W bosons is applied to derive a signal-depleted control region, the so-called dileptonic region. This dileptonic region

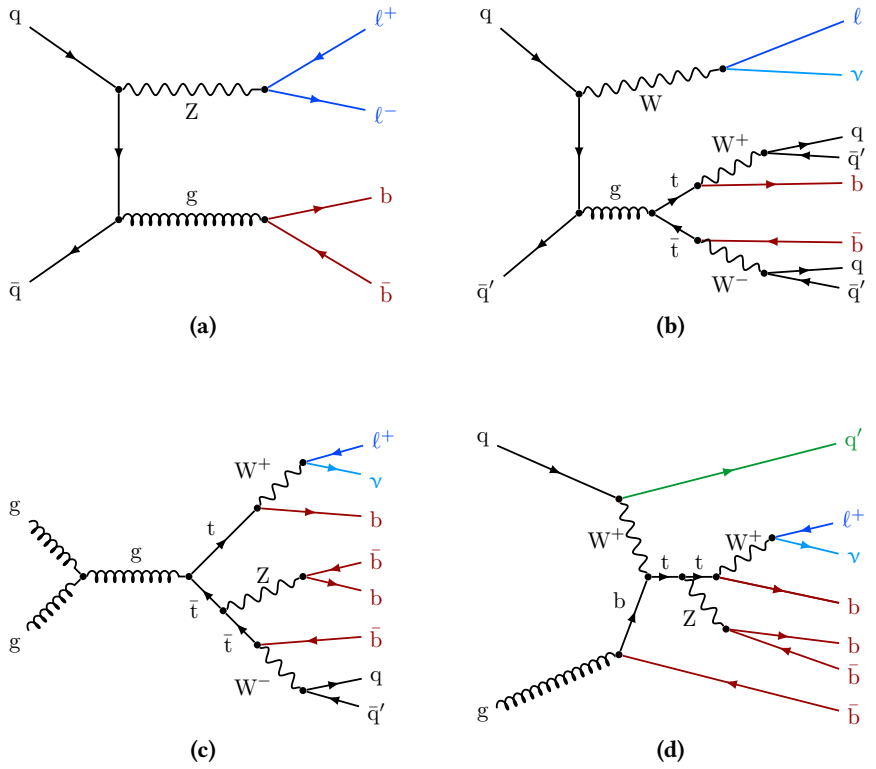


Figure 8.6.: Exemplary Feynman diagrams of minor background processes in the search for tH production. The production of Z bosons in association with jets (a) is a non-negligible background process for the dileptonic event selection. Top quark pairs can also be produced in association with vector bosons, i. e., W bosons (b) and Z bosons (c), in the latter case the final state is similar to the $t\bar{t}H$ process. Similarly, a single top quark can be produced in association with a Z boson in the t channel (d), which has the same final state as the tHq signal process.

aims at constraining the different $t\bar{t}$+HF components, as their modeling in the signal region is similar and only differs due to the additional leptonic W boson decay, which replaces the hadronic W boson decay. The selection criteria for both regions are outlined in the following.

8.3.1. Single Lepton Region

Events passing the single-lepton selection must pass the online HLT for one isolated muon with a p_T of at least 24 GeV or one electron with a p_T of at least 32 GeV and $|\eta| < 2.1$. The different HLT paths used for this selection are listed in Table 8.1. At least one of these HLT paths must have a match for every event. The offline selection criteria for these leptons are slightly stricter to take the turn-on of the HLT into account. Therefore, the muon (electron) that caused the HLT match is required to have a p_T of at least 27(35) GeV. The muon must satisfy the tight ID criterion, defined in Section 5.1.2, and must have a relative isolation of less than 0.15. For electrons, the tight ID needs to be fulfilled (see Section 5.1.3). Electrons within the transition region between the EB and EE ($1.442 < |\eta| < 1.566$) are excluded from the event selection. To reject events

Table 8.1.: List of all HLT paths used for the event selection process in the single-lepton region. Each event must pass at least one of these triggers, which are designed to identify leptons with high transverse momentum in the event.

Single-lepton channel	Trigger
μ	HLT_IsoMu24_v*
	HLT_IsoTkMu24_v*
e	HLT_Ele32_eta2p1_WPTight_Gsf_v*

with more than one leptonically decaying W boson or with leptonically decaying Z boson, an event veto is enforced if additional loose leptons can be identified: events are rejected if they contain one loose muon or one loose electron in addition to the tight lepton described above. Loose electrons are defined by passing the loose electron ID, having a transverse momentum of at least 15 GeV and an absolute pseudorapidity of less than 2.5. Every loose muon must satisfy the loose muon ID selection criterion, $p_T > 10$ GeV and have a relative isolation of less than 0.25.

Selected jets are required to either have a p_T of at least 30 GeV in the central region of the detector ($|\eta| < 2.4$) or $p_T > 40$ GeV if they are located in the forward region ($2.4 \leq |\eta| < 4.7$). Each jet must fulfill the PF jet ID requirements outlined in Section 5.1.4. Jets are rejected if their distance ΔR in the η-ϕ plane is less than 0.4 with respect to the selected tight lepton. To identify jets stemming from bottom quarks, the CSVv2 b tagging algorithm is utilized with its medium working point, as defined in Section 5.1.5. To reject most Z+jets production and possible QCD multijet event background, a lepton-flavor dependent selection on the missing transverse momentum is applied. For events with a tight muon, events are required to have a missing transverse momentum of at least 35 GeV, while for events with tight electrons this requirement is raised to 45 GeV, since the QCD multijet contribution is more pronounced in the electron channel.

Events in the single-lepton region are then categorized according to their number of b-tagged jets. This selection is motivated by the final state of the tHq process. Therefore, the two main signal regions consist of either exactly three (3 tag) or four b-tagged jets (4 tag), depending on whether the additional jet from the gluon splitting falls into the detector acceptance. The overall number of jets per event must be at least one jet higher than the number of b-tagged jets per event to take the light-flavored jet in forward direction into account. An additional region with exactly two b-tagged jets (2 tag) is defined to validate the modeling of variables. Therefore, the number of jets in the 2 tag region is the same as for the 3 tag region, while the different b jet multiplicity ensures the orthogonality of this control region to the signal region. All selection criteria are summarized in Table 8.2. The event yields for the 3 and 4 tag signal region are provided in Table 8.5. Some basic kinematic distributions of reconstructed objects in the 2 tag control region are shown in Fig. 8.7.

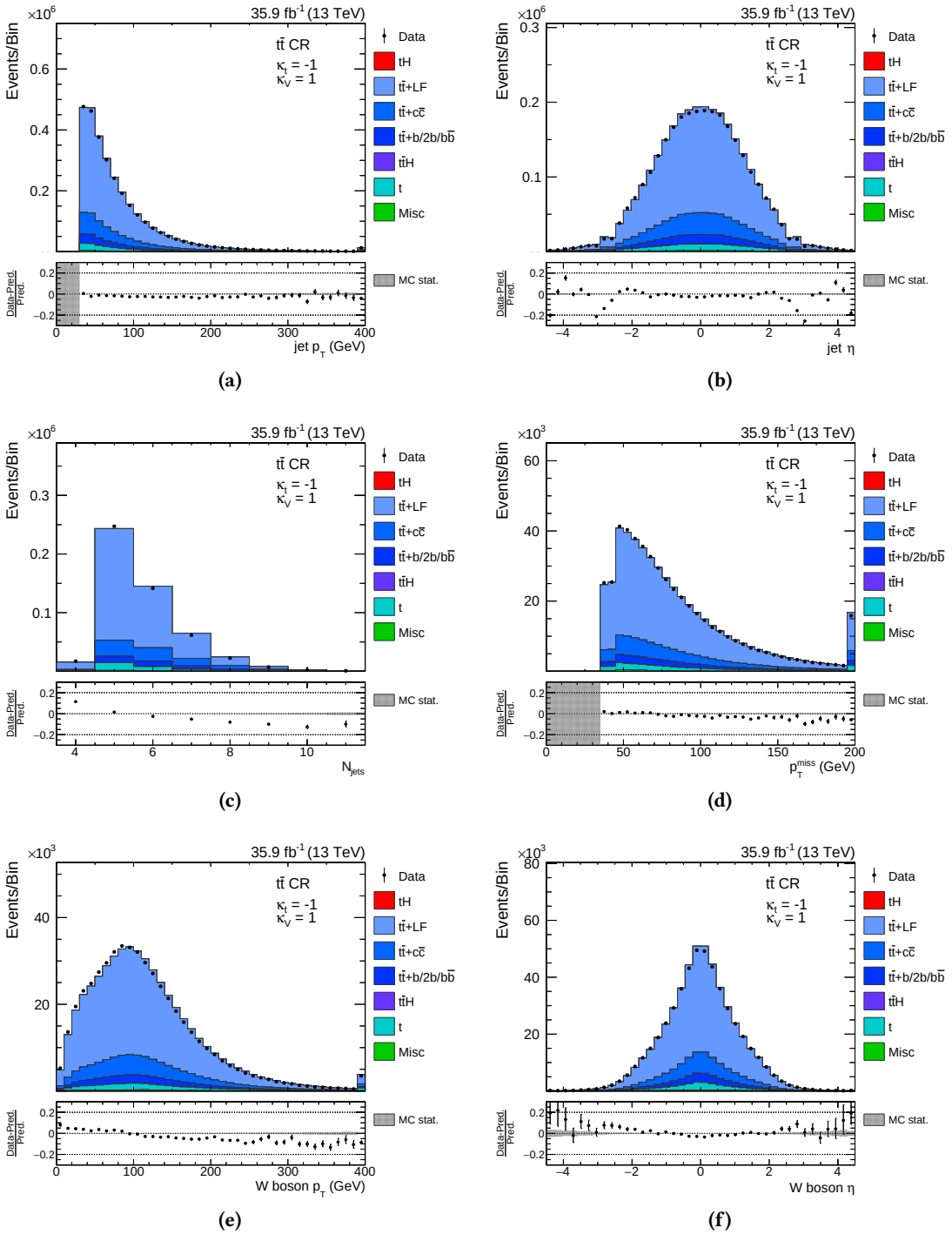


Figure 8.7.: Kinematic distributions of the transverse momentum and pseudorapidity of selected jets ((a) and (b)), the number of selected jets (c), the missing transverse momentum (d), as well as the transverse momentum and pseudorapidity of the leptonically decaying W boson ((e) and (f)) in the 2 tag control region ($t\bar{t}$ CR). The simulation is scaled to the expected number of events.

Table 8.2.: List of all requirements an event has to pass to fall into either the 2 tag control region or the 3 and 4 tag signal regions.

	2 tag control region	3 tag signal region	4 tag signal region
Trigger match	see Table 8.1	see Table 8.1	see Table 8.1
Number of tight leptons	1	1	1
Number of additional loose leptons	0	0	0
Number of selected jets	≥ 4	≥ 4	≥ 5
Number of b-tagged jets	2	3	4
Missing transverse momentum	$> 45/35$ GeV (e/μ)	$> 45/35$ GeV (e/μ)	$> 45/35$ GeV (e/μ)

8.3.2. Dileptonic Region

The dileptonic region is used to distinguish different $t\bar{t}$+HF components and reduce the uncertainty of these processes in the signal region. Thus, the event selection is optimized for dileptonic top quark pair production and not for the signal processes. The HLT selection is more complex compared to that of the single-lepton case, as there are different HLT paths available for each different combination of electrons and muons. All HLT paths utilized for the dileptonic region are listed in Table 8.3. Similarly to the single-lepton selection, the dileptonic region also requires one tight lepton, but in contrast to the previous selection, events with additional loose leptons are not vetoed. Instead events are required to have a loose lepton in addition to the tight lepton. Tight and loose leptons are following the same ID requirements as in the single-lepton selection, but the requirements on the transverse momentum are relaxed due to the lower thresholds of the different HLT paths. For tight (loose) electrons, the p_T is required to be at least 20(15) GeV and no requirement on the pseudorapidity is enforced, while for muons the requirement on the transverse momentum for loose and tight muons is 20 GeV. In general, also two tight leptons can be selected, as the tight electron or muon ID automatically fulfills also the corresponding loose ID. The background contribution from Z+jets events is rejected by imposing a requirement on the missing transverse momentum in the event of at least 40 GeV.
Since the purpose of the dileptonic region is to select dileptonic top quark pair events, which contain exactly two bottom quarks in the final state, only selection criteria on b-tagged jets are applied. Each b-tagged jet must have a p_T of at least 30 GeV and must be located in the central region of the detector to make use of b tagging. Each event in the dileptonic region must have at least two b-tagged jets according to the medium CSVv2 and at least one additional b-tagged jet according to the loose CSVv2 working point. The selection based on the medium working point also includes the loose working point, as outlined in Section 5.1.5. The motivation for the additional loose b-tagged jet criterion is to result in an equal amount of different $t\bar{t}$+HF compo-

Table 8.3.: List of all HLT paths used for the event selection in the dileptonic region. The different trigger paths account either for a lepton pair of the same flavor (ee or μμ) or opposite flavor (eμ or μe), in which case different triggers are used dependent on which lepton flavor has higher transverse momentum.

Dileptonic channel	Trigger
ee	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_vX
μμ	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_vX
	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_vX
eμ	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_vX
	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_vX
	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_vX
	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_vX
μe	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_vX
	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_vX

nents in the dileptonic region. All selection criteria for the dileptonic region are summarized in Table 8.4. The event yields for the dileptonic region are provided in Table 8.5.

8.4. Simulation of Events and Corrections

The simulated events for both single top signal processes, i. e., the tHq and tHW processes, are generated with the event generator MadGraph5_aMC@NLO version 2.2.3 [172]. The different coupling scenarios and the resulting changes of the final-state object kinematics properties normally require a dedicated sample for each scenario. To avoid the demands of such high resources, an alternative approach is chosen. Both simulation samples are generated for the ITC scenario and all other scenarios are recovered by applying an LHE event weight, corre-

Table 8.4.: Summary of event selection requirements for the dileptonic region.

	Dileptonic control region
Trigger match for	see Table 8.3
Number of tight leptons	1
Number of additional loose leptons	1
Number of b-tagged jets (medium WP)	≥ 2
Number of b-tagged jets (loose WP)	≥ 3
Missing transverse momentum	40 GeV

Table 8.5.: Predicted and observed event yields for the 3 and 4 tag signal region, as well as the dileptonic control region, after applying the event selection. The uncertainties on the predictions include the complete set of uncertainties considered in the search, i. e., the statistical and all systematic uncertainties.

	3 tag signal region	4 tag signal region	dileptonic control region
$t\bar{t}$+LF	24 127 ± 5812	320 ± 181	5248 ± 998
$t\bar{t}$+$c\bar{c}$	8521 ± 4869	339 ± 256	2084 ± 1204
$t\bar{t}$+b	3946 ± 2116	183 ± 113	766 ± 427
$t\bar{t}$+$b\bar{b}$	4115 ± 2265	777 ± 429	745 ± 436
$t\bar{t}$+2b	2299 ± 1148	138 ± 88	401 ± 228
Single top	1979 ± 353	78.4 ± 25.8	285 ± 37
$t\bar{t}Z$	202 ± 30	32.0 ± 6.6	54.8 ± 7.3
$t\bar{t}W$	90.3 ± 22.8	4.2 ± 2.8	31.4 ± 5.9
tZq	28.3 ± 5.7	2.9 ± 2.3	—
Z+jets	—	—	69.0 ± 31.5
Sum of backgrounds	45 308 ± 8279	1875 ± 551	9684 ± 1695
tHq (SM)	11.1 ± 3.3	1.3 ± 0.3	0.31 ± 0.08
tHW (SM)	7.6 ± 1.1	1.1 ± 0.3	1.4 ± 0.2
$t\bar{t}H$ (SM)	268 ± 31	62.0 ± 9.9	48.9 ± 5.9
tHq (ITC)	160 ± 38	19.1 ± 5.2	3.9 ± 1.0
tHW (ITC)	91.9 ± 11.9	13.7 ± 2.3	17.6 ± 2.2
$t\bar{t}H$ (ITC)	268 ± 31	62.0 ± 9.9	48.9 ± 5.9
Observed	44 311	2035	9065

sponding to the desired scenario. The effect of this reshaping on the kinematic distributions is illustrated in Appendix C.2 for the simulated tHq and tHW sample. The correlation between the transverse momentum and the pseudorapidity of the Higgs boson and the top quark can be found in Appendix C.3, both for the ITC and the reweighted SM scenario. The $t\bar{t}H$ signal process and the top quark pair production background are modeled with POWHEG version 2 [176, 177, 318]. For the simulation of single top backgrounds, POWHEG has been used as well, version 2 for the t-channel process [259] and version 1 for the associated tW production [244]. The remaining background processes, namely $t\bar{t}W$, $t\bar{t}Z$, Z+jets and tZq, have been simulated with MADGRAPH5_AMC@NLO. All processes involving top quarks are generated with a top quark mass of 172.5 GeV. The parton shower and the hadronization are modeled using PYTHIA version 8.2 [180]. The underlying event is either modeled with the tune CUETP8M1 [181, 182] or, in case of the top quark pair background, with tune CUETP8M2T4, which provides significant improvements to the modeling of top quark pair events. For all processes, the NNPDF30_-

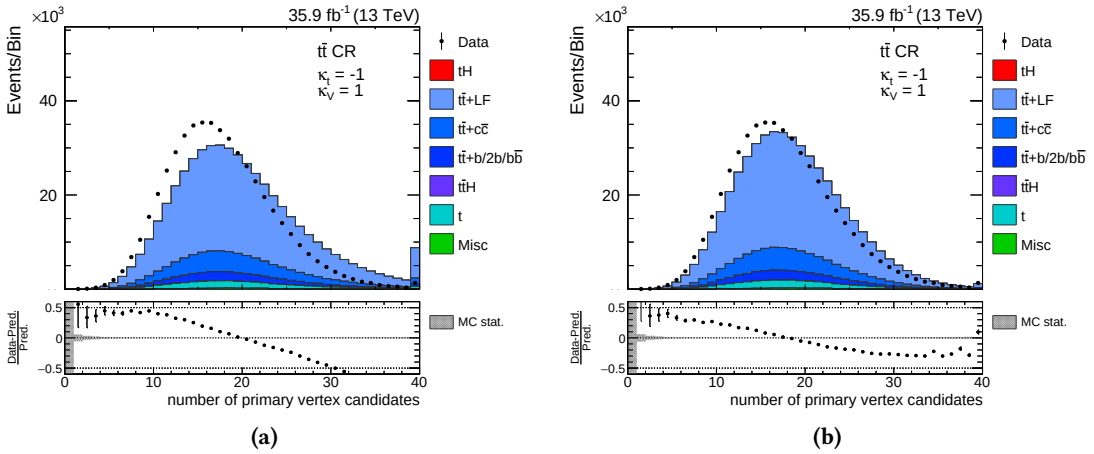


Figure 8.8.: The distribution of the number of primary vertices is different for simulation and data (a) in the 2 tag control region ($t\bar{t}$ CR). To correct this disagreement, the simulated events are reweighted to match the observed distribution (b). A slight improvement is achieved, but disagreement can still be observed. The simulation is scaled to the integral of data.

NLO_AS_0118 PDF set has been used [159]. Pileup effects from multiple interactions at the same bunch crossing are added on top to each simulation. The full list of simulation samples for this analysis is provided in Appendix C.1.

8.4.1. Number of Pileup Interactions

The additional pileup interactions per event in the simulation are defined in advance to the data-taking period with a preliminary estimate. As pileup interactions usually involve only soft interactions, the modeling of objects with high transverse momentum is unaffected. Only certain variables, such as the number of primary vertices, show an observable difference in simulation and data. Therefore, events are reweighted based on an event level with a pileup profile obtained from minimum-bias data with a cross section of 69.2 mb [245, 246]. The effect of this reweighting on the distribution of the number of primary vertices is shown in Fig. 8.8. A slight improvement compared to the unweighted distribution is observed. Yet, the agreement is not perfect and a mismodeling of this particular variable remains. The agreement between data and simulation can be improved by changing the minimum-bias cross section to higher values. However, this method is disfavored by the Luminosity Physics Object Group (LUM POG) [319] of the CMS Collaboration, as the minimum-bias cross section is measured accurately. The remaining differences are still under investigation by the LUM POG.

8.4.2. Lepton Efficiencies

Reconstructed leptons undergo various selection steps, which yield different efficiencies for simulation and data. These differences are corrected by applying dedicated scale factors to events, based on their leptonic content.

For the reconstruction of muons, an overall muon efficiency is calculated. In addition to the already known efficiencies for trigger, isolation and ID, the tracking efficiency is now also taken into account [290]:

$$\epsilon = \epsilon_{\text{Tracking}} \cdot \epsilon_{\text{ID|Tracking}} \cdot \epsilon_{\text{Iso|ID}} \cdot \epsilon_{\text{Trigger|Iso}}. \tag{8.1}$$

This is necessary because of the increased number of pileup interactions in 2016 compared to 2015. All scale factors are provided by the MUO POG with a tag-and-probe method [248–250]. For electrons, the situation is similar, corrections are applied to take the efficiency of the trigger, reconstruction and ID into account. Most scale factors are centrally provided by the E/gamma POG of the CMS Collaboration, except for the single electron trigger in this analysis. The scale factors for the `HLT_Ele32_eta2p1_WPTight_Gsf_v*` trigger are measured specifically for various single top analyses [292, 293]. The results of the lepton efficiency corrections for muons and electrons are shown in Fig. 8.9.

8.4.3. b Tagging Efficiencies

Since this analysis heavily relies on b-tagged jets, the employed b tagging algorithm needs to describe simulation and data accurately. To correct for different efficiencies in simulation and data, a dedicated reweighting approach is used. Compared to the method described in the two previous chapters, this approach aims to correct the entire distribution of the b tagging algorithm in simulation. This is necessary, as the analysis utilizes the distribution of the classifier to separate different event hypotheses or signal processes from background processes. Event-based scale factors are derived with a tag-and-probe method [320, 321]. The correction factors for heavy-flavored jets, dependent on the transverse momentum and the pseudorapidity of each jet, are defined by

$$\text{SF}_{\text{HF}}(p_{\text{T}}, \eta, \text{CSV}) = \frac{\text{Data} - \text{MC}_{\text{LF}}}{\text{MC}_{\text{HF}}}, \tag{8.2}$$

and equivalent for light-flavored jets

$$\text{SF}_{\text{LF}}(p_{\text{T}}, \eta, \text{CSV}) = \frac{\text{Data} - \text{MC}_{\text{HF}}}{\text{MC}_{\text{LF}}}. \tag{8.3}$$

These jet-based scale factors are obtained from a dileptonic $\text{t}\bar{\text{t}}$-enriched control region for heavy-flavored jets or from Z+jets events for light-flavored jets. Each jet in simulated events receives a scale factor based on its hadron flavor, in particular SF_{HF} is applied to jets from b hadrons and SF_{LF} to jets from light-flavored jets. No correction is applied to jets from c hadrons. The final event weight is then obtained by the product of all selected jets:

$$\text{SF} = \prod_{i}^{N_{\text{jets}}} \text{SF}_{\text{jet},\,i}. \tag{8.4}$$

The effect of this reweighting procedure on the output distribution of the CSVv2 b tagging algorithm is shown in Fig. 8.10.

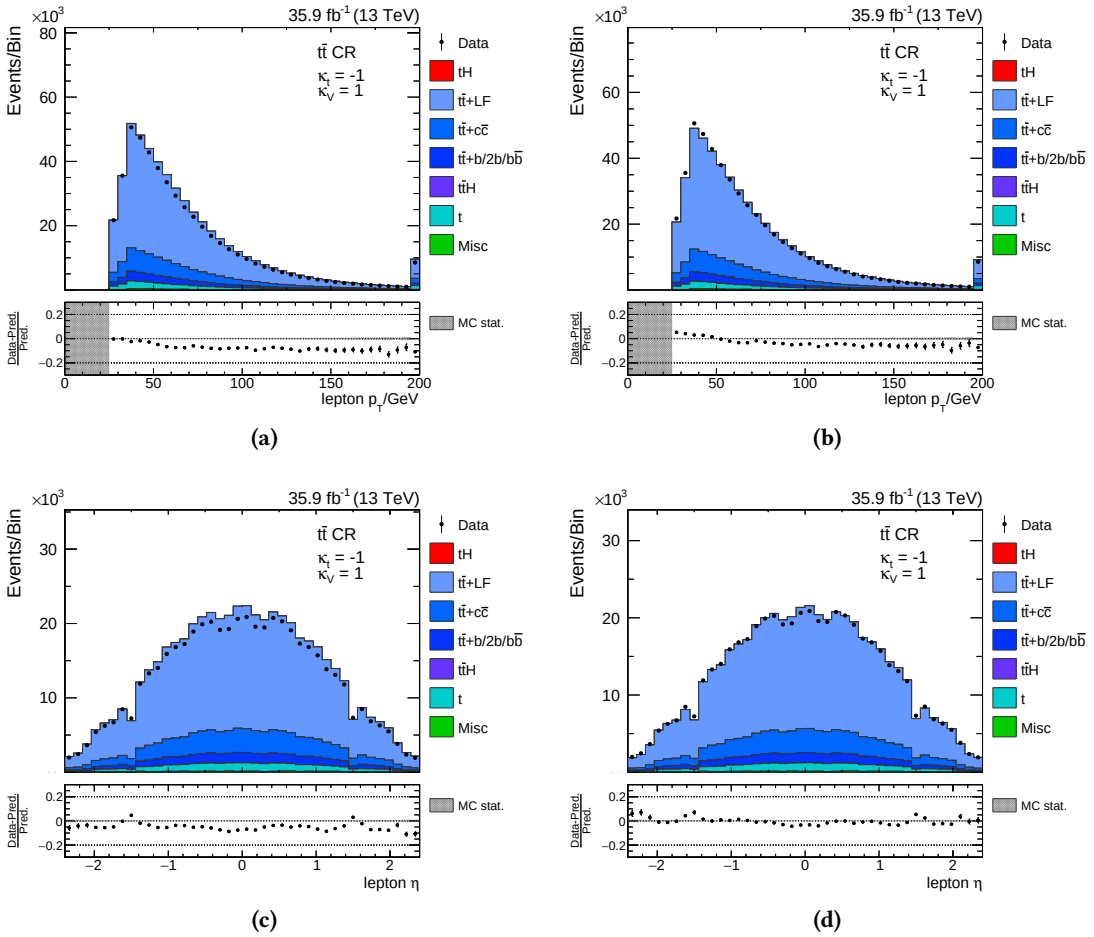


Figure 8.9.: The effect of correcting the differences in lepton efficiencies between simulation and data in the 2 tag control region ($t\bar{t}$ CR). The simulation of the transverse momentum (a) and pseudorapidity (c) of selected leptons is corrected by applying event weights for trigger, isolation, identification and reconstruction efficiencies of leptons. The reweighted distributions show a slight improvement of the modeling of muons (b) and electrons (d). The simulation is scaled to the expected number of events.

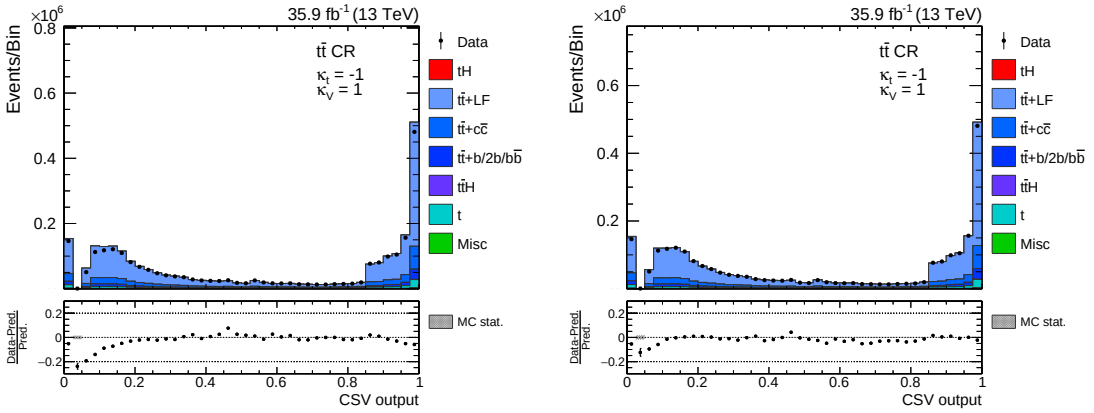


Figure 8.10.: Effect of the b tagging reweighting procedure on the distribution of the b tagging classifier in the 2 tag control region ($t\bar{t}$ CR). The disagreement between data and simulation (a) is corrected by applying jet-dependent scale factors to each jet in an event and calculating an event-based scale factor. The corrected simulation (b) describes the observation adequately.

8.5. Event Reconstruction

After selecting events based on their typical final-state composition, each event has to be fully reconstructed to be able to perform a classification of events and to separate the signal from background processes. Compared to the single top quark analyses described in the two previous chapters, this is not an easy task as the jet-to-quark assignment is not trivial due to the high jet multiplicity. The standard approach to accomplish this task is to reconstruct each possible assignment and to choose the combination that minimizes a χ^2-based metric.
This analysis instead employs a multivariate approach based on BDTs to perform a jet-to-quark assignment. The jet-to-quark assignment is performed for three different interpretations: tHq signal event, tHW signal event or semileptonic $t\bar{t}$ background event. By utilizing different hypotheses for every simulated event and data, characteristic variables can be derived, which strongly depend on the correct jet-to-quark assignment and whether the corresponding variable has a physical meaning. For instance, the invariant mass of the reconstructed Higgs boson from two b jets shows a distinctive peak at around 125 GeV for correct jet-to-quark assignments of simulated tHq and tHW events, but has a broader distribution for wrong assignments of simulated tHq and tHW events and for every possible jet-to-quark assignment of simulated $t\bar{t}$ background events. With this method, reconstruction-dependent variables are derived, which can be used in an event classification to separate signal and background processes.
Each BDT is trained with simulated events of the corresponding hypothesis. In a first step, the reconstructed jets are tried to be matched to the final-state partons for every simulated event based on a ΔR distance criterion. A matchable event is found when each jet can be assigned to a parton with a distance ΔR of less than 0.3. Additional requirements, such as a comparison of the reconstructed transverse momentum of the jet with the transverse momentum of the final-state object, have proven to not yield any improvements [322]. For every matchable event, a correct and wrong interpretation is then extracted. Correct interpretations are defined by the jet-to-quark assignment that minimizes the sum of ΔR distances among all objects in the event and the wrong assignment is chosen by picking a random wrong assignment of this event. Unmatchable events are not used to derive correct and wrong assignments. This results in a data set of correct and wrong event interpretations that serve as 'signal' and 'background' input for the training of the BDT. The BDTs for each event interpretation are trained in the 3 tag signal region. Each BDT is trained exclusively for the ITC scenario. A training for each individual scenario results in a degrade of performance due to the reweighting of events and the associated uncertainty. The settings used for the training of the BDTs for the event reconstruction procedure are summarized in Appendix C.4. The resulting BDT is then applied to all simulated processes and on data. Each possible jet-to-quark assignment for each event is evaluated and the assignment with the highest BDT value is chosen as the interpretation for this event. This is done for all three possible event interpretations independently. The whole event reconstruction technique is illustrated in Fig 8.11 exemplarily for the tHq event interpretation.

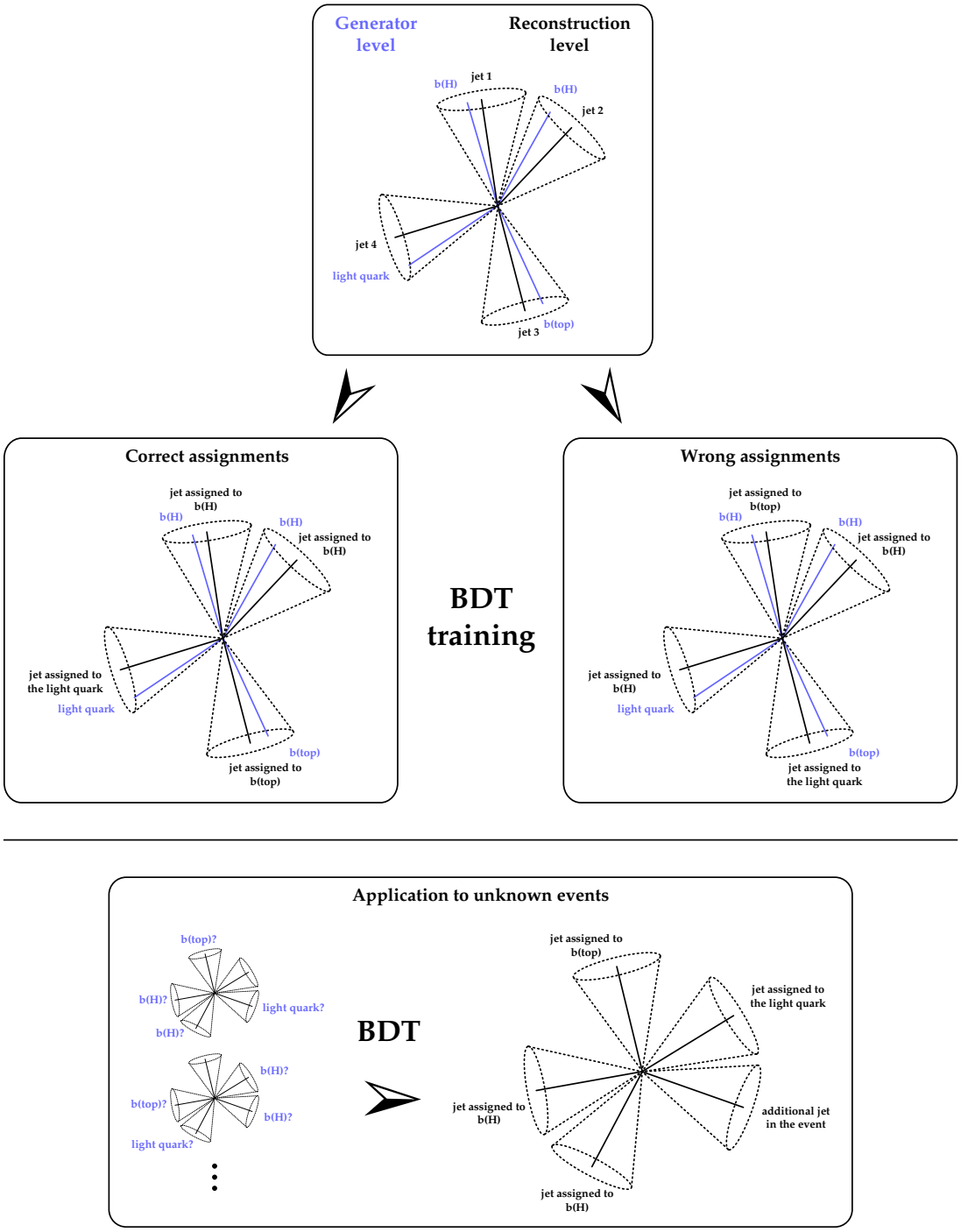


Figure 8.11.: Illustration of the event reconstruction approach. Matchable events are defined in which each reconstructed jet can be assigned to a final-state quark on generator level. Each of these events provides a correct and a randomly chosen wrong jet-to-quark assignment, from which a BDT is trained to distinguish both cases. The BDTs are then applied to unknown events and evaluated for all possible assignments, where the jet-quark assignment with the highest BDT value is chosen for the event.

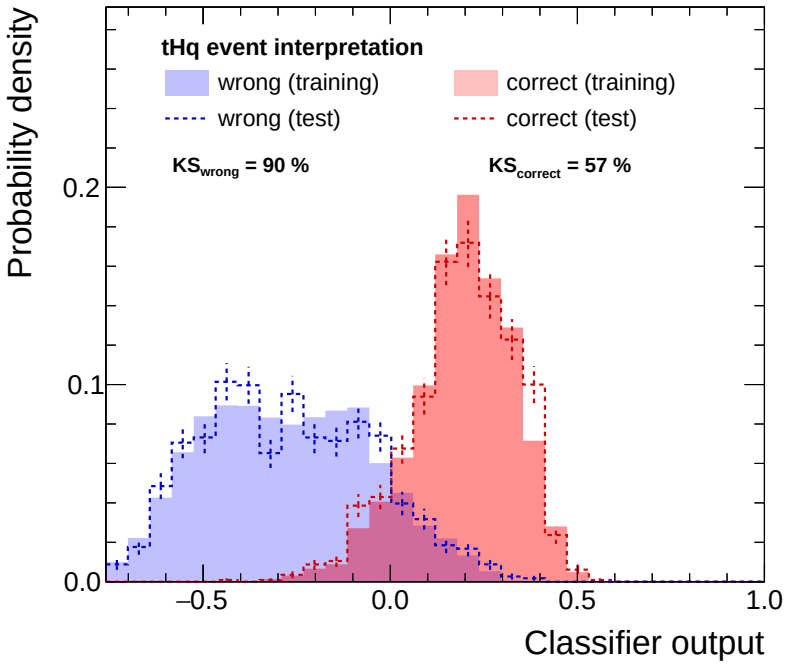


Figure 8.12.: The training of the tHq event reconstruction is tested against a possible overtraining by comparing the result of a training data set to a statistically independent test data set. Similar performance is observed for both data sets, therefore no overtraining is present.

8.5.1. Reconstruction under the tHq Hypothesis

The first event interpretation is the reconstruction under the tHq hypothesis. In the 3 tag signal region, at least four jets are present, exactly three of them b-tagged, which need to be assigned to the final-state quarks of the tHq process. These quarks are the two bottom quarks from the Higgs boson decay, the bottom quark from the top quark decay and the light-flavored quark in forward direction. To reduce the number of possible assignments, jets assigned to bottom quarks must be located in the central region of the detector ($|\eta| < 2.4$) and the jet assigned to the recoiling quark must not be b-tagged. This restriction is fulfilled by 99% of all correct combinations.
The BDT is trained with a total of 15 different input variables, where every variable is sensitive to the given jet-to-quark assignment. The description of all input variables and their ranking in the training of the BDT can be found in Table 8.6. The resulting BDT distribution for correct and wrong assignments is shown in Fig. 8.12. No sign of overtraining is observed. Shape distributions of input variables for correct and wrong jet-to-quark assignments in the tHq event hypothesis are provided in Appendix C.5. The jet-to-quark assignment with the highest BDT value for each event is shown in Fig. 8.13 for the 3 tag and 4 tag signal regions, both for the ITC and SM scenarios.

Table 8.6.: List of all input variables in the tHq event reconstruction. Each variable adds separation power between correct and wrong jet-to-quark assignments. The variables are ranked by their importance in the training.

Variable	Description
log m(Higgs)/GeV	Invariant mass of the reconstructed Higgs boson
cos $\theta(\mathrm{b_t}, \ell)$	Cosine of the angle between the jet assigned to the bottom quark from the top quark decay and the lepton
$\Delta R(\mathrm{b_t}, \ell)$	ΔR between the jet assigned to the bottom quark from the top decay and the lepton
ΔR (Higgs jets)	ΔR between the two jets from the Higgs boson decay
log m(top)/GeV	Invariant mass of the reconstructed top quark
CSV(Higgs jet 1)	Output of the b tagging discriminant for the jet assigned to the Higgs boson with the highest transverse momentum
$\lvert\eta$(light jet)$\rvert$	Absolute pseudorapidity of the light forward jet
CSV($\mathrm{b_t}$)	Output of the b tagging discriminant for the jet assigned to the bottom quark from the top quark decay
$\lvert\eta(\mathrm{b_t})\rvert$	Absolute pseudorapidity of the jet assigned to the bottom quark from the top decay
relative H_T	Fraction of the total transverse momenta (jets, lepton, $p_\mathrm{T}^\mathrm{miss}$) that falls to the b jet from the top quark, Higgs jets and light forward jet
$\lvert\eta$(light jet) $- \eta(\mathrm{b_t})\rvert$	Absolute difference of pseudorapidities of the light forward jet and the b jet from the top quark decay
ΔE (light jet, $\mathrm{b_t}$)	Jet energy difference of the light forward jet and the jet assigned to the bottom quark from the top quark decay
CSV(Higgs jet 2)	Output of the b tagging discriminant for the jet assigned to the Higgs boson with the second highest transverse momentum
log min(p_T (Higgs jets))/GeV	Lower transverse momentum of the two jets assigned to the Higgs boson decay products
$\lvert\eta$(top) $- \eta$(Higgs)$\rvert$	Absolute difference of pseudorapidities of the reconstructed top quark and the reconstructed Higgs boson

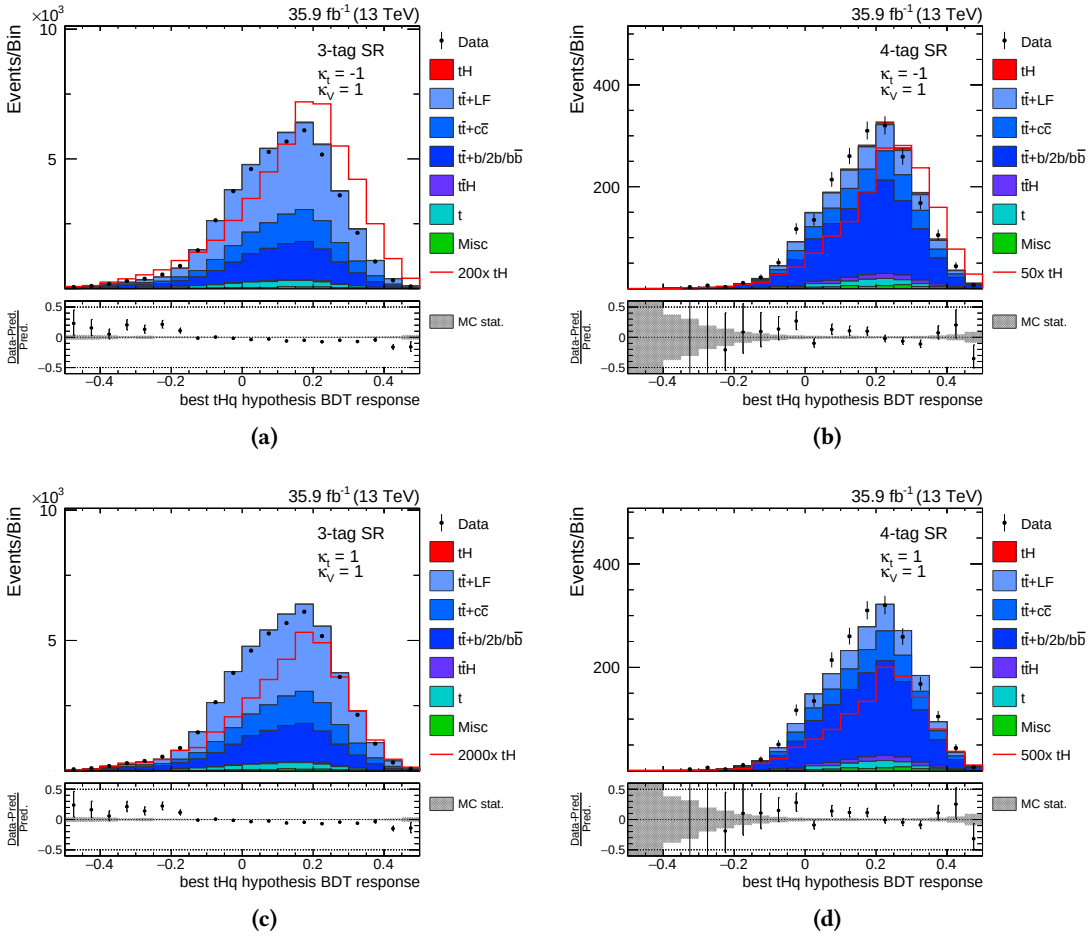


Figure 8.13.: Result of the tHq event interpretation. The best possible jet-to-quark assignment in each event is shown in the 3 tag (a) and 4 tag (b) signal regions for the ITC scenario. The same distributions for the SM scenario ((c) and (d)) are obtained by applying the corresponding LHE weight. The simulation is scaled to the expected number of events. Sufficient agreement between simulation and data is observed.

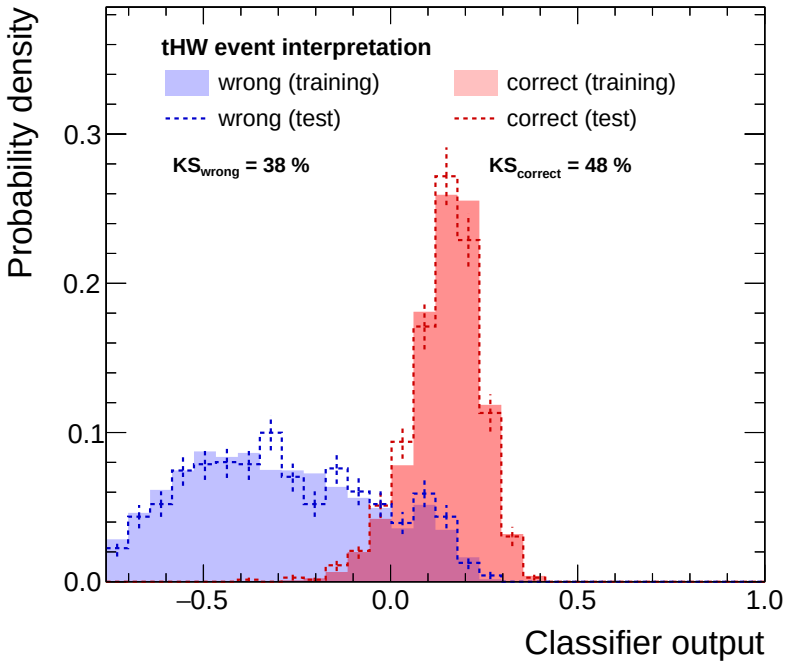


Figure 8.14.: By comparing the result of the tHW event reconstruction BDT obtained for the training data set and a statistically independent test data set, the BDT is checked against possible overtraining, which would result in different distributions. No deviations between both data sets is observed.

8.5.2. Reconstruction under the tHW Hypothesis

The reconstruction of the tHW process is similar to the one of the tHq process. The final state of the tHW process assumes one W boson to decay leptonically and the other W boson to decay hadronically. This results in a final state with three b-tagged jets, one from the top quark decay and two from the Higgs boson decay, and two additional jets from the hadronically decaying W boson. However, not every event in the 3 tag signal region can be fully reconstructed under the tHW event interpretation, as the final state of the tHW process in the 5FS has five quarks in total and the 3 tag signal region requires at least four jets. Although the tHW final state is more complicated and the relative contribution in the two signal regions is smaller, both compared to the tHq processes, a dedicated reconstruction for the tHW process yields a non-negligible improvement of the results [323].
Thirteen different input variables are used for the training of the tHW reconstruction BDT, each reflecting different properties of the chosen jet-to-quark assignment. All input variables are listed in Table 8.7. The distribution of the final classifier is presented in Fig. 8.14, along with the check for possible overtraining. The shapes of all input variables for the training of the BDT are shown in Appendix C.6. Similar to the tHq reconstruction, the jet-to-quark assignment with the highest BDT value in the 3 tag and 4 tag signal regions is shown in Fig. 8.15.

Table 8.7.: All input variables used for the tHW event reconstruction. The variables are ranked according to their importance to the overall separation power.

Variable	Description
log m(Higgs)/GeV	Invariant mass of the reconstructed Higgs boson
ΔR (Higgs jets)	ΔR between the two jets from the Higgs boson decay
log m(top)/GeV	Invariant mass of the reconstructed top quark
log $m(\mathrm{W_b})$/GeV	Invariant mass of the W boson from the top quark production vertex
relative H_T	Percentage of the total transverse momenta (jets, lepton, $p_\mathrm{T}^\mathrm{miss}$) that falls to b jet of the top quark, Higgs jets and light forward jet
CSV(Higgs jet 2)	Output of the b tagging discriminant for the jet assigned to the Higgs boson with the second highest transverse momentum
$\lvert\eta(\mathrm{top}) - \eta(\mathrm{Higgs})\rvert$	Absolute difference of pseudorapidities of the reconstructed top quark and the reconstructed Higgs boson
ΔR ($\mathrm{b_t}$, $\mathrm{W_{lep}}$)	ΔR between the jet assigned to the bottom quark from the top decay and the leptonically decaying W boson
cos θ($\mathrm{b_t}$, ℓ)	Cosine of the angle between the jet assigned to the bottom quark from the top quark decay and the lepton
$\lvert\eta(\mathrm{b_t})\rvert$	Absolute pseudorapidity of the jet assigned to the bottom quark of the top decay
CSV($\mathrm{b_t}$)	Output of the b tagging discriminant for the jet assigned to the bottom quark from the top quark decay
CSV(Higgs jet 1)	Output of the b tagging discriminant for the jet assigned to the Higgs boson with the highest transverse momentum
$\lvert\eta(\mathrm{top}) - \eta(\mathrm{W})\rvert$	Absolute difference of pseudorapidities of the reconstructed top quark and the reconstructed W boson from the top quark production vertex

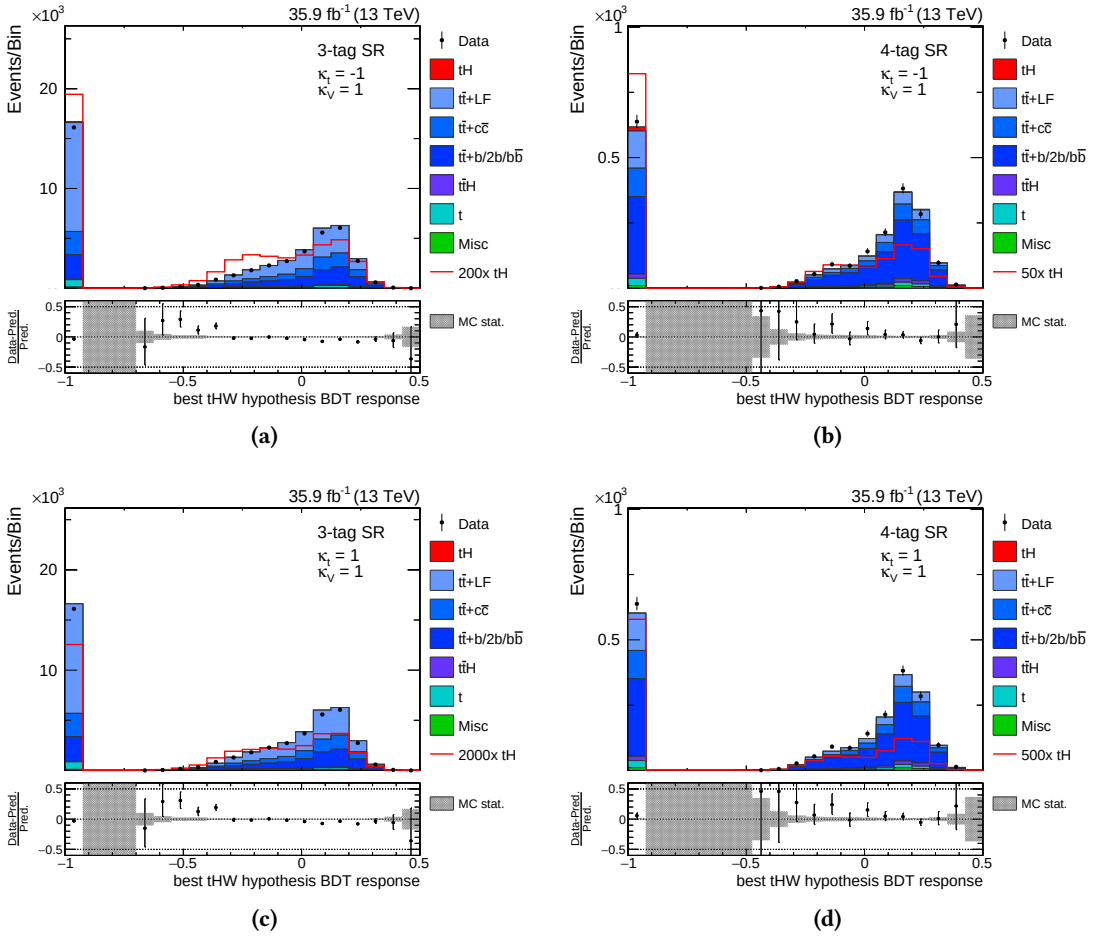


Figure 8.15.: Distribution of the tHW event reconstruction BDT in the 3 tag (a) and 4 tag (b) signal regions for the ITC scenario. The SM scenario ((c) and (d)) is recovered by applying the corresponding LHE weights. The simulation is scaled to the expected number of events. The distributions for simulation and data are in agreement.

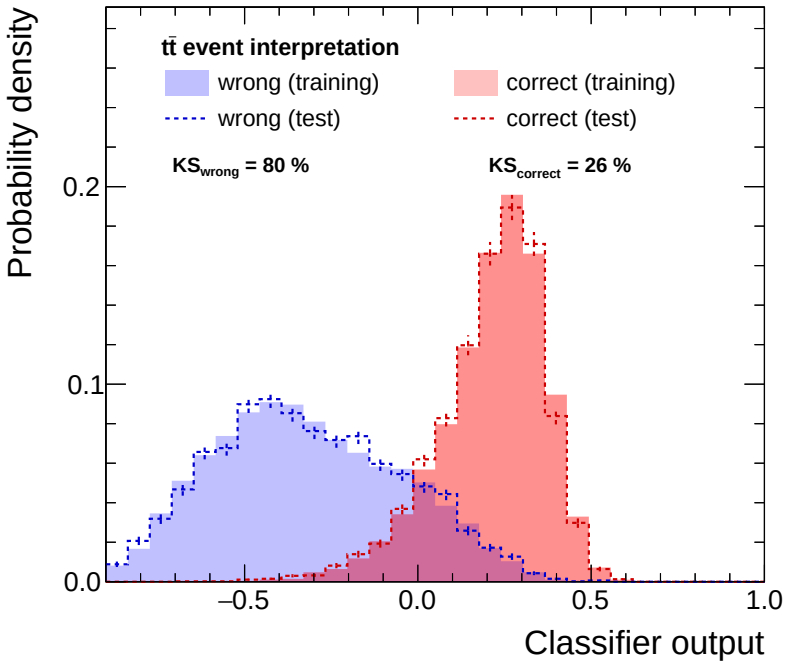


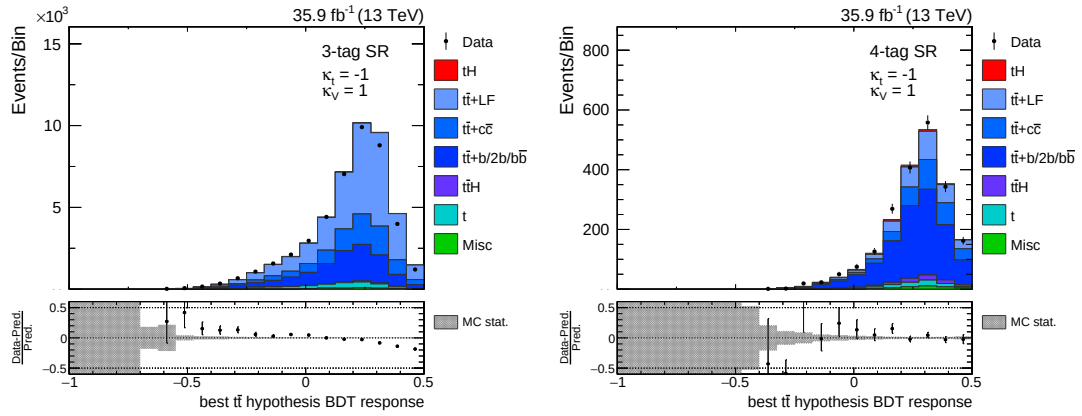
Figure 8.16.: A possible overtraining of the employed BDT is examined by comparing the performance of the top quark pair event reconstruction on the trained data set with the performance on an independent test data set. Similar separation power is observed, therefore no overtraining is observed.

8.5.3. Reconstruction under the Top Quark Pair Hypothesis

The third event hypothesis is not a signal interpretation, but the interpretation of an event to be a top quark pair production event. This is important, as these events are the dominant source of background processes in the signal regions. In particular, the semileptonic decay of the top quark pairs has the largest contribution. Therefore, every event is also reconstructed as a semileptonic top quark pair event, in addition to the tHq and tHW event reconstruction. The final state of this process consists of the two b-tagged jets from both top quark decays and two jets from the hadronically decaying W boson. The charged lepton and the corresponding neutrino of the leptonically decaying W boson are not of interest for the reconstruction, but they are the main reason for the high contribution of top quark pair production to the signal regions. To reduce the number of possible jet-to-quark assignments, jets assigned to the bottom quarks from the top quark decay are required to be b-tagged.
For the training of the BDT, nine different input variables are used to distinguish correct from wrong assignments. They are listed in Table 8.8, ranked by their importance in the training of the BDT. The shape distributions of all input variables can be found in Appendix C.7. The best top quark pair event reconstruction hypothesis for every event is shown in Fig. 8.17 for the 3 tag and 4 tag signal regions. As the reconstruction is independent from the chosen coupling scenario, the same hypothesis is chosen for every scenario.

Table 8.8.: List of all input variables employed in the top quark pair event reconstruction. Each variable provides separation power to distinguish between correct and wrong jet-to-quark assignments. The variables are ranked by their importance in the training of the BDT.

Variable	Description
log ($m(\mathrm{t_{had}})$-$m(\mathrm{W_{had}})$)/GeV	Difference between the invariant masses of the reconstructed hadronically decaying top quark and the reconstructed hadronically decaying W boson
log $m(\mathrm{W_{had}})$/GeV	Invariant mass of the two jets assigned to the reconstructed W boson from the hadronically decaying top quark
log $m(\mathrm{t_{lep}})$/GeV	Invariant mass of the reconstructed leptonically decaying top quark
CSV($\mathrm{W_{had}}$ jet 2)	Output of the b tagging discriminant for the jet with the second highest transverse momentum assigned to the hadronically decaying W boson
ΔR ($\mathrm{b_{t_{had}}}$, $\mathrm{W_{had}}$)	ΔR between the jet assigned to the bottom quark from the hadronically decaying top quark and the reconstructed W boson from the hadronically decaying top quark
relative H_T	Percentage of the total transverse momentum (jets, lepton, $p_\mathrm{T}^\mathrm{miss}$) that falls to the reconstructed hadronically and leptonically decaying top quarks
log $p_\mathrm{T}(\mathrm{t_{lep}})$/GeV	Transverse momentum of the reconstructed leptonically decaying top quark
log $p_\mathrm{T}(\mathrm{t_{had}})$/GeV	Transverse momentum of the reconstructed hadronically decaying top quark
CSV($\mathrm{W_{had}}$ jet 1)	Output of the b tagging discriminant for the jet with the highest transverse momentum assigned to the hadronically decaying W boson



(a) (b)

Figure 8.17.: Distribution of the highest BDT value for the chosen top quark pair event reconstruction in each event in the 3 tag (a) and 4 tag (b) signal regions. The simulation is scaled to the expected number of events. Agreement between simulation and data is observed.

8.6. Event Classification

To be able to make use of the different event reconstruction hypotheses and the characteristic variables, a dedicated BDT has to be trained to separate events into a signal- and background-like category. This is achieved by the classification BDT for the single-lepton region. An additional BDT in the dileptonic region is trained to distinguish different top quark pair background components, as they are the main backgrounds of the analysis. Since the populations in the different regions only depend on the well-understood W boson decay, the dileptonic region can be used to constrain the different $t\bar{t}$+HF components in the final maximum-likelihood fit.

8.6.1. Single Lepton Region

To make use of the different reconstruction hypotheses and their respective variables, the classification BDT is trained with a mixture of variables from the tHq, tHW and top quark pair reconstruction. In addition, also global event variables are used, which do not depend on any particular reconstruction hypothesis. In total, 18 different input variables are used in the training of the BDT, they are listed in Table 8.9. The settings for the training of the BDT can be found in Appendix C.4. As a signal process for the training, a mixture of tHq and tHW events is used, whereas background processes are taken from top quark pair events. The outcome of the training of the classification BDT is shown in Fig. 8.18. No obvious signs of overtraining are observed. The shapes of the three most discriminating variables, along with a comparison of simulated events and data, are provided in Fig. 8.19. The remaining distributions are shown in Appendix C.8. A comparison of simulated events and data of the final BDT distribution is shown in Fig. 8.20 for the 3 tag and 4 tag signal regions, both for the ITC and SM scenarios.

Table 8.9.: List of input variables for the single-lepton event classification BDT. The list consists of global event variables, independent of any reconstruction hypothesis, and reconstruction-dependent variables derived from the tHq, tHW and top quark pair reconstruction hypotheses. The variables in each category are ordered according to their importance for the training.

Variable	Description
$\log m_3$/GeV	Invariant mass of the three jets that result in the highest transverse momentum when combined
aplanarity	Aplanarity of the event [324]
Fox-Wolfram #1	First Fox-Wolfram moment of the event [296, 297]
$q(\ell)$	Electric charge of the lepton
$\log m(\mathrm{t_{had}})$/GeV	Invariant mass of the reconstructed hadronically decaying top quark
CSV($\mathrm{W_{had}}$ jet 2)	Output of the b tagging discriminant for the jet assigned to the hadronically decaying W boson with the second highest transverse momentum
CSV($\mathrm{W_{had}}$ jet 1)	Output of the b tagging discriminant for the jet assigned to the hadronically decaying W boson with the highest transverse momentum
ΔR($\mathrm{W_{had}}$ jets)	ΔR between the two light jets assigned to the hadronically decaying W boson
tHW BDT response	Output of the tHW reconstruction BDT
$\log p_\mathrm{T}$(H)/GeV	Transverse momentum of the reconstructed Higgs boson candidate
$\lvert\eta$(light jet)$\rvert$	Absolute pseudorapidity of the light forward jet
$\log m$(H)/GeV	Invariant mass of the reconstructed Higgs boson candidate
CSV(Higgs jet 1)	Output of the b tagging discriminant for the jet assigned to the Higgs boson candidate with the highest transverse momentum
CSV(Higgs jet 2)	Output of the b tagging discriminant for the jet assigned to the Higgs boson candidate with the second highest transverse momentum
$\cos\theta(\mathrm{b_t}, \ell)$	Cosine of the angle between the b-tagged jet from the top quark decay and the lepton.
$\cos\theta^*$	Cosine of the angle between the light forward jet and the lepton in the top quark rest frame
$\lvert\eta$(H) - η(t)$\rvert$	Absolute pseudorapidity difference of the reconstructed Higgs boson and reconstructed top quark
$\log p_\mathrm{T}$(light jet)/GeV	Transverse momentum of the light forward jet

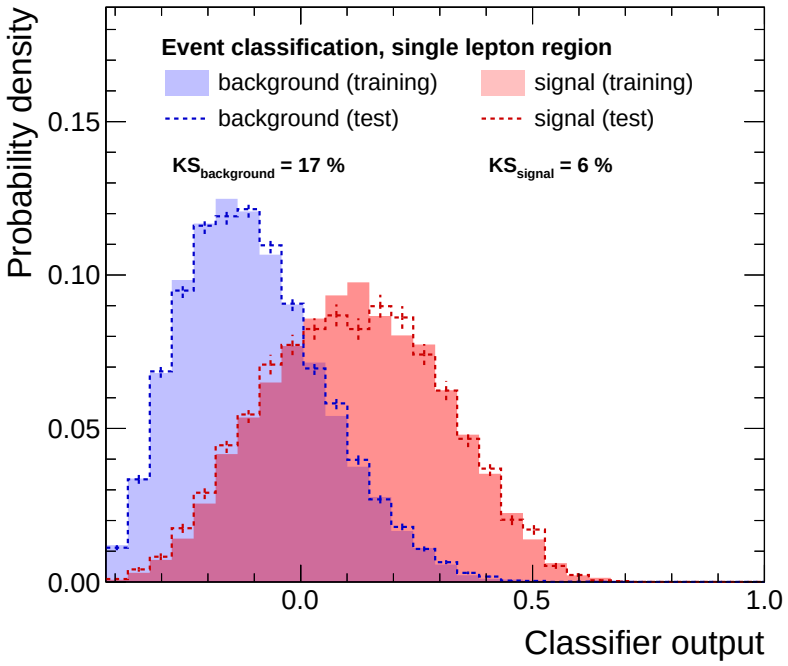


Figure 8.18.: Overtraining check for the BDT of the single-lepton event classification. The BDT shows similar performance for the training data set and a statistically independent testing data set. No signs of overtraining are observed.

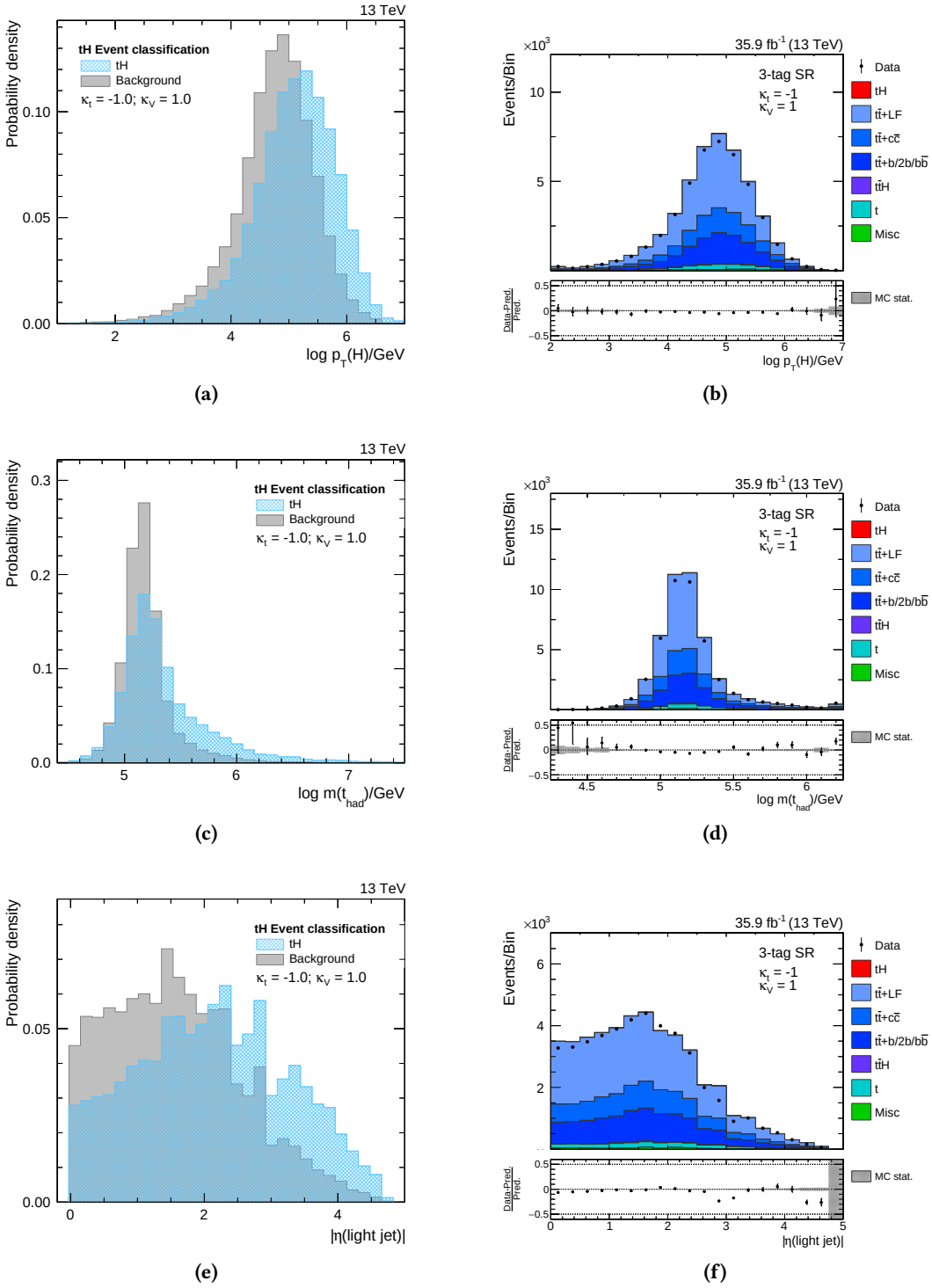


Figure 8.19.: The three most important variables in the training of the single-lepton event classification. For the distributions on the right, the simulation is scaled to the expected number of events. All input variables are modeled well by the simulation.

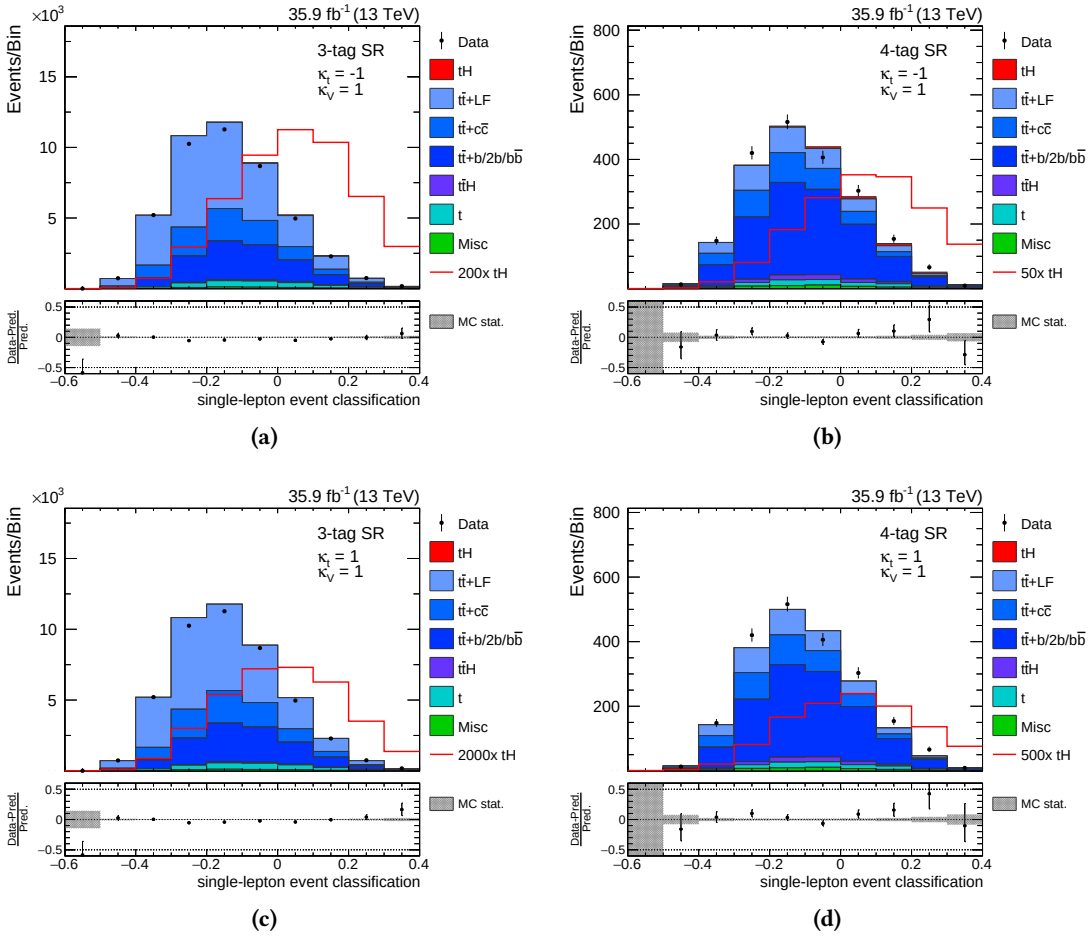


Figure 8.20.: Result of the single-lepton event classification in the 3 tag (a) and 4 tag (b) signal regions for the ITC scenario. The SM scenario for both regions ((c) and (d)) is recovered by applying the dedicated LHE weights. The simulation is scaled to the expected number of events. Sufficient agreement between simulation and data is observed.

8.6.2. Dileptonic Region

An event classification is also performed in the dileptonic control region with a dedicated BDT. This event classification does not aim to distinguish signal processes involving Higgs bosons from the top quark pair background, but to separate different top quark pair contributions. More specifically, a BDT is trained to distinguish heavy-flavor components from light-flavor components of additional jets in top quark pair production. In particular, heavy-flavor components include all top quark pair production contributions with additional bottom quarks, whereas in the light-flavor component these additional jets stem from light-flavored quarks (up, down and strange quarks) or gluons. Components with charm quark contribution are not explicitly used in the training of the BDT, but they can nevertheless be identified as their response is neither signal- nor background-like. The inclusion of this region has been found to be a significant improvement for this analysis [325].
To achieve a separation of light-flavor events and heavy-flavor events, b tagging information is exploited for the training of the BDT. The light-flavor component is used as a signal process and the heavy-flavor component as a background process in the training. The definition of signal and background process is arbitrary and the opposite definition could also be used. Additional separation can be gained by including information of a dedicated c tagging algorithm, which is optimized in identifying jets originating from charm quarks [219]. In total, the BDT is trained with eight different variables, which are listed in Table 8.10. The settings used in the training of the BDT are provided in Appendix C.4. The results of the training is shown in Fig. 8.21. A good separation of the light-flavor and heavy-flavor component can be achieved and no hints of overtraining are observed. Shape distributions and a comparison between simulation and data for the three most important input variables in the training are provided in Fig. 8.22. The remaining five variables can be found in Appendix C.9. A comparison of the final BDT distribution between simulation and data is provided in Fig. 8.23.

Table 8.10.: List of all input variables used in the training of the dileptonic event classification. The variables are ranked by their importance for the training of the BDT.

Variable	Description
CSV(b-tagged jet 3)	Output of the b tagging discriminant for the b-tagged jet with the third highest b tagging value in the event
n_{jets}(tight)	Number of jets in the event passing the tight working point of the b tagging algorithm
CvsL(jet p_T 3)	Output of the charm vs. light flavor tagging algorithm for the jet with the third highest transverse momentum in the event
CSV(b-tagged jet 2)	Output of the b tagging discriminant for the b-tagged jet with the second highest b tagging value in the event
CvsL(jet p_T 4)	Output of the charm vs. light flavor tagging algorithm for the jet with the fourth highest transverse momentum in the event
CvsB(jet p_T 3)	Output of the charm vs. bottom flavor tagging algorithm for the jet with the third highest transverse momentum in the event
CSV(b-tagged jet 4)	Output of the b tagging discriminant for the b-tagged jet with the fourth highest b tagging value in the event
n_{jets}(loose)	Number of jets in the event passing the loose working point of the b tagging algorithm

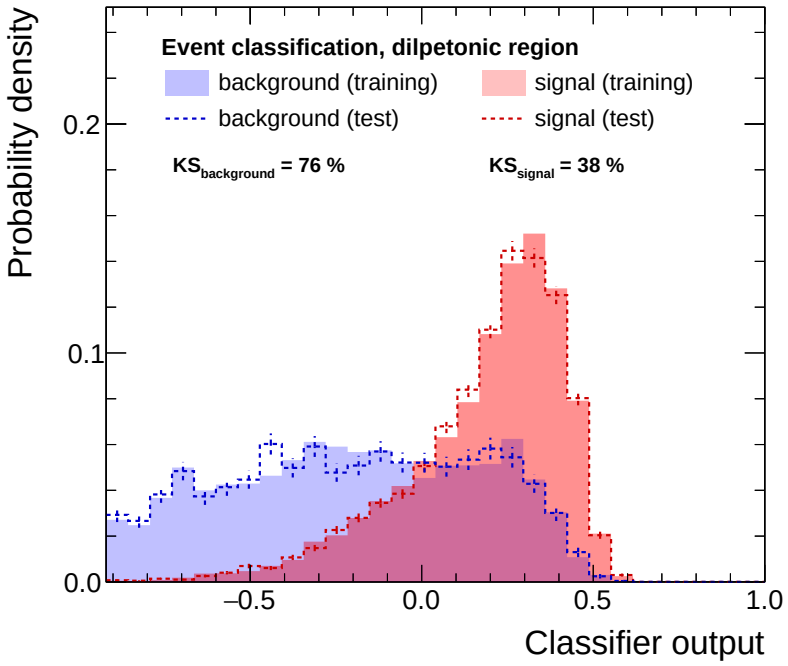
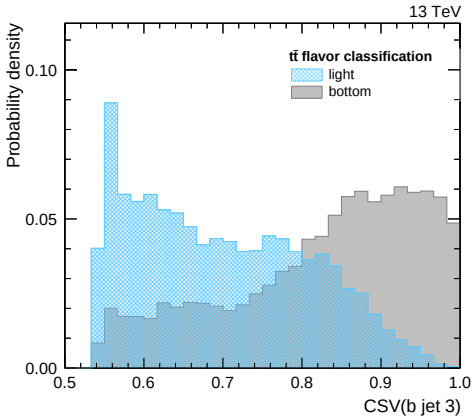
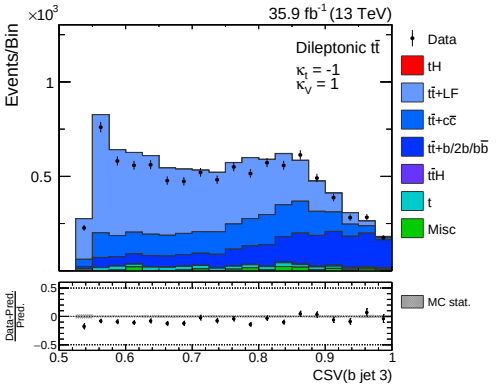


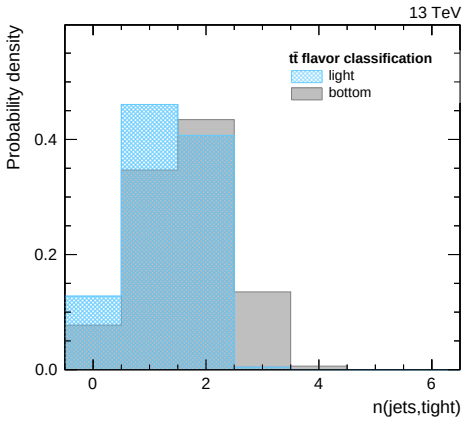
Figure 8.21.: The performance of the dileptonic event classification BDT is tested against overtraining by comparing the result on the training data set with a data set that has not been used in training process. Similar separation power is observed, no overtraining is present.



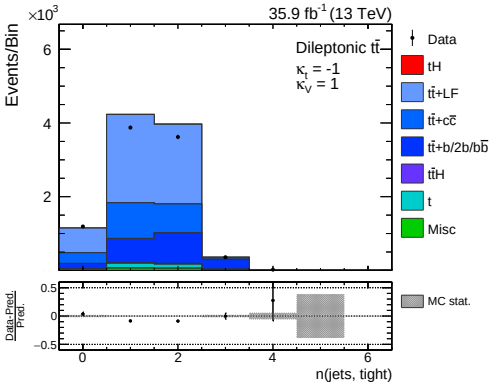
(a)



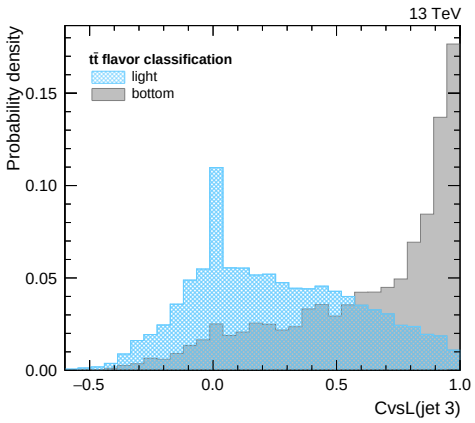
(b)



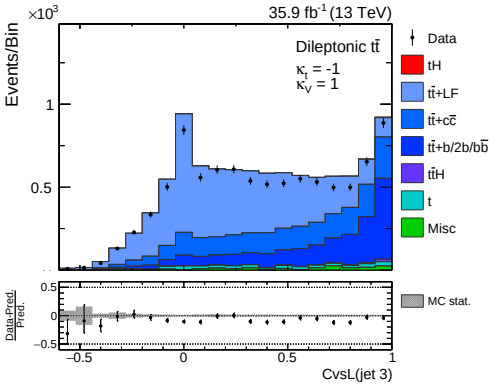
(c)



(d)



(e)



(f)

Figure 8.22.: The three most important variables in the training of the dileptonic event classification. For the distributions on the right, the simulation is scaled to the expected number of events. All input variables are modeled well by the simulation.

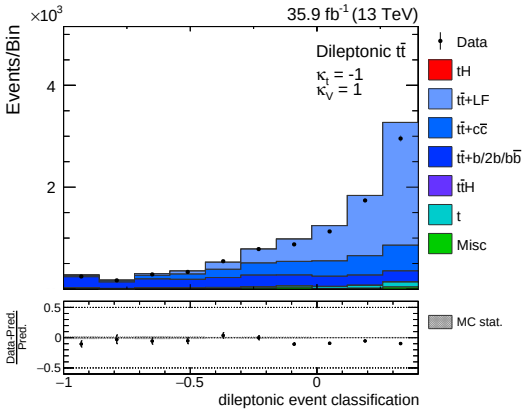


Figure 8.23.: The resulting distribution of the dileptonic event classification BDT in the dileptonic control region. The simulation is scaled to the expected number of events and the last bin of the distribution also contains all events that have larger values. The data is well described by the simulation.

8.7. Systematic Uncertainties

Different sources of systematic uncertainties affect the search for tH production. All considered systematic uncertainty sources are explained in the following, split by experimental and theoretical origin.

8.7.1. Experimental Uncertainties

- **Jet energy scale (JES):** The jet energy correction applied to jets are varied within their given uncertainties [215, 216]. Since this analysis relies on a good description of jets in simulation and data due to the high jet multiplicity, this jet energy scale uncertainty is treated with special care. Instead of one single nuisance parameter describing the variation of the jet energy scale, 26 different nuisance parameters are used, each reflecting a different source to the overall jet energy scale uncertainty [326]. The effect of each individual source affects not only the final BDT distributions, but also causes migration effects between different event categories. This is taken into account by repeating the analysis on dedicated samples, for which the four-momentum of jets has been shifted. The effect on the missing transverse momentum due to the shifted four-momentum of jets is also propagated.
- **Jet energy resolution (JER):** The corrections applied to jets in simulated events to improve the description of data is shifted by the given uncertainties in Ref. [232]. The effect is also propagated through the entire analysis with dedicated simulation samples for up- and down-shifts of this systematic uncertainty.
- **Unclustered energy:** Each particle contributing to the calculation of the missing transverse momentum has a distinct resolution. This resolution is varied for each particle contributing to the missing transverse momentum [221]. Similar to the previous two systematic uncertainties, this effect is propagated to the whole analysis to take migration effects into account.
- **b tagging:** The method applied to correct the prediction of b-tagged jets described in Section 8.4.3, has different systematic uncertainties associated with it [320, 321]. Each systematic variation provides a different set of correction factors for b-tagged jets. A nuisance parameter is assigned to a variation taking into account impurities of the control samples, from which the scale factors were derived. Two nuisance parameters describe the statistical fluctuations of the scale factor determination. Both the systematic uncertainty on the impurity and the statistical uncertainties are evaluated for heavy- and light-flavor scale factors independently. Two additional nuisance parameters control the contribution of charm-induced jets in the scale factor calculation.
- **Lepton efficiencies:** All corrections applied to the description of electrons and muons, such as efficiency of the trigger, isolation and identification (see Section 8.4.2), are switched with systematically-shifted corrections. These shifted correction are derived from systematic effects of the scale factor determination, but also include the statistical uncertainty of the scale factors.

- **Pileup:** An uncertainty to the correction of pileup interaction, described in Section 8.4.1, is made by shifting the nominal value of the minimum-bias cross section of 69.2 mb by ±4.6% [245, 246].
- **Luminosity:** The integrated luminosity of 35.9 fb^{-1} can only be measured with a certain precision. An uncertainty of 2.5% is assigned to the normalization of any simulated process to take this effect into account [298].
- **Limited size of MC simulation samples:** As the templates used in the maximum-likelihood fit are created from MC simulation samples with finite sample size, the uncertainty due to this limitation needs to be considered. This is achieved by a the so-called Barlow–Beeston method [258, 300]. In contrast to the original proposal, the method employed adds only one general nuisance parameter per region and bin for all processes, based on the number of simulated events. This significantly reduces the computing resources and increases the stability of the fit. The method is automatically provided by the COMBINE package.

8.7.2. Theoretical Uncertainties

- **Heavy-flavor component of top quark pair production:** The individual contribution of each heavy-flavor component of the top quark pair production background, i. e., $\text{t}\bar{\text{t}}$ +$\text{b}\bar{\text{b}}$, $\text{t}\bar{\text{t}}$ +2b, $\text{t}\bar{\text{t}}$ +b and $\text{t}\bar{\text{t}}$ +$\text{c}\bar{\text{c}}$, is varied by 50% of the nominal cross section.
- **Factorization and renormalization scales:** The value for the factorization and renormalization scales chosen in the production of simulated samples influences the jet multiplicity and is therefore expected to have an impact for this analysis. The effect is estimated by varying both scales independently for each process by applying a reweighting procedure on LHE basis [260]. For both scales a value of half or double the nominal value is chosen. The final uncertainty for each process is derived from the envelope of all possible combinations, while omitting both cases where the two scales are varied in different directions. This method is not available for the simulation sample of the associated tW single top production. Instead, a rate uncertainty of 3% is assigned, based on the accuracy of the theory prediction. The effect of this systematic uncertainty is evaluated for each process independently.
- **PDF and α_s:** Uncertainties related to the choice of the PDF and the value of α_s have found to be of minor impact to the analysis. Therefore, a conservative rate uncertainty is assigned to each process, based on the respective initial state. The individual values are listed in Table 8.11.
- **Modeling of the top quark p_T:** It has been observed that the distribution of the transverse momentum of top quarks in top quark pair production is significantly shifted to higher values in simulation, compared to data [261]. A correction for this behavior exists, however, it is not recommended by the Top Quark Physics Analysis Group of the CMS Collaboration [303] to apply this correction for the analysis. Nevertheless, it is recommended to estimate the effect of applying this correction as an additional uncertainty.

Table 8.11.: List of variations used for the PDF uncertainty of each process. Variations of the same initial state share the same nuisance parameter and are therefore correlated. The magnitude of each

Process	gg PDF (%)	qg PDF (%)	qq PDF (%)
tHq	—	3.7	—
tHw	—	6.1	—
$t\bar{t}H$	3.6	—	—
$t\bar{t}$	4.2	—	—
Single top	—	4.0	—
$t\bar{t}Z$	—	—	2.0
$t\bar{t}W$	—	—	2.0
tZq	—	4.0	—
Z+jets	—	—	2.0

The correction is applied by reweighting events based on a two-parameter exponential function, dependent on the top quark transverse momentum. Only top quarks from SM top quark pair production are corrected for this uncertainty.

8.8. Results and Interpretation

A combined maximum-likelihood fit (see Section 2.2.3) is performed to the output distribution of the single-lepton event classification in the 3 tag and 4 tag signal regions, as well as to the output distribution of the dileptonic event classification in the dileptonic control region. All three distributions are fitted simultaneously. The corresponding LHE weights are applied to the tHq and tHW processes to recover each given coupling scenario. No significant excess beyond the background-only hypothesis is observed. Therefore, exclusion limits for a confidence level of 95% are derived. In the following, the different limit scenarios are explained.

8.8.1. tH Production

In the search for tH production, the exclusion limits are given on the signal strength $\mu = \sigma_{\text{obs}}/\sigma_{\text{pred}}$, where σ_{obs} is the cross section that can be excluded and σ_{pred} the predicted cross section for the given scenario. Two scenarios are examined: the ITC scenario and the SM scenario. The result of the combined fit in all three regions for the ITC scenario is shown in Fig. 8.24 and for the SM scenario in Fig. 8.25. The combined result for each scenario is shown in Fig. 8.26, where all bins are arranged according to their signal significance. The exclusion limits for both scenarios are listed in Table 8.12. Assuming SM couplings, the expected exclusion limit for tH is found to be $\mu^{\text{tH}}_{\text{exp,SM}} = 41.5$ and the observed limit is found to be $\mu^{\text{tH}}_{\text{obs,SM}} = 89.6$, which is still compatible within two standard deviations of the expectation. For the ITC scenario, the expected exclusion limit on tH production yields $\mu^{\text{tH}}_{\text{exp,ITC}} = 2.95$ and the observation is $\mu^{\text{tH}}_{\text{obs,ITC}} = 5.83$. All observed exclusion limits are compatible with the expectation.

Table 8.12.: Exclusion limits on the signal strength of tH production for the SM and ITC scenarios. In addition to the combined exclusion limit, a separate estimation has been made for each region individually.

Scenario	Region	Observed limit	Expected limit	±1 standard deviation	±2 standard deviations
SM ($\kappa_t = +1$)					
	3 tag	92.6	65.4	[42.4, 95.4]	[30.5, 145]
	4 tag	92.4	47.0	[32.0, 72.2]	[23.3, 110]
	dilep.	767	629	[456, 882]	[346, 1198]
	comb.	89.6	41.5	[28.2, 62.7]	[20.4, 92.8]
ITC ($\kappa_t = -1$)					
	3 tag	6.31	4.08	[2.78, 6.20]	[2.01, 9.36]
	4 tag	6.30	3.68	[2.52, 5.60]	[1.84, 8.43]
	dilep.	82.9	75.6	[55.5, 104]	[42.5, 138]
	comb.	5.83	2.95	[2.00, 4.46]	[1.45, 6.64]

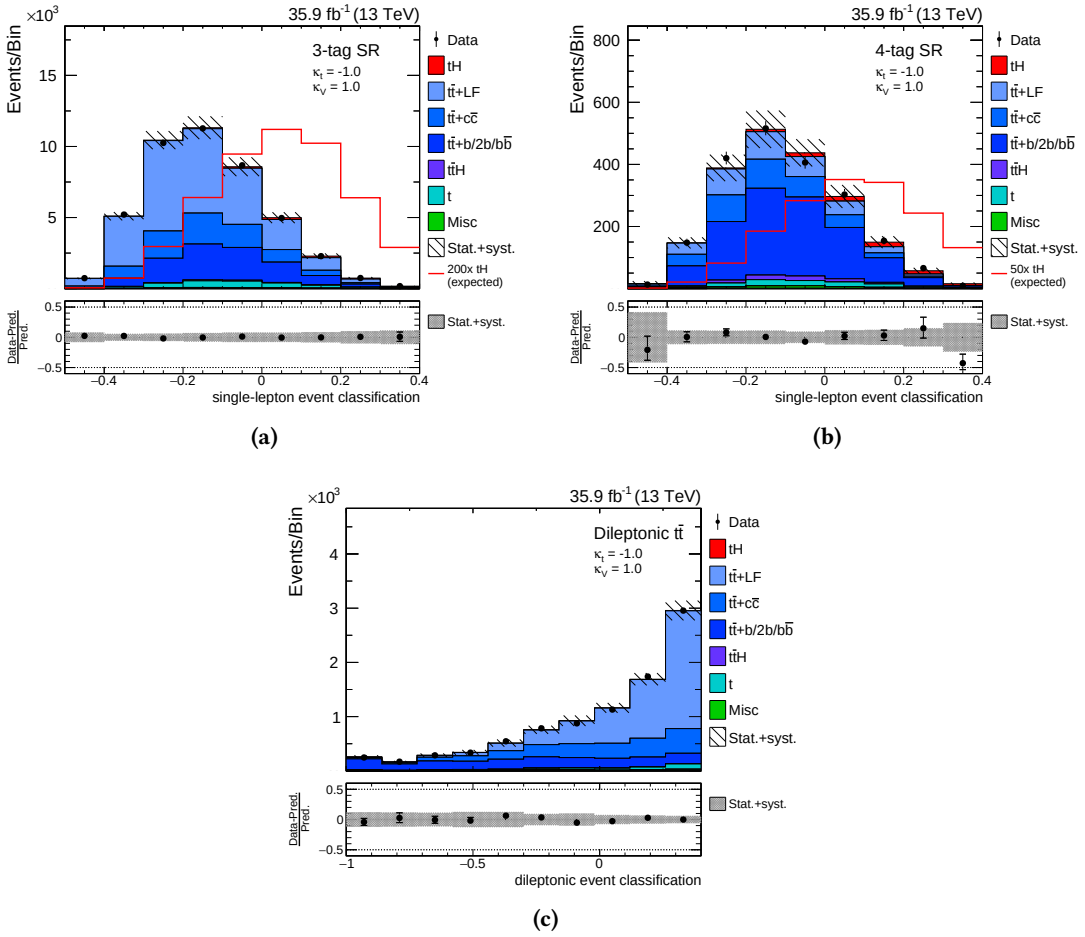


Figure 8.24.: The resulting distributions after the simultaneous maximum-likelihood fit of the single-lepton event-classification output in the 3 tag (a) and 4 tag (b) signal regions, as well as the dileptonic event classification output in the dileptonic control region (c), assuming the ITC coupling scenario and tH production as signal. The simulation is scaled to the result of the fit.

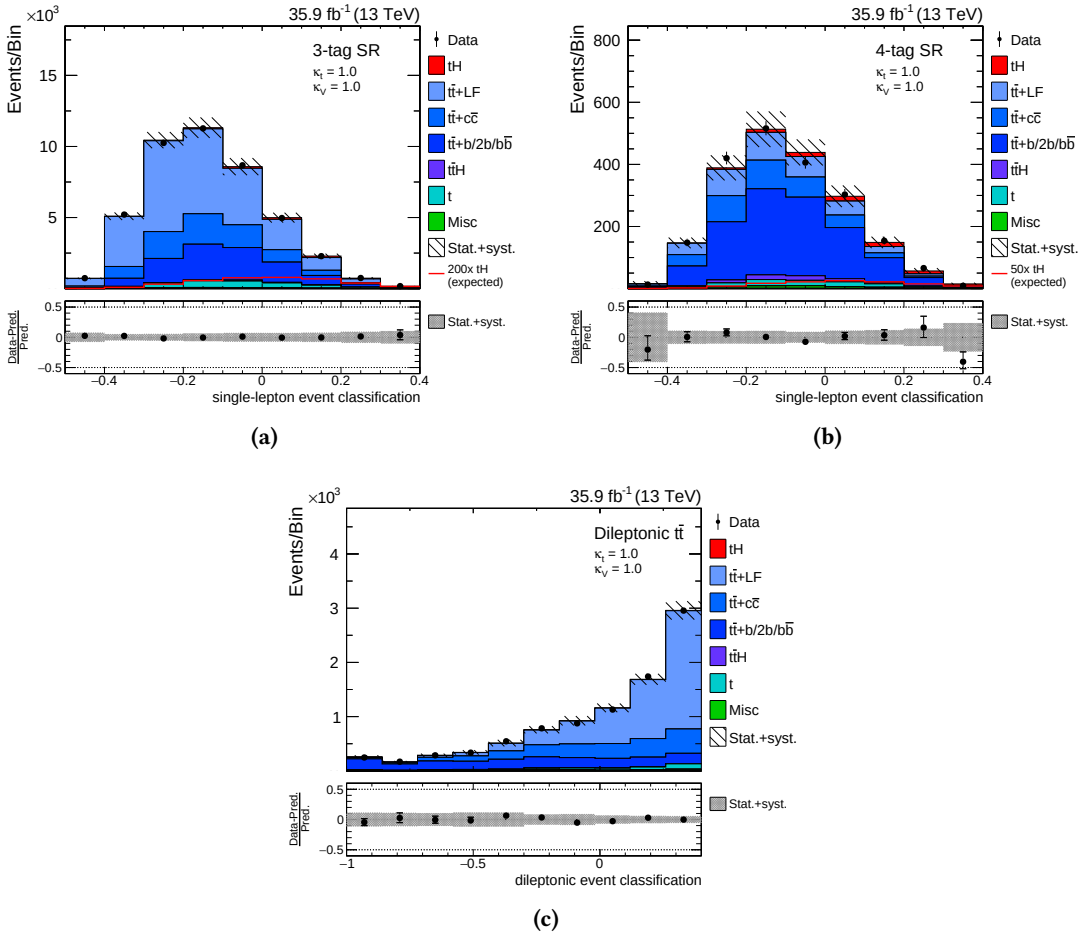
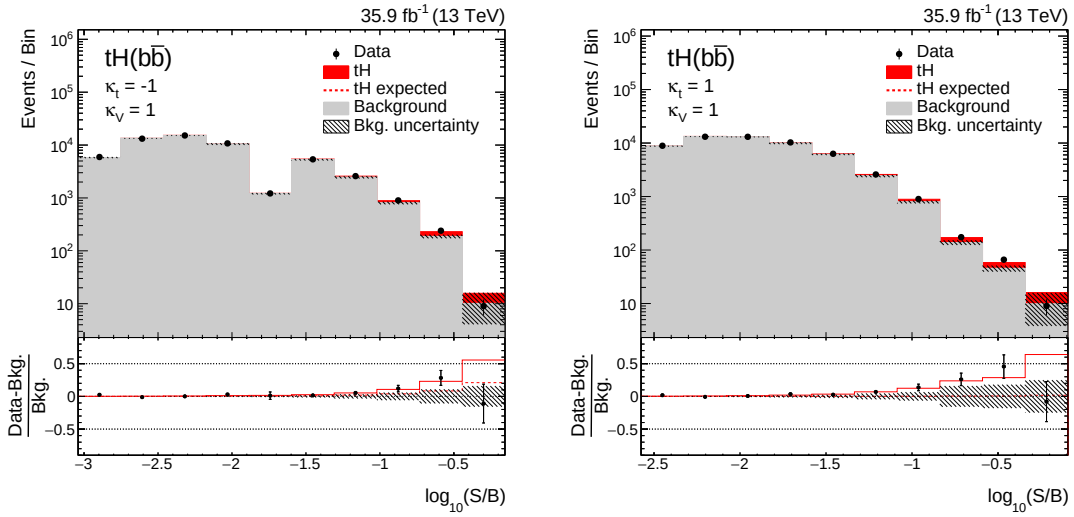


Figure 8.25.: The resulting distributions after the simultaneous maximum-likelihood fit of the single-lepton event-classification output in the 3 tag (a) and 4 tag (b) signal regions, as well as the dileptonic event classification output in the dileptonic control region (c), assuming the SM coupling scenario and tH production as signal. The simulation is scaled to the result of the fit.



(a) **(b)**

Figure 8.26.: Combined comparison of prediction and data for the ITC scenario (a) and the SM scenario (b). All bins of the fitted distributions are rearranged according to their signal-to-background ratio such that an excess above the background-only hypothesis would be visible to the right. No significant excess is observed.

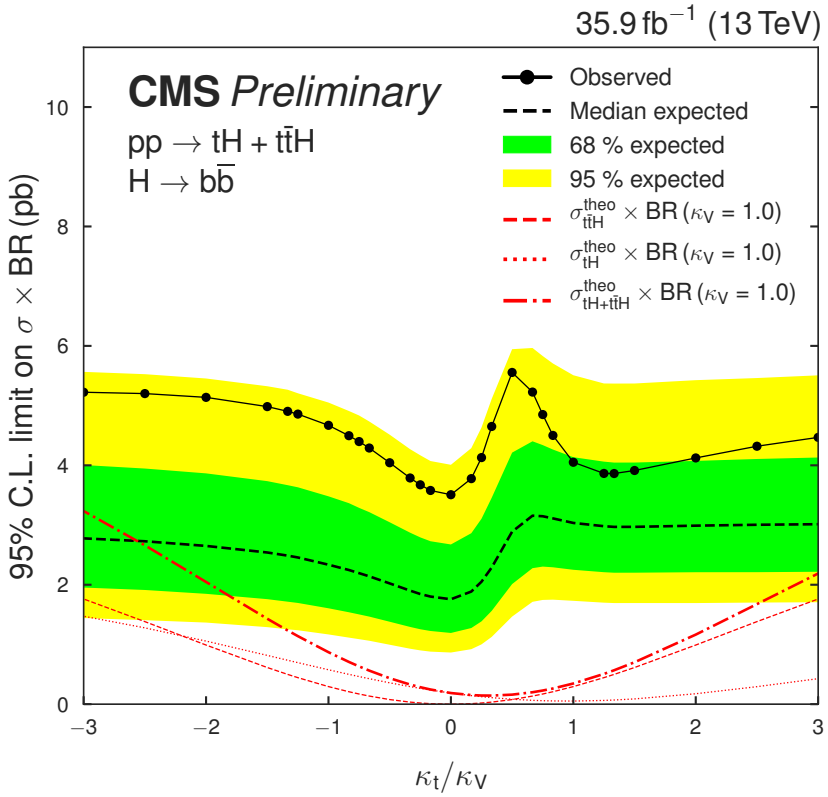


Figure 8.27.: Observed and expected 95% C.L. exclusion limits obtained for the scan over different κ_t/κ_V ratios on the combined tH+t$\bar{t}$H cross section times branching ratio. In addition, the exclusion curves for tH, t$\bar{t}$H and tH+t$\bar{t}$H production are shown in red (assuming $\kappa_V = 1$).

8.8.2. κ_t/κ_V Coupling Ratios

Variations of the coupling modifiers other than $\kappa_t = -1$ affect also the production cross section of the t$\bar{t}$H processes, as well as the branching ratio of the Higgs boson decay into a bottom quark-antiquark pair in general. For the broader scan over modifications of κ_t and κ_V, or more precisely κ_t/κ_V, the exclusion limits are given on the combined tH+t$\bar{t}$H production cross section times branching ratio. To incorporate the correct behavior for each studied κ_t/κ_V point, a special physics model in the COMBINE framework is used. This model automatically adapts both the cross section and the branching ratio of the given scenario. The complete list of exclusion limits on the cross section of tH+t$\bar{t}$H times branching ratio for all studied points are given in Table 8.13. In Fig 8.27, all exclusion limits are shown as a function of κ_t/κ_V. In addition, the predicted cross section times branching ratio for tH, t$\bar{t}$H and tH+t$\bar{t}$H production are shown for $\kappa_V = 1$.

8.8.3. CP-Mixing

In addition to exclusion limits on κ_t/κ_V, exclusion limits are set on a possible CP-odd admixture of the Higgs boson. Since the modified top-Higgs coupling also affects the t$\bar{t}$H production, the

Table 8.13.: Complete list of all exclusion limits for different κ_t/κ_V coupling ratios. The exclusion limits are given on the combined cross section of tH+t$\bar{t}$H production times the branching ratio in units of pb.

κ_t/κ_V	Observed limit	Expected limit	±1 standard deviation	±2 standard deviations
−6.00	5.21	2.88	[2.05, 4.10]	[1.53, 5.61]
−4.00	5.22	2.83	[2.00, 4.06]	[1.48, 5.60]
−3.00	5.22	2.78	[1.95, 4.00]	[1.44, 5.57]
−2.50	5.20	2.73	[1.91, 3.95]	[1.41, 5.53]
−2.00	5.14	2.65	[1.85, 3.87]	[1.37, 5.45]
−1.50	4.98	2.54	[1.76, 3.73]	[1.29, 5.32]
−1.33	4.90	2.49	[1.72, 3.67]	[1.25, 5.27]
−1.25	4.86	2.46	[1.70, 3.63]	[1.24, 5.21]
−1.00	4.67	2.34	[1.61, 3.48]	[1.17, 5.05]
−0.83	4.50	2.25	[1.54, 3.37]	[1.11, 4.91]
−0.75	4.40	2.19	[1.50, 3.29]	[1.09, 4.83]
−0.67	4.29	2.14	[1.47, 3.22]	[1.06, 4.73]
−0.50	4.04	2.02	[1.38, 3.05]	[1.00, 4.51]
−0.33	3.79	1.90	[1.29, 2.87]	[0.93, 4.27]
−0.25	3.67	1.84	[1.25, 2.79]	[0.90, 4.17]
−0.17	3.58	1.80	[1.23, 2.73]	[0.88, 4.07]
0.00	3.51	1.76	[1.19, 2.67]	[0.86, 4.01]
0.17	3.78	1.89	[1.28, 2.86]	[0.92, 4.29]
0.25	4.13	2.05	[1.39, 3.11]	[1.00, 4.63]
0.33	4.65	2.29	[1.56, 3.46]	[1.13, 5.11]
0.50	5.56	2.89	[2.01, 4.21]	[1.47, 5.94]
0.67	5.23	3.16	[2.28, 4.40]	[1.71, 5.97]
0.75	4.85	3.15	[2.30, 4.34]	[1.75, 5.83]
0.83	4.50	3.11	[2.29, 4.26]	[1.75, 5.71]
1.00	4.05	3.03	[2.25, 4.14]	[1.73, 5.51]
1.25	3.86	2.98	[2.21, 4.06]	[1.70, 5.37]
1.33	3.86	2.97	[2.20, 4.05]	[1.69, 5.37]
1.50	3.91	2.97	[2.20, 4.05]	[1.69, 5.37]
2.00	4.12	2.99	[2.21, 4.07]	[1.69, 5.43]
2.50	4.32	3.00	[2.21, 4.11]	[1.70, 5.46]
3.00	4.47	3.01	[2.22, 4.13]	[1.71, 5.51]
4.00	4.65	3.02	[2.21, 4.16]	[1.70, 5.55]
6.00	4.82	3.00	[2.19, 4.16]	[1.68, 5.58]

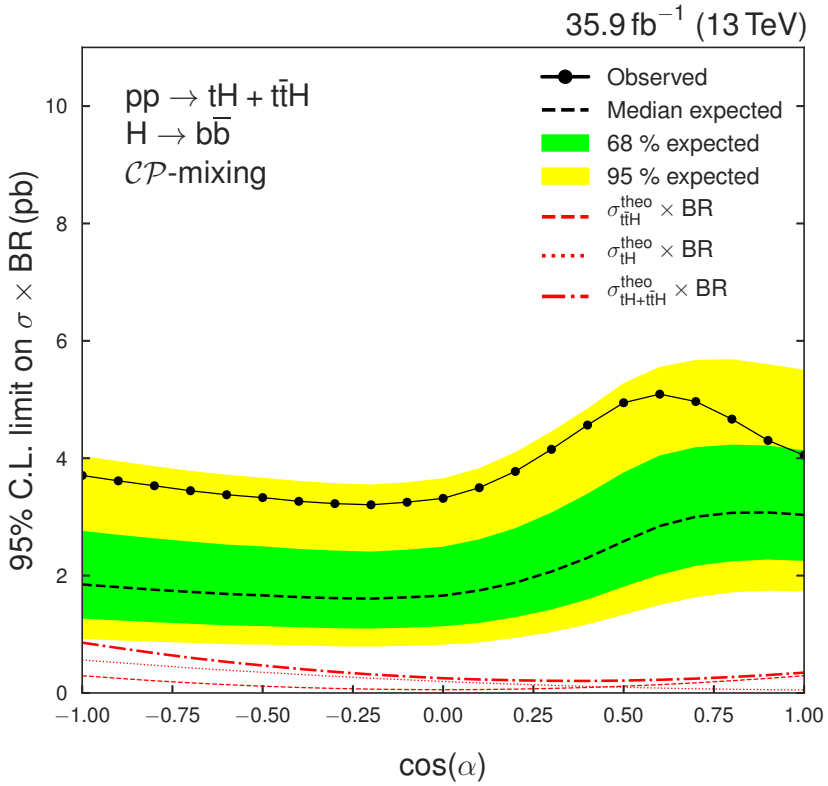


Figure 8.28.: Observed and expected 95% C.L. exclusion limits for different CP-mixing scenarios as a function of the mixing angle α. In addition, the exclusion curves for tH, t$\bar{t}$H and tH+t$\bar{t}$H production are shown in red.

exclusion limits are given on the combined tH+t$\bar{t}$H cross section times branching ratio. The limits are given as a function of the mixing angle α. A list of all exclusion limits is provided in Table 8.14 and Fig. 8.28.

Table 8.14.: List of exclusion limits for different CP-mixing scenarios, parameterized by the mixing angle α. The exclusion limits are given on the combined cross section of tH+t$\bar{t}$H production times the branching ratio in units of pb.

$\cos\alpha$	Observed limit	Expected limit	±1 standard deviation	±2 standard deviations
−1.0	3.71	1.85	[1.27, 2.76]	[0.92, 4.03]
−0.9	3.61	1.80	[1.24, 2.70]	[0.89, 3.95]
−0.8	3.53	1.76	[1.20, 2.64]	[0.87, 3.86]
−0.7	3.45	1.72	[1.18, 2.58]	[0.85, 3.78]
−0.6	3.38	1.69	[1.15, 2.53]	[0.83, 3.72]
−0.5	3.33	1.66	[1.14, 2.50]	[0.82, 3.67]
−0.4	3.27	1.63	[1.12, 2.45]	[0.81, 3.61]
−0.3	3.23	1.62	[1.10, 2.43]	[0.80, 3.57]
−0.2	3.21	1.61	[1.10, 2.41]	[0.79, 3.55]
−0.1	3.25	1.63	[1.11, 2.44]	[0.80, 3.59]
0.0	3.32	1.66	[1.14, 2.49]	[0.82, 3.66]
0.1	3.50	1.75	[1.19, 2.62]	[0.86, 3.83]
0.2	3.77	1.88	[1.29, 2.81]	[0.94, 4.10]
0.3	4.15	2.07	[1.42, 3.07]	[1.03, 4.46]
0.4	4.57	2.30	[1.59, 3.40]	[1.17, 4.85]
0.5	4.95	2.58	[1.81, 3.76]	[1.33, 5.28]
0.6	5.09	2.84	[2.01, 4.05]	[1.50, 5.56]
0.7	4.97	3.00	[2.17, 4.19]	[1.63, 5.68]
0.8	4.67	3.07	[2.24, 4.23]	[1.71, 5.69]
0.9	4.30	3.08	[2.27, 4.22]	[1.74, 5.60]
1.0	4.05	3.03	[2.25, 4.14]	[1.73, 5.51]

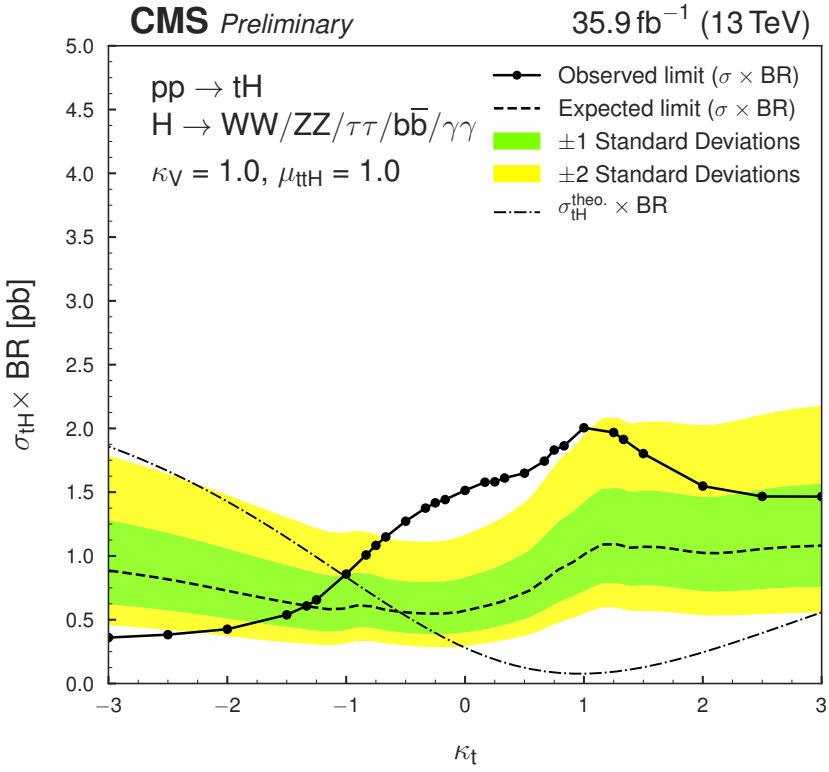


Figure 8.29.: Observed and expected 95% C.L. exclusion limits obtained for the scan over different κ_t/κ_V ratios on the tH production cross section after combining different Higgs boson decay channels. An exclusion curve for the tH production cross section times branching ratio is shown as dotted-dashed line.

8.9. Combination of the Results with Different Higgs Boson Decay Modes

The former exclusion limits are derived only in the decay mode of the Higgs boson to a bottom quark-antiquark pair. To increase the sensitivity, the results obtained in this chapter are combined with other analyses searching for tH production, but focusing on a final state with multiple leptons [327] or the decay of the Higgs boson into two photons [328].
For the multilepton analysis, the decay of the Higgs boson into a pair of W bosons, Z bosons or tau leptons is investigated, where at least one of the two decay products has to decay into leptons. Together with the leptonic decay of the top quark in tHq and tHW process, the final state consists of at least two charged leptons.
For the Higgs to diphoton analysis, events are required to contain two photons with high transverse momentum. The results in Ref. [328] include all possible Higgs boson production mechanisms. These results are reinterpreted to set upper limits on the tH+t$\bar{t}$H production cross section using the tag categories for t$\bar{t}$H production.
The result of the combination for tH production is shown in Fig. 8.29 under the assumption of no SM Higgs boson (similar to Fig. 8.27). In addition to the exclusion limits for tH production

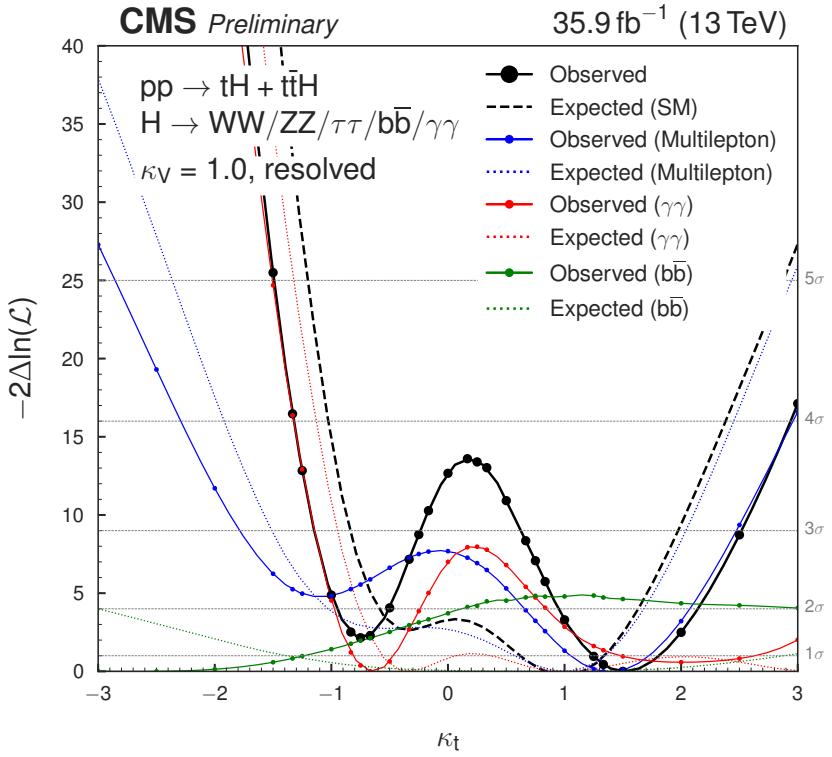


Figure 8.30.: Profile likelihood scan to determine κ_t for a fixed value of $\kappa_V = +1$. The observation is shown as solid line for the combined result (black), only the multilepton channel (blue), only the diphoton channel (red) and only the bottom quark-antiquark channel (green). The expected values for each channel are shown as dotted lines and the expectation from a SM Higgs boson as dashed line. Values outside of $[-0.9, -0.5]$ and $[0.9, 2.2]$ can be excluded for κ_t with a confidence level of 95% (2σ).

with different values of κ_t and κ_V, a profile likelihood fit is performed to derive constraints on κ_t, while fixing κ_V to +1. The result of this scan is shown in Fig. 8.30. With the combined result, values outside of $[-0.9, -0.5]$ and $[0.9, 2.2]$ can be excluded for κ_t with a confidence level of 95%. The ITC scenario is therefore excluded with the combined result of different Higgs boson decay channels.

8.10. Summary and Outlook

A search for tH production with the CMS experiment at a center-of-mass energy of 13 TeV has been presented in this chapter. The analysis exploits the decay channel of the Higgs boson into a bottom quark-antiquark pair and utilizes the complete data set of 2016, corresponding to an integrated luminosity of 35.9 fb^{-1}. The analysis has been published by the CMS Collaboration as *Physics Analysis Summary* [329]. The combination of this measurement with measurements for different Higgs boson final states is going to be published in *Physical Review D*. The combined likelihood scan is able to exclude κ_t values outside of $[-0.9, -0.5]$ and $[0.9, 2.2]$, assuming

$\kappa_V = +1$. The ITC scenario is therefore also excluded. No dedicated searches for tH production have been performed by the ATLAS Collaboration, only a combined result for tH+t$\bar{t}$H production in the diphoton channel with data from Run I is available [330]. This analysis obtains compatible results, but the sensitivity is not at the same level due to the lower cross section at 7 and 8 TeV.

For future analyses of Higgs boson production in association with single top quark and the Higgs to bottom quark-antiquark decay channel, it is not sufficient to just add more data to increase the sensitivity, as the current analysis is limited by systematic uncertainties. Instead, the overall systematic uncertainties have to be reduced. One promising topic for this reduction is the implementation of b jet energy regression [331]. Since jets from bottom quarks involve semileptonic decays of hadrons, the jet is associated with a larger amount of missing transverse momentum that cannot be clustered into the jet itself and therefore spoils the measured transverse momentum of the jet. The b jet energy regression provides a multidimensional correction from a DNN, specifically derived for b jets that correct and improve the modeling of these jets. This allows to reduce the uncertainties from jet energy corrections associated with b jets.

Another way to improve the existing analysis is to implement different methods for the background treatment. For instance, the prospect of a data-driven approach for the top quark pair production background can be reevaluated, which has been already used for the analysis at 8 TeV [314]. In this approach, the modeling of the top quark pair production background is directly taken from data in the 2 tag control region, which is pure in top quark pair events. The contribution to the to 3 tag and 4 tag signal region is then extrapolated by calculating a probability that an event would have an additional b-tagged jet. By employing this method, uncertainties in the modeling can be drastically reduced. However, this method has the drawback to then rely even more on systematic uncertainties of the b tagging algorithm.

The recent progress in more sophisticated machine learning could also be exploited to improve the jet-to-quark assignment in the existing analysis. In general, the BDTs for the jet-to-quark assignment could be replaced by dedicated DNNs and instead of specifically chosen variables also the complete four-momenta of jets could be used as training variables. With enough nodes and layers, the DNN should outperform the BDT in choosing the correct jet-to-quark assignment [332]. Instead of reconstructing each event separately under different hypotheses and performing an event classification afterwards, the jet-to-quark assignment could be combined in a single DNN with the event classification. Such a network would consist of a few nodes in the first layer with variable input length per event [333, 334], which represent the jets in each event and replace the dedicated jet-to-quark assignment BDTs. These variable-size nodes have already been successfully tested in high energy physics [335]. The limiting factor for most of these machine learning approaches is the sample size of the training data set, which is usually not sufficient at the analysis level.

9. Conclusion

With the observation of the Higgs boson at the LHC in 2012, the existence of the last fundamental particle predicted by the standard model has been finally experimentally verified by the ATLAS and CMS Collaborations. The discovery of the Higgs boson nearly 50 years after its first prediction and the absence of other new unpredicted particles so far is a huge success for the standard model of particle physics. Although the standard model works excellent at the energy range of current particle colliders, theoretical aspects and evidence from other experiments suggest that there is physics beyond the standard model realized in nature. However, no direct evidence for new physics was found at the LHC in Run I, which lasted until the end of 2012.
In 2015, a new era of the LHC began with an increase of the center-of-mass energy from 8 to 13 TeV. The higher energy allowed to quickly push the direct exclusion limits of many popular theories of new physics up to several TeV. Another approach apart from increasing the center-of-mass energy even further to search for new physics is to look for subtle deviations in high-precision standard model predictions. Many theories of new physics are motivated to resolve the open questions of the standard model and are therefore expected to interact with standard model particles.

One way to search for such deviations is to study single top quark production at the LHC. Unlike the more likely gluon-induced production of top quark pairs, single top quarks are produced through electroweak interaction. The electroweak unification of the electromagnetic and weak force in the standard model is also related to the Higgs mechanism, responsible for the masses of fermions and gauge boson, whose associated particle is the Higgs boson. Experimentally, the event topology of the different single top quark production modes is similar to the topology of top quark pair production, which makes top quark pair production the dominant background process. Therefore, measurements with single top quarks cannot be realized by simply counting events, but instead require the usage of multivariate analysis techniques to provide meaningful results.

The first analysis presented in this thesis was the measurement of the t-channel single top quark production cross section. Early evidence for this process was observed with the first proton-proton collisions at a center-of-mass energy of 13 TeV in the summer of 2015. This early data set is only a tiny fraction of the complete data set of 2015, which has been also analyzed and described in this thesis. The cross section was found to be $\sigma_{t-\mathrm{ch.,t}} = 154 \pm 8\,(\mathrm{stat}) \pm 9\,(\mathrm{exp}) \pm 19\,(\mathrm{theo}) \pm 4\,(\mathrm{lumi})\,\mathrm{pb} = 154 \pm 22\,\mathrm{pb}$ for single top quark production and $\sigma_{t-\mathrm{ch.,\bar{t}}} = 85 \pm 10\,(\mathrm{stat}) \pm 4\,(\mathrm{exp}) \pm 11\,(\mathrm{theo}) \pm 2\,(\mathrm{lumi})\,\mathrm{pb} = 85 \pm 16\,\mathrm{pb}$ for single top antiquark production. This resulted in a ratio of $R_{t-\mathrm{ch.}} = 1.81 \pm 0.18\,(\mathrm{stat}) \pm 0.15\,(\mathrm{syst})$, which is directly related to the ratio of up- and down-type quarks in the proton. The inclusive cross section of single top quark and antiquark production was found to be $\sigma_{t-\mathrm{ch.,t+\bar{t}}} =$

238 ± 13 (stat) ± 12 (exp) ± 26 (theo) ± 5 (lumi) pb $= 238 \pm 32$ pb, from which the CKM matrix element $|f_{\mathrm{LV}} \cdot V_{\mathrm{tb}}| = 1.05 \pm 0.07$ (exp) ± 0.02 (theo) was determined. All measured values are in agreement with the predictions from the standard model [263].

With the increased amount of data recorded in 2016 and 2017, compared to 2015, it is possible to search for more rare single top quark production modes at the LHC. Such a search was presented as the second analysis of this thesis, a search for the *s*-channel single top quark production. The *s* channel is the rarest of all single top quark production modes in proton-proton collisions at the LHC and was not observed at center-of-mass energies of 7, 8 and 13 TeV so far. Not only is this production mode more than 20 times less likely than single top quark production via the *t* channel, but also higher center-of-mass energies increase the relative background contribution with respect to the signal process. For this reason, the majority of the recorded proton-proton collision data at Run II of the LHC, i. e., the combined data sets of 2016 and 2017, was used to search for the *s*-channel single top quark production mode. Strong evidence for this process was found with an observed significance of 4.63 standard deviations, while a significance of 4.45 standard deviations was expected assuming a standard-model-like signal. This resulted in a production cross section for the *s*-channel single top quark production of $\sigma_{s\ \mathrm{ch.}} = 11.66\ ^{+3.50}_{-3.38}$ (syst + stat) pb. No significant deviations from the predictions of the standard model were observed. Since the search is still limited by the available data set, an observation of this process, i. e., an observed significance of at least five standard deviations, should be possible by including the data set of 2018.

The third and final analysis presented in this thesis was the search for the associated production of a single top quark with a Higgs boson. This production mode provides a way to study the interplay between the top quark and the Higgs boson due to interference effects and to determine the top-Yukawa coupling, as well as a possible CP-violating coupling of the Higgs boson to the top quark. A dedicated analysis for the Higgs boson decay channel into a bottom quark-antiquark pair was presented, along with the combination of other Higgs boson final states. Assuming a standard model coupling of the Higgs boson to vector bosons ($\kappa_{\mathrm{V}} = +1$), values of the coupling modifier for the top-Yukawa coupling κ_{t} outside of $[-0.9, -0.5]$ and $[0.9, 2.2]$ were excluded. This excludes a possible negative sign of the top-Yukawa coupling in the standard model, the so-called *inverted top coupling* scenario, which is indistinguishable from the coupling with positive sign for most channels that are sensitive to this parameter. A possible CP-violating coupling of the Higgs boson to top quarks was not observed, but could also not be excluded with the available data set [329].

The data sets analyzed in this thesis were recorded during a time period of three years, starting from the first proton-proton collisions at a center-of-mass energy of 13 TeV and an integrated luminosity of $42\,\mathrm{pb}^{-1}$ in 2015, up to the combined data sets of 2016 and 2017 with a total of $77.4\,\mathrm{fb}^{-1}$. The integrated luminosity of these data sets ranges over four orders of magnitude and illustrates the path of the LHC from a domain of statistically limited results to a new precision era. Additional $60\,\mathrm{fb}^{-1}$ to $80\,\mathrm{fb}^{-1}$ are expected to be recorded during 2018 before the LHC undergoes a two-year shutdown for preparing an increase of the center-of-mass energy up

to 14 TeV for Run III. This data-taking period will also last three years, beginning in 2021 and running until the end of 2023, expected to provide around 150 fb^{-1} of proton-proton collision data. The data recorded up to this date will be only a glimpse of the data that the LHC will provide in the future, as the LHC will be upgraded into the High-Luminosity Large Hadron Collider (HL-LHC) [336–338] in 2026, raising the instantaneous luminosity up to $5 \cdot 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$. This will allow the HL-LHC to accumulate more than 3 ab^{-1} of proton-proton collisions in the following ten years. However, this comes with the drawback of increasing the mean number of interactions per bunch crossing up to 200, which requires significant upgrades of the particle detectors to be able to resolve the primary hard scattering from multiple pileup interactions.

Although additional collision data increases the precision of single top quark measurements, the impact of the statistical uncertainty will diminish at some point, which, in case for the t-channel production mode, has already been reached with the available data set. Instead, the sensitivity of each measurement is determined by the systematic uncertainties. Improvements on some experimental uncertainties are expected in the future, as most experimental systematic uncertainties also include a non-negligible statistical component. However, the current theoretical systematic uncertainties are dominant for the measurements presented in this thesis and are independent of the size of the analyzed data. A more in-depth understanding of these uncertainties will become even more important in the future to further increase the precision of measurements. As an alternative to the predictions from event generators, data-driven techniques could be used that do not rely on uncertainties of an underlying theory. On the other hand, utilizing these approaches requires profound understanding of experimental uncertainties.

Moving away from the LHC, single top quark production can also be observed at a possible future lepton collider with sufficient energy [339, 340]. Here, the non-hadronic initial state provides a much cleaner experimental environment for studying single top quark production. The production of single top quarks is not only of interest for collider physics, but also plays a role in neutrino physics for instance, where high-energetic atmospheric neutrinos can create multi-lepton signatures in the *IceCube* detector [341].

Regardless of how single top quark production will be studied in the future, it will provide a powerful tool to probe the properties of the standard model and to search for possible deviations as a sign of new physics.

A. Appendix: Measurement of the t-Channel Single Top Quark Production Cross Section at $\sqrt{s} = 13$ TeV

A.1. Data Sets and MC Simulation Samples

Table A.1.: The data set of 2015 used for the analysis with the corresponding run range and the integrated luminosity.

Period	Run range	Integrated luminosity
/SingleMuon/Run2015D-16Dec2015-v1/MINIAOD	256630–260627	$2.2\,\text{fb}^{-1}$

Table A.2.: List of all MC simulation samples used in the analysis of 2015 data with the number of generated events and the corresponding cross section times branching ratio. The data set name corresponds to the primary data set name, the fragments RunIIFall15MiniAODv2-PU25nsData2015v1_76X_mcRun2_asymptotic_v12_ext1-v* and MINIAODSIM have been omitted for better readability. If not stated otherwise, the cross section is taken from the generator.

Data set name	Events	Cross section × BR (pb)
ST_t-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	29 257 464	216.99×0.3258 (NLO [3, 37])
TT_TuneCUETP8M1_13TeV-powheg-pythia8	97 818 391	831.76 (NNLO [34])
ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	1 000 000	35.85 (NNLO [37])
ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	999 400	35.85 (NNLO [37])
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	28 751 199	5765 (NNLO [342])
WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	237 398 193	61527 (NLO)

Table A.3.: List of all additional MC simulation samples used for systematic variations in the analysis of 2015 data with the number of generated events and the corresponding cross section times branching ratio. The data set name corresponds to the primary data set name, the fragments RunIIFall15MiniAODv2-PU25nsData2015v1_76X_mcRun2_asymptotic_v12_ext1-v* and MINIAODSIM have been omitted for better readability. If not stated otherwise, the cross section is taken from the generator.

Data set name	Events	Cross section × BR (pb)
ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	64 957 724	136.02 (NLO [37])
ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	38 932 192	80.95 (NLO [37])
ST_t-channel_4f_leptonDecays_13TeV-amcatnlo-herwigpp_TuneEE5C	19 390 065	216.99×0.3258 (NLO [3,37])
ST_t-channel_4f_scaleup_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	28 635 525	216.99×0.3258 (NLO [3,37])
ST_t-channel_4f_scaledown_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	29 775 108	216.99×0.3258 (NLO [3,37])
TTJets_13TeV-amcatnloFXFX-pythia8	115 050 051	831.76 (NNLO [34])
TT_TuneEE5C_13TeV-powheg-herwigpp	18 639 024	831.76 (NNLO [34])
TT_TuneCUETP8M1_13TeV-powheg-scaleup-pythia8	38 507 969	831.76 (NNLO [34])
TT_TuneCUETP8M1_13TeV-powheg-scaledown-pythia8	39 461 147	831.76 (NNLO [34])

A.2. Input Variables in the 2-Jets-1-Tag Region

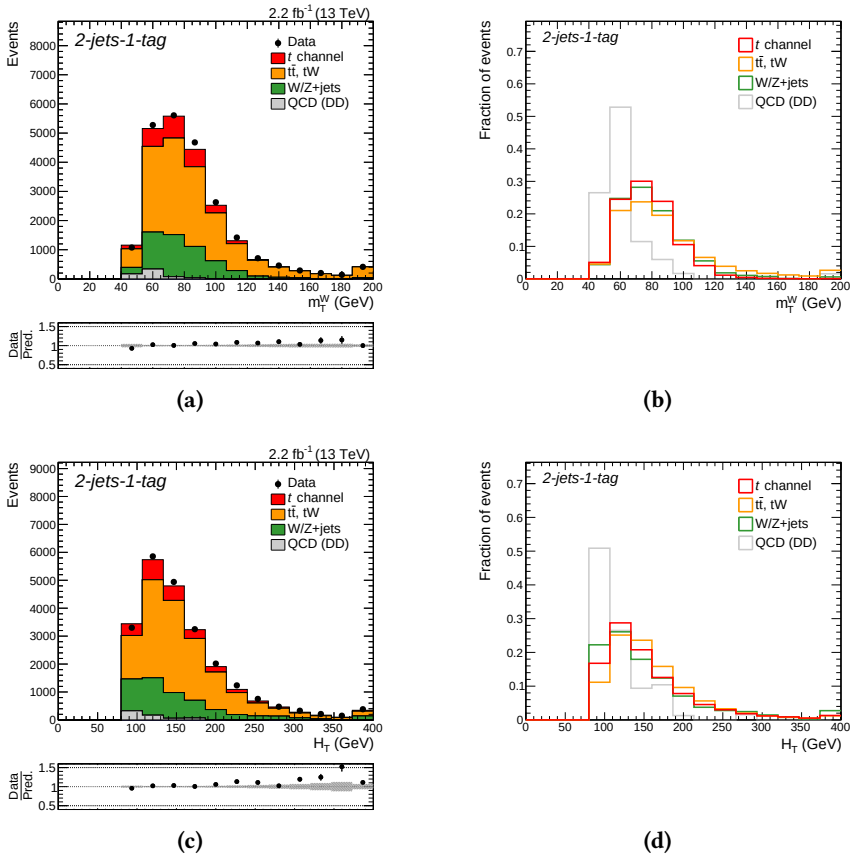


Figure A.1.: Distributions of input variables for the event classification in the 2-jets–1-tag region, ranked 4$^{\text{th}}$ and 5$^{\text{th}}$ in the training of the neural net: the transverse mass of the reconstructed W boson ((a) and (b)) and the scalar sum of the transverse momenta of the two jets with the highest transverse momentum ((c) and (d)). The distributions on the left are scaled to the expected number of events.

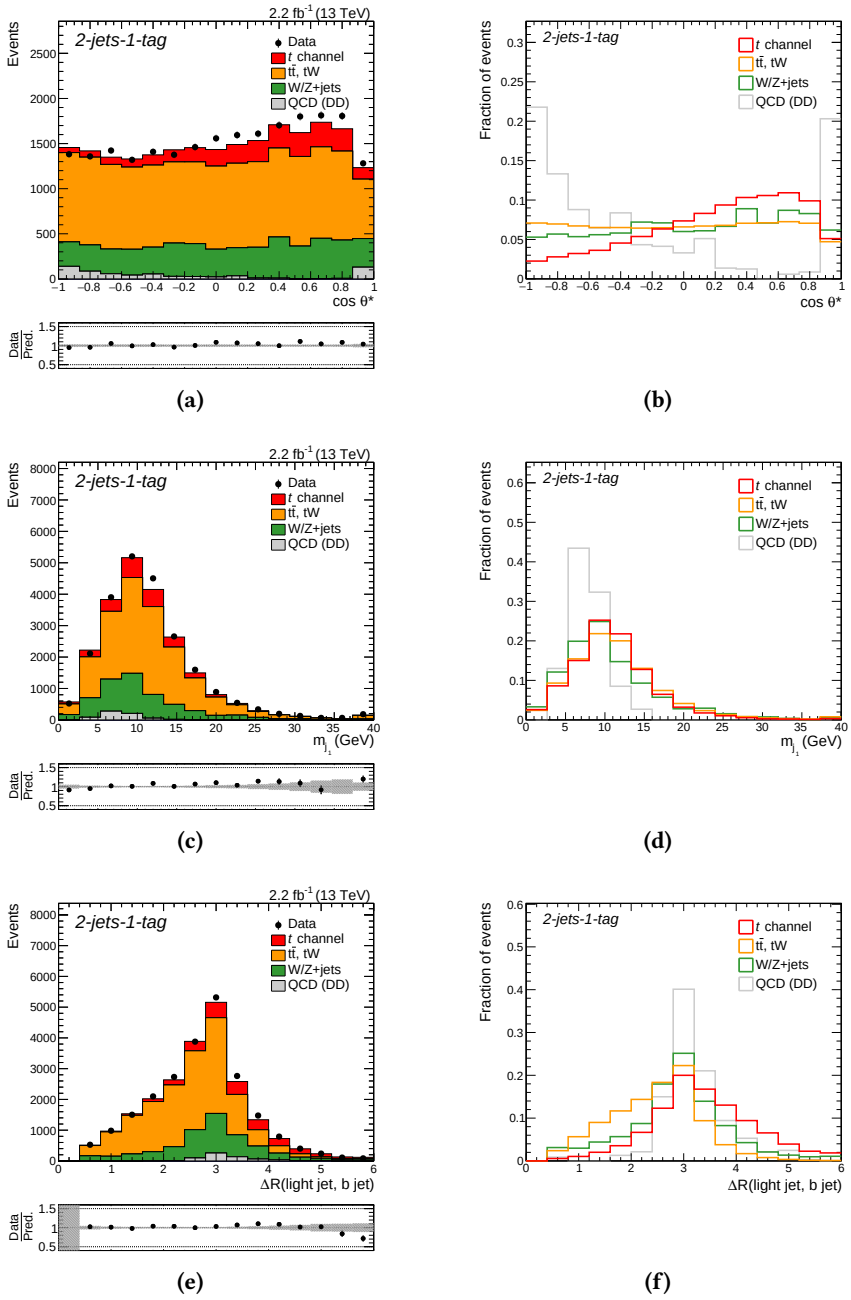


Figure A.2.: Distributions of input variables for the event classification in the 2-jets–1-tag region, ranked 6^{th} to 8^{th} in the training of the neural net: the cosine of the angle between the muon and the light-flavored quark in the rest frame of the top quark ((a) and (b)), the invariant mass of the jet with the largest transverse momentum ((c) and (d)) and ΔR between the momentum vectors of the light-flavored jet and the b-tagged jet ((e) and (f)). The distributions on the left are scaled to the expected number of events.

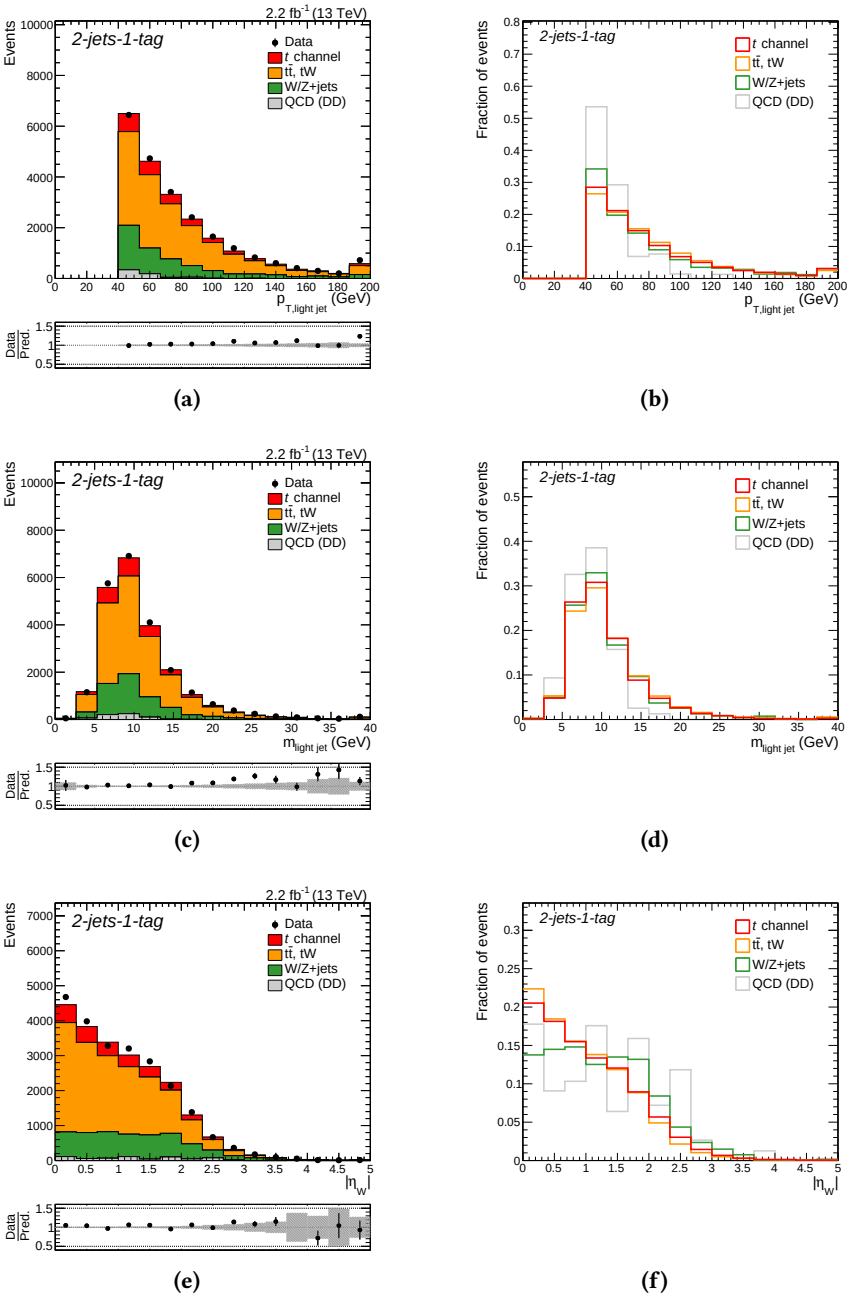


Figure A.3.: Distributions of input variables for the event classification in the 2-jets–1-tag region, ranked 9th to 11th in the training of the neural net: the transverse momentum of light-flavored jet ((a) and (b)), the invariant mass of the light-flavored jet ((c) and (d)) and the absolute value of the pseudorapidity of the reconstructed W boson ((e) and (f)). The distributions on the left are scaled to the expected number of events.

A.3. Input Variables in the 3-Jets-1-Tag Region

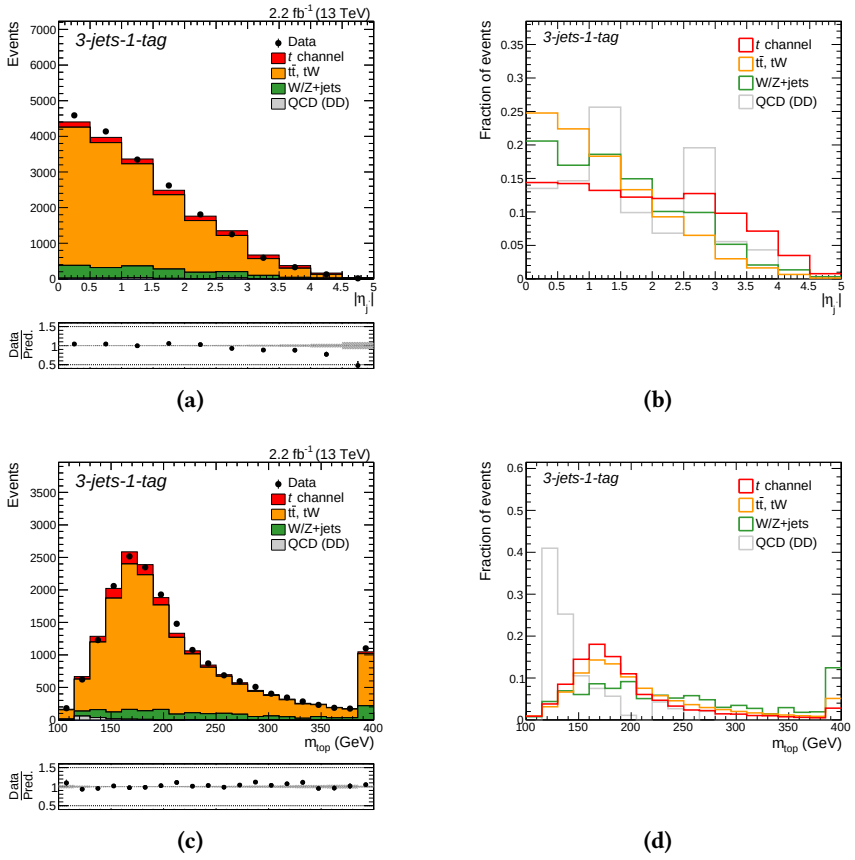


Figure A.4.: Distributions of input variables for the event classification in the 3-jets–1-tag region, ranked 1st and 2nd in the training of the neural net: the absolute value of the pseudorapidity of the light-flavored jet ((a) and (b)) and the invariant mass of the reconstructed top quark ((c) and (d)). The distributions on the left are scaled to the expected number of events.

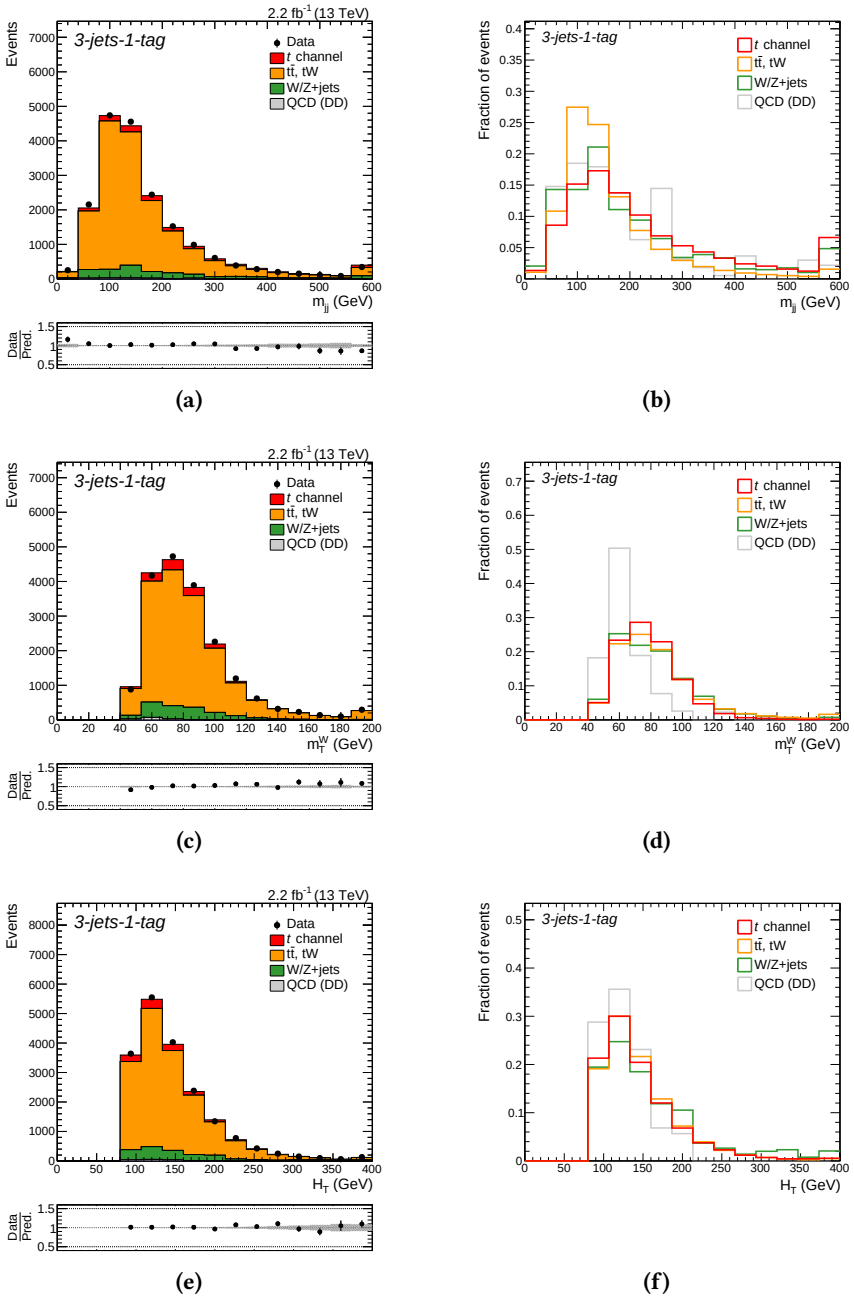


Figure A.5.: Distributions of input variables for the event classification in the 3-jets–1-tag region, ranked 3rd to 5th in the training of the neural net: the invariant mass of the two jets with the highest transverse momentum ((a) and (b)), the transverse mass of the reconstructed W boson ((c) and (d)) and the scalar sum of the transverse momenta of the two jets with the highest transverse momentum ((e) and (f)). The distributions on the left are scaled to the expected number of events.

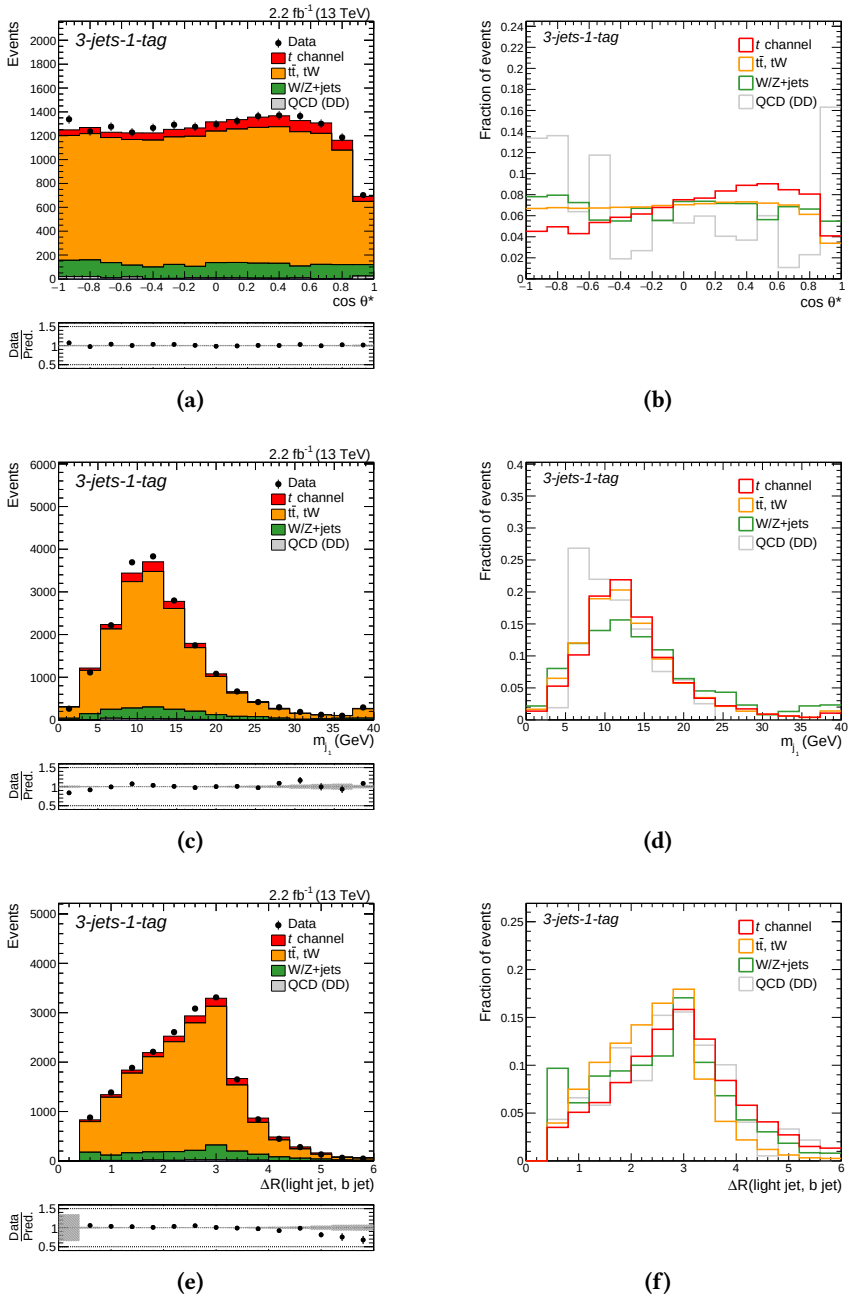


Figure A.6.: Distributions of input variables for the event classification in the 3-jets–1-tag region, ranked 6$^{\text{th}}$ to 8$^{\text{th}}$ in the training of the neural net: the cosine of the angle between the muon and the light-flavored quark in the rest frame of the top quark ((a) and (b)), the invariant mass of the jet with the largest transverse momentum ((c) and (d)) and ΔR between the momentum vectors of the light-flavored jet and the b-tagged jet ((e) and (f)). The distributions on the left are scaled to the expected number of events.

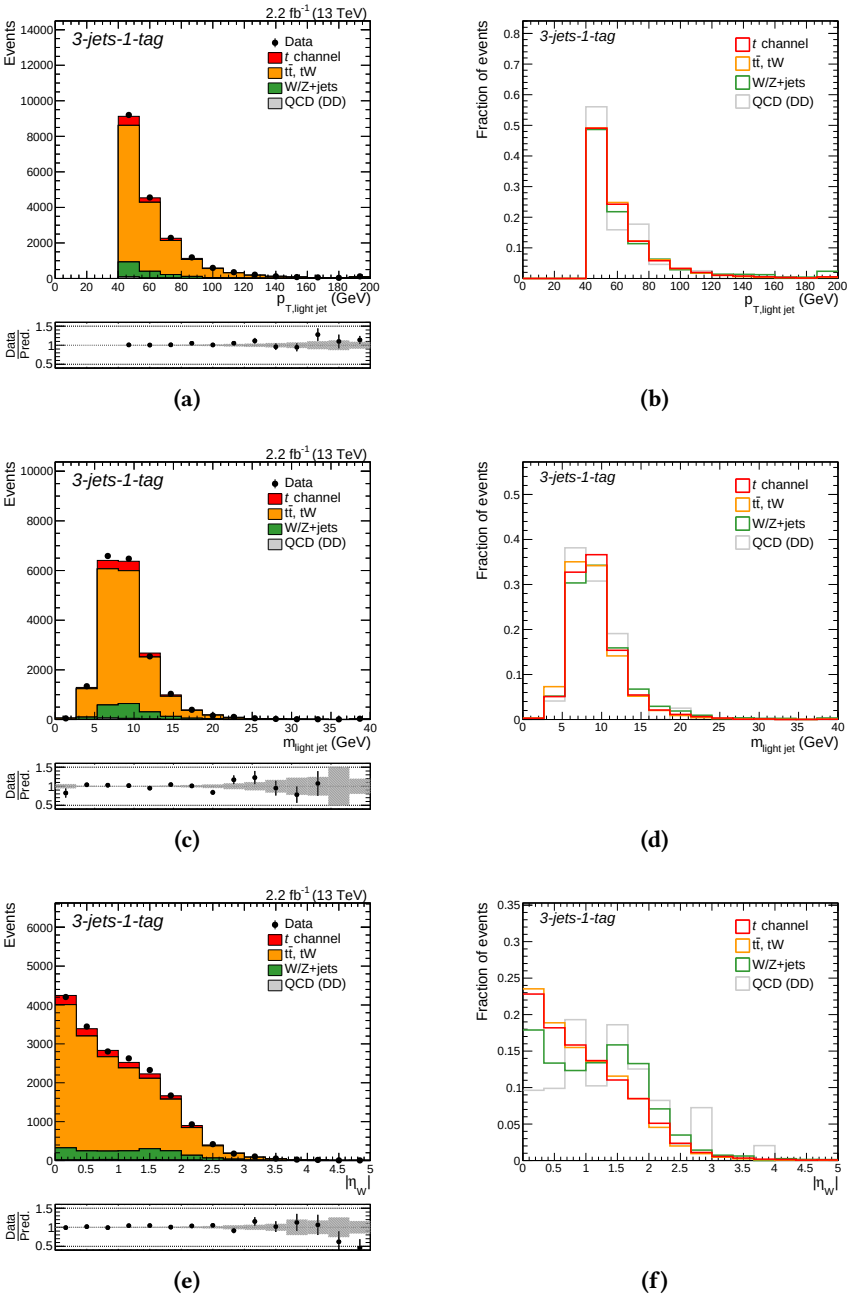


Figure A.7.: Distributions of input variables for the event classification in the 3-jets–1-tag region, ranked 9th to 11th in the training of the neural net: the transverse momentum of light-flavored jet ((a) and (b)), the invariant mass of the light-flavored jet ((c) and (d)) and the absolute value of the pseudorapidity of the reconstructed W boson ((e) and (f)). The distributions on the left are scaled to the expected number of events.

A.4. Input Variables in the 3-Jets-2-Tags Region

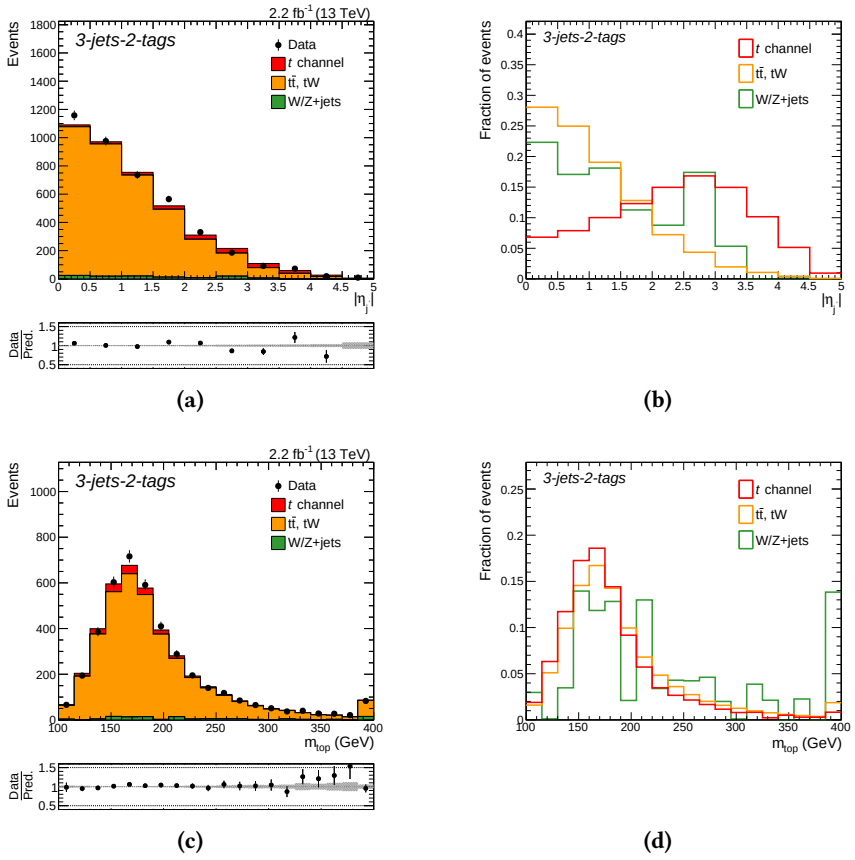


Figure A.8.: Distributions of input variables for the event classification in the 3-jets–2-tags region, ranked 1st and 2nd in the training of the neural net: the absolute value of the pseudorapidity of the light-flavored jet ((a) and (b)) and the invariant mass of the reconstructed top quark ((c) and (d)). The distributions on the left are scaled to the expected number of events.

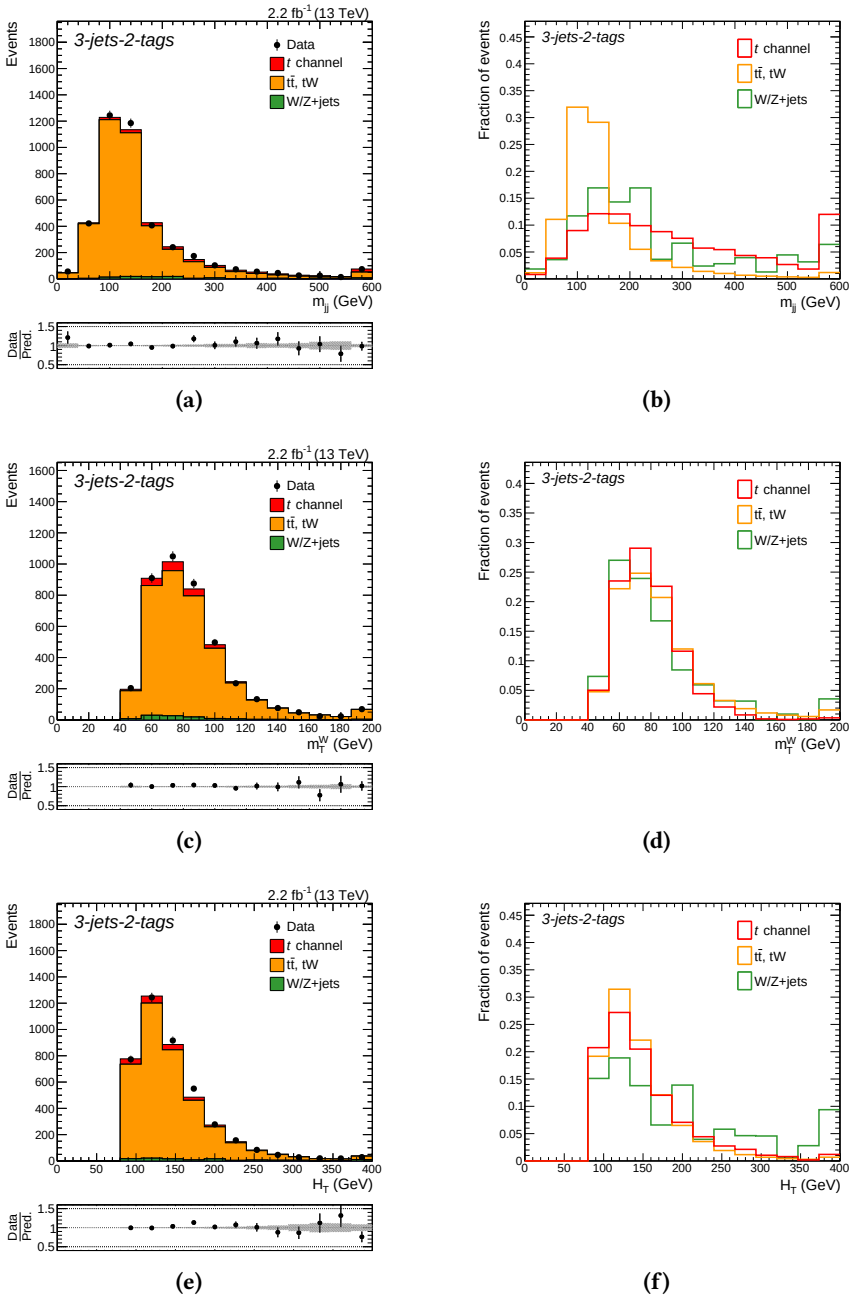


Figure A.9.: Distributions of input variables for the event classification in the 3-jets–2-tags region, ranked 3rd to 5th in the training of the neural net: the invariant mass of the two jets with the highest transverse momentum ((a) and (b)), the transverse mass of the reconstructed W boson ((c) and (d)) and the scalar sum of the transverse momenta of the two jets with the highest transverse momentum ((e) and (f)). The distributions on the left are scaled to the expected number of events.

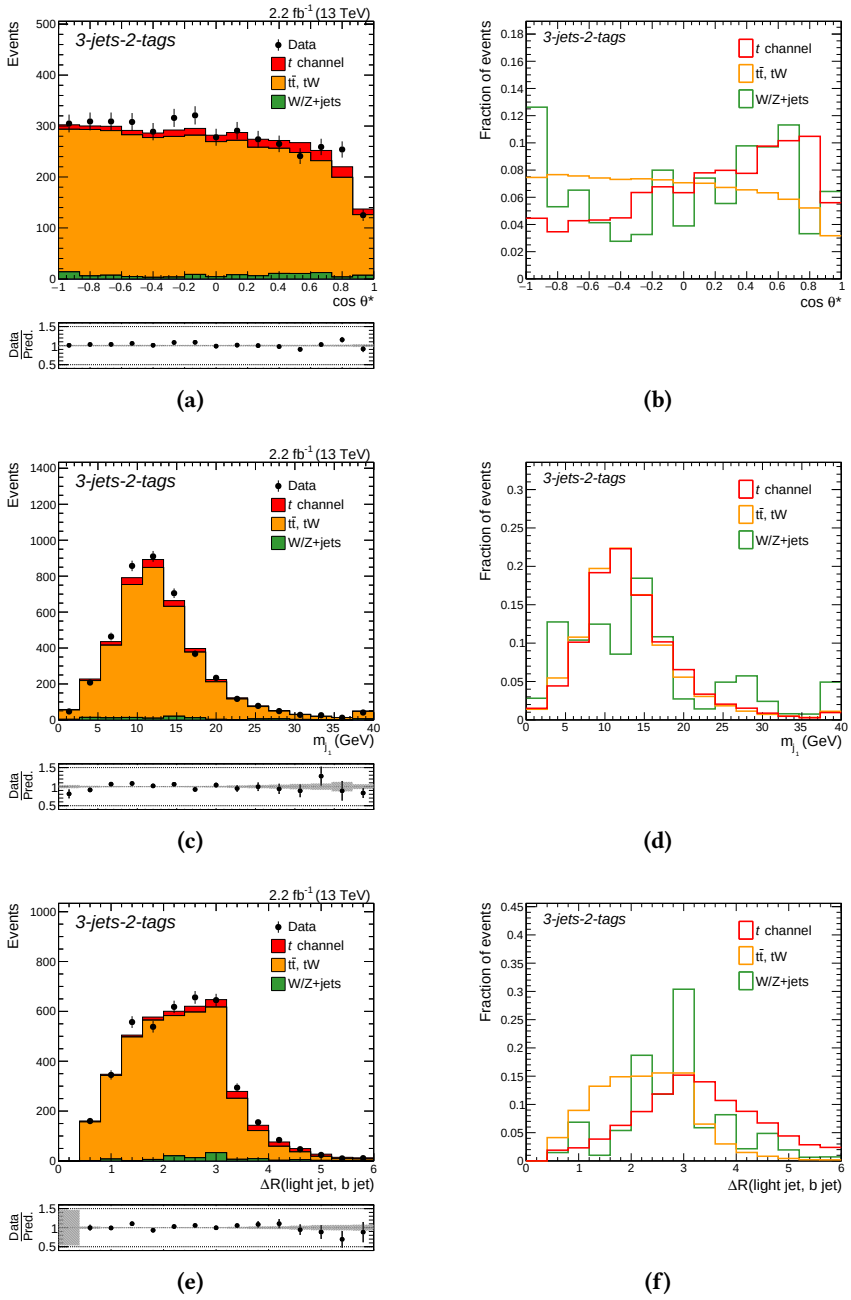


Figure A.10.: Distributions of input variables for the event classification in the 3-jets–2-tags region, ranked 6$^{\text{th}}$ to 8$^{\text{th}}$ in the training of the neural net: the cosine of the angle between the muon and the light-flavored quark in the rest frame of the top quark ((a) and (b)), the invariant mass of the jet with the largest transverse momentum ((c) and (d)) and ΔR between the momentum vectors of the light-flavored jet and the b-tagged jet ((e) and (f)). The distributions on the left are scaled to the expected number of events.

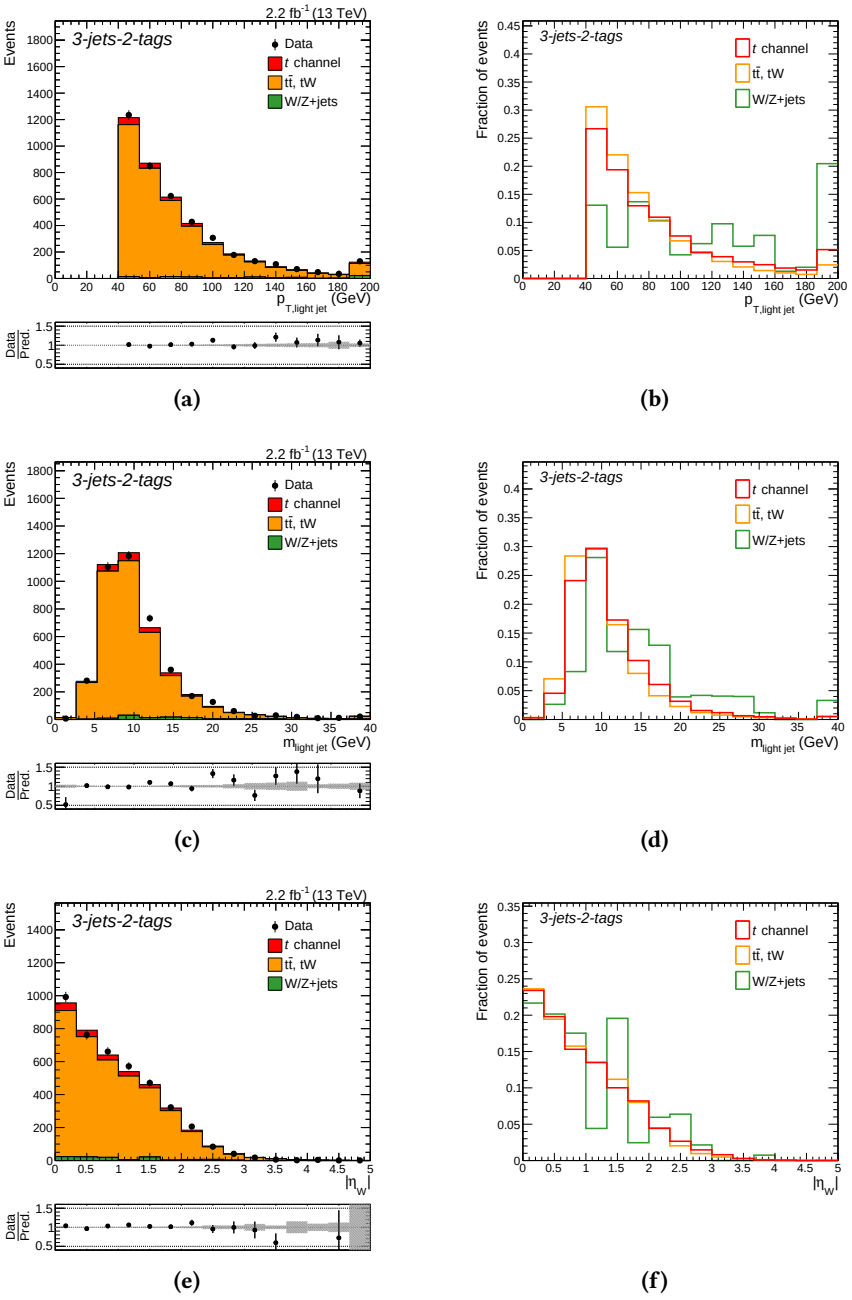


Figure A.11.: Distributions of input variables for the event classification in the 3-jets–2-tags region, ranked 9th to 11th in the training of the neural net: the transverse momentum of light-flavored jet ((a) and (b)), the invariant mass of the light-flavored jet ((c) and (d)) and the absolute value of the pseudorapidity of the reconstructed W boson ((e) and (f)).The distributions on the left are scaled to the expected number of events.

B. Appendix: Search for s-Channel Single Top Quark Production at $\sqrt{s} = 13$ TeV

B.1. Data Sets and MC Simulation Samples

Table B.1.: The data sets of 2016 used for the analysis with the corresponding run ranges and the integrated luminosities.

Period	Run range	Integrated luminosity
/.../Run2016B-03Feb2017_ver2-v2/MINIAOD	272007–275376	5.8 fb^{-1}
/.../Run2016C-03Feb2017-v1/MINIAOD	275657–276283	2.6 fb^{-1}
/.../Run2016D-03Feb2017-v1/MINIAOD	276315–276811	4.2 fb^{-1}
/.../Run2016E-03Feb2017-v1/MINIAOD	276831–277420	4.0 fb^{-1}
/.../Run2016F-03Feb2017-v1/MINIAOD	277772–278808	3.1 fb^{-1}
/.../Run2016G-03Feb2017-v1/MINIAOD	278820–280385	7.5 fb^{-1}
/.../Run2016H-03Feb2017_ver{2,3}-v1/MINIAOD	280919–284044	8.6 fb^{-1}
Total	272007–284044	35.9 fb^{-1}

Table B.2.: The data sets of 2017 used for the analysis with the corresponding run ranges and the integrated luminosities.

Period	Run range	Integrated luminosity
/.../Run2017B-31Mar2018-v1/MINIAOD	297046—299329	4.8 fb^{-1}
/.../Run2017C-31Mar2018-v1/MINIAOD	299368—302029	9.6 fb^{-1}
/.../Run2017D-31Mar2018-v1/MINIAOD	302030—303434	4.2 fb^{-1}
/.../Run2017E-31Mar2018-v1/MINIAOD	303824—304797	9.3 fb^{-1}
/.../Run2017F-31Mar2018-v1/MINIAOD	305040—306462	13.5 fb^{-1}
Total	297046–306462	41.5 fb^{-1}

Table B.3.: List of all MC simulation samples used in the analysis of 2016 data with the number of generated events and the corresponding cross section times branching ratio.The data set name corresponds to the primary data set name, the fragments RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6-v* and MINIAODSIM have been omitted for better readability. If not stated otherwise, the cross section is taken from the generator.

Data set name	Events	Cross section × BR (pb)
ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8	9 651 642	10.32 × 0.3258 (NLO [3, 37])
TTToSemilepton_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	152 720 952	831.76 × 0.4393 (NNLO [3, 34])
TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	79 092 400	831.76 × 0.1061 (NNLO [3, 34])
ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	67 240 808	136.02 (NLO [37])
ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	38 811 017	80.95 (NLO [37])
ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	5 372 991	35.85 × 0.5454 (NNLO [37])
ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	5 425 134	35.85 × 0.5454 (NNLO [37])
WToLNu_0J_13TeV-amcatnloFXFX-pythia8	49 142 195	50132 (NLO)
WToLNu_1J_13TeV-amcatnloFXFX-pythia8	138 948 696	8426 (NLO)
WToLNu_2J_13TeV-amcatnloFXFX-pythia8	258 502 250	3173 (NLO)
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	122 055 388	5765 (NNLO [342])

Table B.4.: List of all additional MC simulation samples used for systematic variations in the analysis of 2016 data with the number of generated events and the corresponding cross section times branching ratio. The data set name corresponds to the primary data set name, the fragments RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6-v* and MINIAODSIM have been omitted for better readability.

Data set name	Events	Cross section × BR (pb)
TT_hdampDOWN_TuneCUETP8M2T4_13TeV-powheg-pythia8	57 862 669	831.76 (NNLO [3, 34])
TT_hdampUP_TuneCUETP8M2T4_13TeV-powheg-pythia8	58 858 606	831.76 (NNLO [3, 34])
TT_TuneCUETP8M2T4_13TeV-powheg-fsrdown-pythia8	155 991 586	831.76 (NNLO [3, 34])
TT_TuneCUETP8M2T4_13TeV-powheg-fsrup-pythia8	152 617 579	831.76 (NNLO [3, 34])
TT_TuneCUETP8M2T4_13TeV-powheg-isrdown-pythia8	146 448 666	831.76 (NNLO [3, 34])
TT_TuneCUETP8M2T4_13TeV-powheg-isrup-pythia8	156 469 815	831.76 (NNLO [3, 34])
TT_TuneCUETP8M2T4down_13TeV-powheg-pythia8	58 338 240	831.76 (NNLO [3, 34])
TT_TuneCUETP8M2T4up_13TeV-powheg-pythia8	58 953 660	831.76 (NNLO [3, 34])
ST_tW_top_5f_MEscaleup_NoFullyHadronicDecays_13TeV-powheg	3 188 774	35.85 × 0.5454 (NNLO [37])
ST_tW_top_MEscaledown_5f_NoFullyHadronicDecays_13TeV-powheg	3 052 085	35.85 × 0.5454 (NNLO [37])
ST_tW_top_5f_PSscaleup_NoFullyHadronicDecays_13TeV-powheg-pythia8	3 124 934	35.85 × 0.5454 (NNLO [37])
ST_tW_top_PSscaledown_5f_NoFullyHadronicDecays_13TeV-powheg-pythia8	3 181 663	35.85 × 0.5454 (NNLO [37])
ST_tW_antitop_5f_MEscaleup_NoFullyHadronicDecays_13TeV-powheg	1 607 001	35.85 × 0.5454 (NNLO [37])
ST_tW_antitop_MEscaledown_5f_NoFullyHadronicDecays_13TeV-powheg	1 575 183	35.85 × 0.5454 (NNLO [37])
ST_tW_antitop_5f_PSscaleup_NoFullyHadronicDecays_13TeV-powheg-pythia8	1 628 527	35.85 × 0.5454 (NNLO [37])
ST_tW_antitop_PSscaledown_5f_NoFullyHadronicDecays_13TeV-powheg-pythia8	1 628 344	35.85 × 0.5454 (NNLO [37])

Table B.5.: List of all MC simulation samples used in the analysis of 2017 data with the number of generated events and the corresponding cross section times branching ratio. The data set name corresponds to the primary data set name, the fragments RunIIFall17MiniAODv2-PU2017_12Apr2018_94X_mc2017_realistic_-v14-v* and MINIAODSIM have been omitted for better readability. If not stated otherwise, the cross section is taken from the generator.

Data set name	Events	Cross section × BR (pb)
ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8	9 567 962	10.32 × 0.3258 (NLO [3, 37])
ST_s-channel_4f_leptonDecays_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	9 906 084	10.32 × 0.3258 (NLO [3, 37])
TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8	111 316 892	831.76 × 0.4393 (NNLO [3, 34])
TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8	66 973 657	831.76 × 0.1061 (NNLO [3, 34])
ST_t-channel_top_4f_inclusiveDecays_TuneCP5_13TeV-powhegV2-madspin-pythia8	5 865 565	136.02 (NLO [37])
ST_t-channel_antitop_4f_inclusiveDecays_TuneCP5_13TeV-powhegV2-madspin-pythia8	3 939 818	80.95 (NLO [37])
ST_tW_top_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	7 659 497	35.85 (NNLO [37])
ST_tW_antitop_5f_inclusiveDecays_TuneCP5_PSweights_13TeV-powheg-pythia8	7 993 109	35.85 (NNLO [37])
WJetsToLNu_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8	130 432 877	8426 (NLO)
WJetsToLNu_2J_TuneCP5_13TeV-amcatnloFXFX-pythia8	97 825 503	3173 (NLO)
DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8	182 349 718	5765 (NNLO [342])

Table B.6.: List of all additional MC simulation samples used for systematic variations in the analysis of 2017 data with the number of generated events and the corresponding cross section times branching ratio. The data set name corresponds to the primary data set name, the fragments RunIIFall17MiniAODv2-PU2017_12Apr2018_94X_mc2017_realistic_v14-v* and MINIAODSIM have been omitted for better readability.

Data set name	Events	Cross section × BR (pb)
TTToSemiLeptonic_hdampDOWN_TuneCP5_PSweights_13TeV-powheg-pythia8	26 365 829	831.76 × 0.4393 (NNLO [3, 34])
TTToSemiLeptonic_hdampUP_TuneCP5_PSweights_13TeV-powheg-pythia8	27 189 487	831.76 × 0.4393 (NNLO [3, 34])
TTToSemiLeptonic_TuneCP5down_PSweights_13TeV-powheg-pythia8	26 748 353	831.76 × 0.4393 (NNLO [3, 34])
TTToSemiLeptonic_TuneCP5up_PSweights_13TeV-powheg-pythia8	26 164 030	831.76 × 0.4393 (NNLO [3, 34])
TTTo2L2Nu_hdampDOWN_TuneCP5_PSweights_13TeV-powheg-pythia8	5 336 170	831.76 × 0.1061 (NNLO [3, 34])
TTTo2L2Nu_hdampUP_TuneCP5_PSweights_13TeV-powheg-pythia8	5 413 099	831.76 × 0.1061 (NNLO [3, 34])
TTTo2L2Nu_TuneCP5down_PSweights_13TeV-powheg-pythia8	5 474 710	831.76 × 0.1061 (NNLO [3, 34])
TTTo2L2Nu_TuneCP5up_PSweights_13TeV-powheg-pythia8	5 499 509	831.76 × 0.1061 (NNLO [3, 34])

B.2. Corrections to Simulated Events

B.2.1. Number of Pileup Interactions

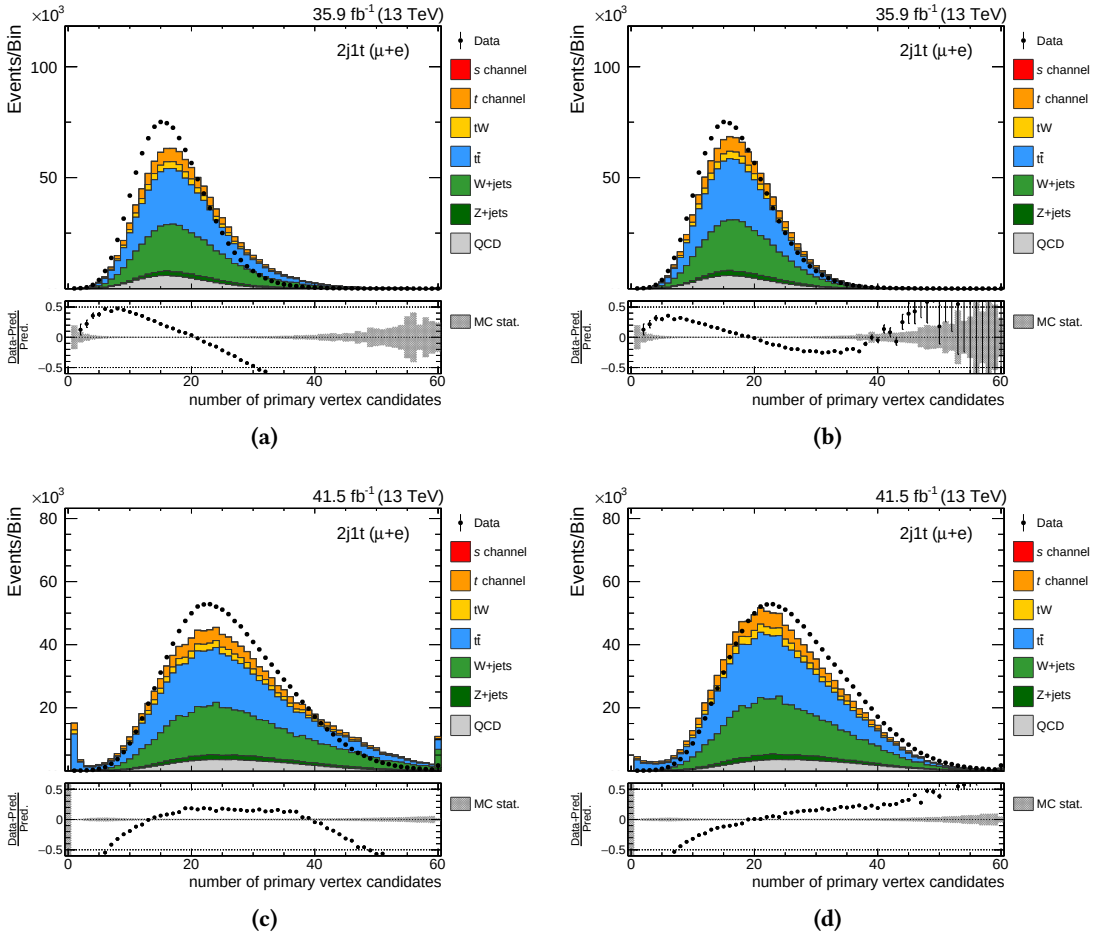


Figure B.1.: Effect of the pileup reweighting on the distribution of the number of primary vertices in an event. The distribution shows disagreement between data and simulation for data taken in 2016 (a) and 2017 (c). The corresponding simulation is corrected to resemble the observed distribution ((b) and (d)). The simulation is scaled to the expected number of events.

B.2.2. Lepton Efficiencies

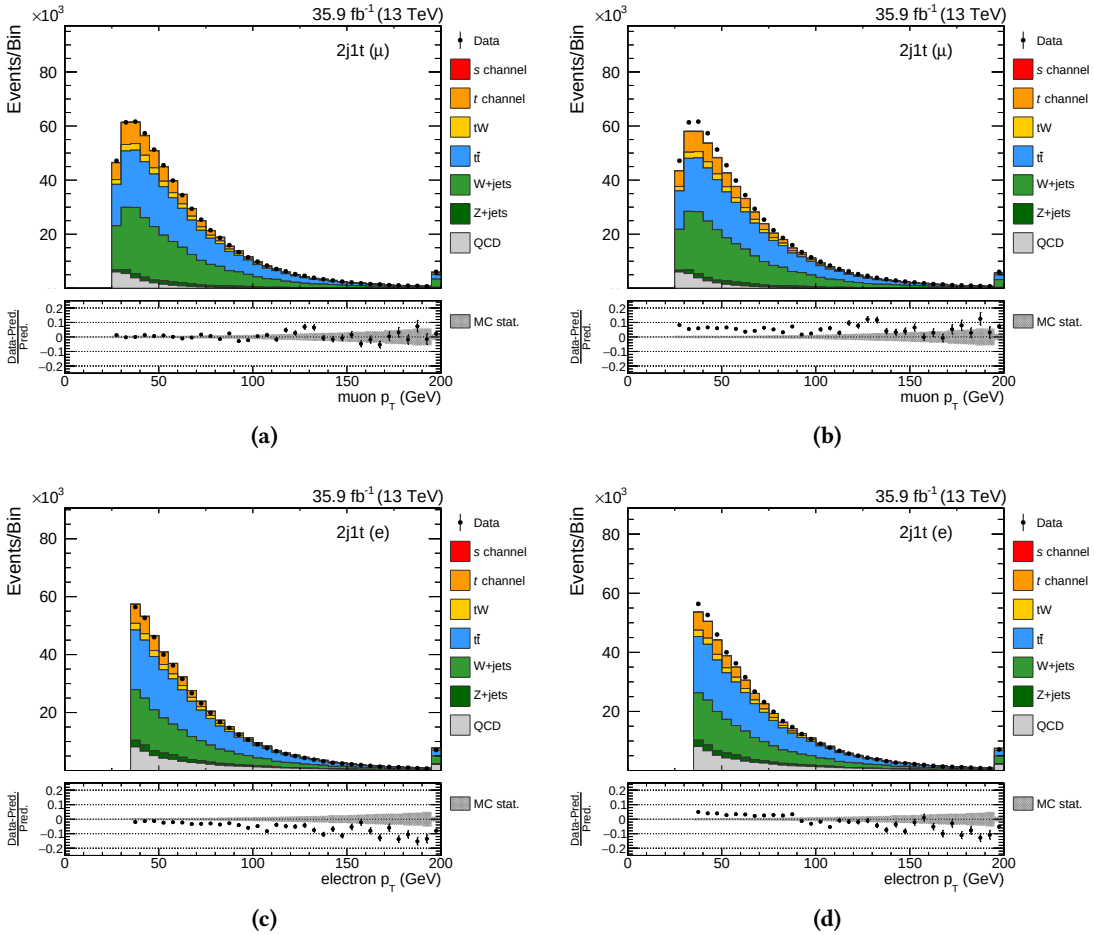


Figure B.2.: Effect of correcting the simulation for different lepton efficiencies in simulation and data of 2016. Dedicated scale factors are applied to the muon (a) and electron (c) channel to provide a more accurate description ((b) and (d)). The simulation is scaled to the expected number of events.

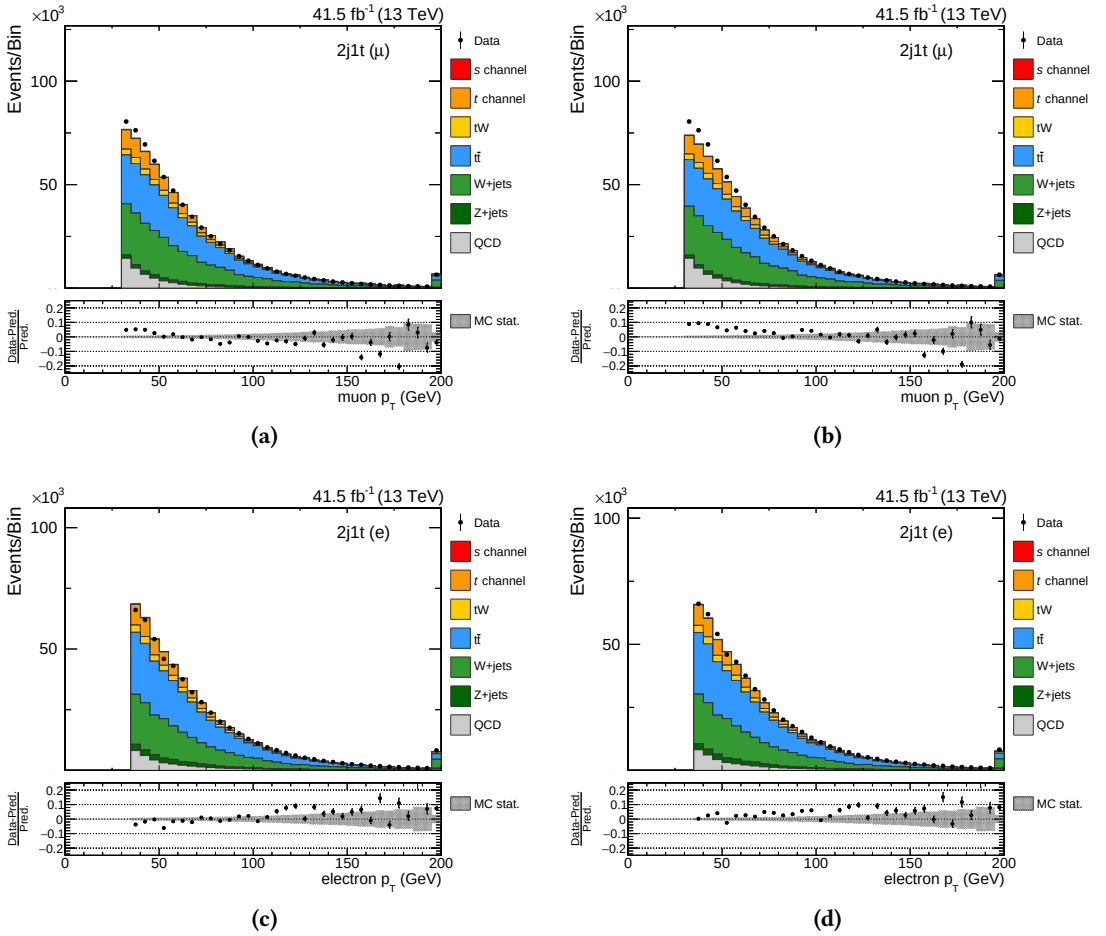


Figure B.3.: Effect of correcting the simulation for different lepton efficiencies in simulation and data of 2017. Dedicated scale factors are applied to the muon (a) and electron (c) channel to provide a more accurate description ((b) and (d)). The simulation is scaled to the expected number of events.

B.2.3. b Tagging Efficiencies

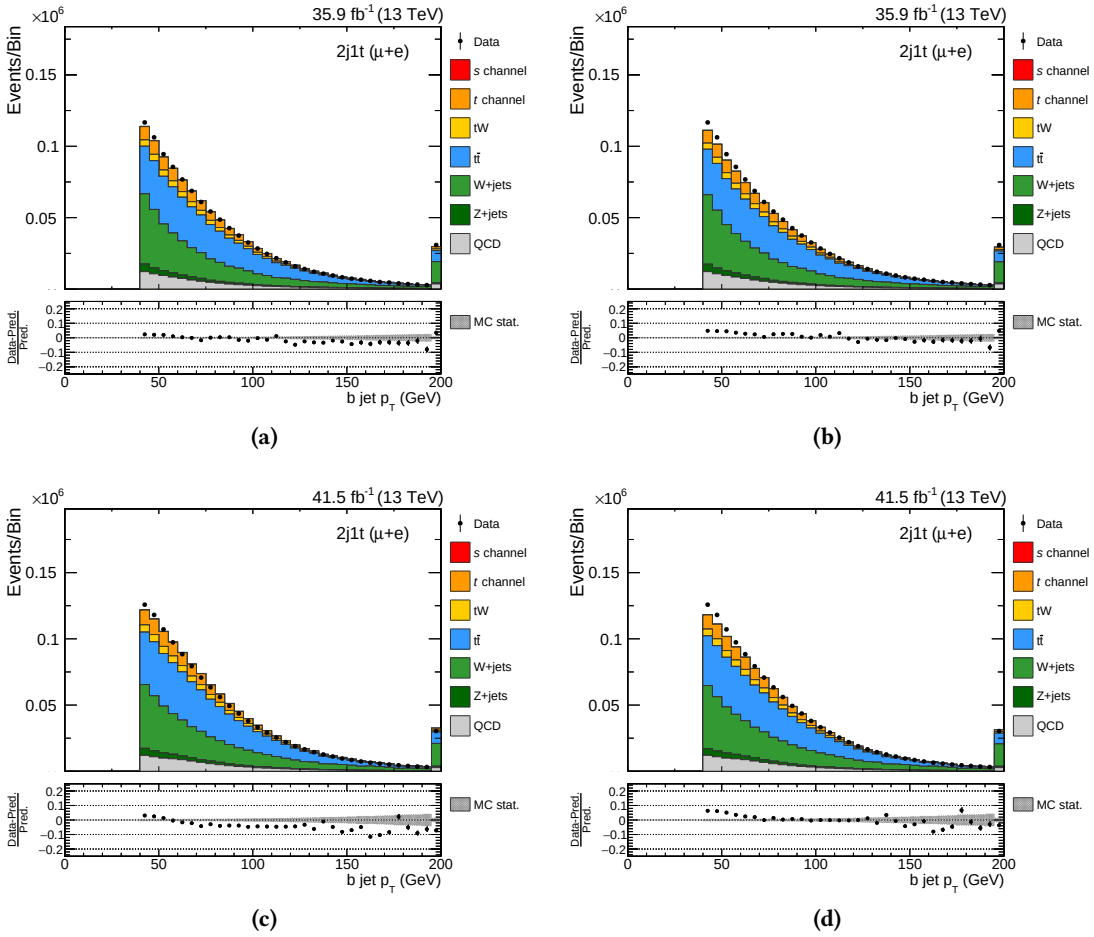


Figure B.4.: Distribution of the transverse momentum of the b-tagged jet in the 2-jets–1-tag region for uncorrected simulation in 2016 (a) and 2017 (c). The simulation is corrected ((b) and (d)) to take into account different tagging efficiencies for simulation and data. The simulation is scaled to the expected number of events.

B.3. QCD Estimation in the 2-jets–1-tag Control Region

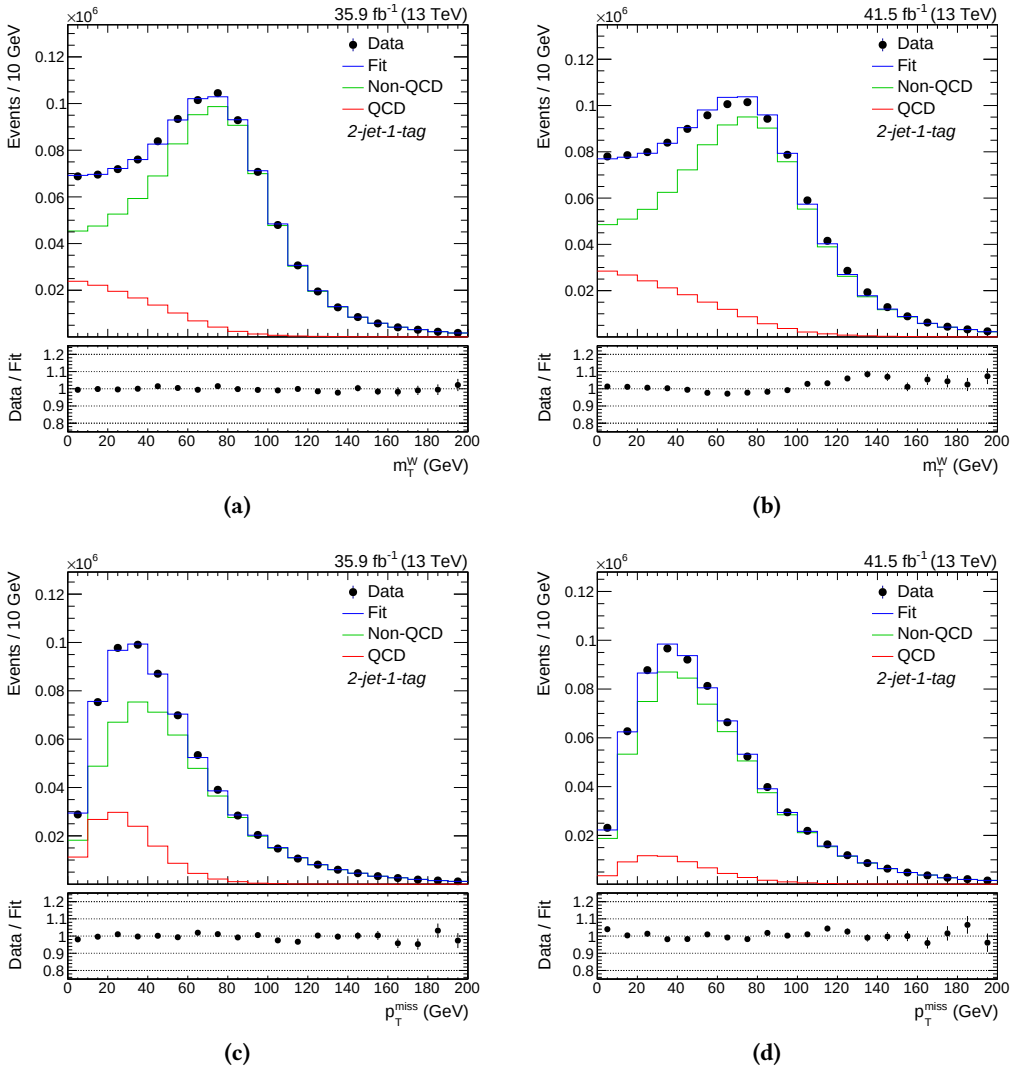


Figure B.5.: QCD template fit in the 2-jets–1-tag control region for the muon channel in 2016 (a), the muon channel in 2017 (c), the electron channel in 2016 (c) and the electron channel in 2017 (d). All fits show sufficient agreement between the prediction and data.

B.4. QCD Estimation in the 3-jets–1-tag Control Region

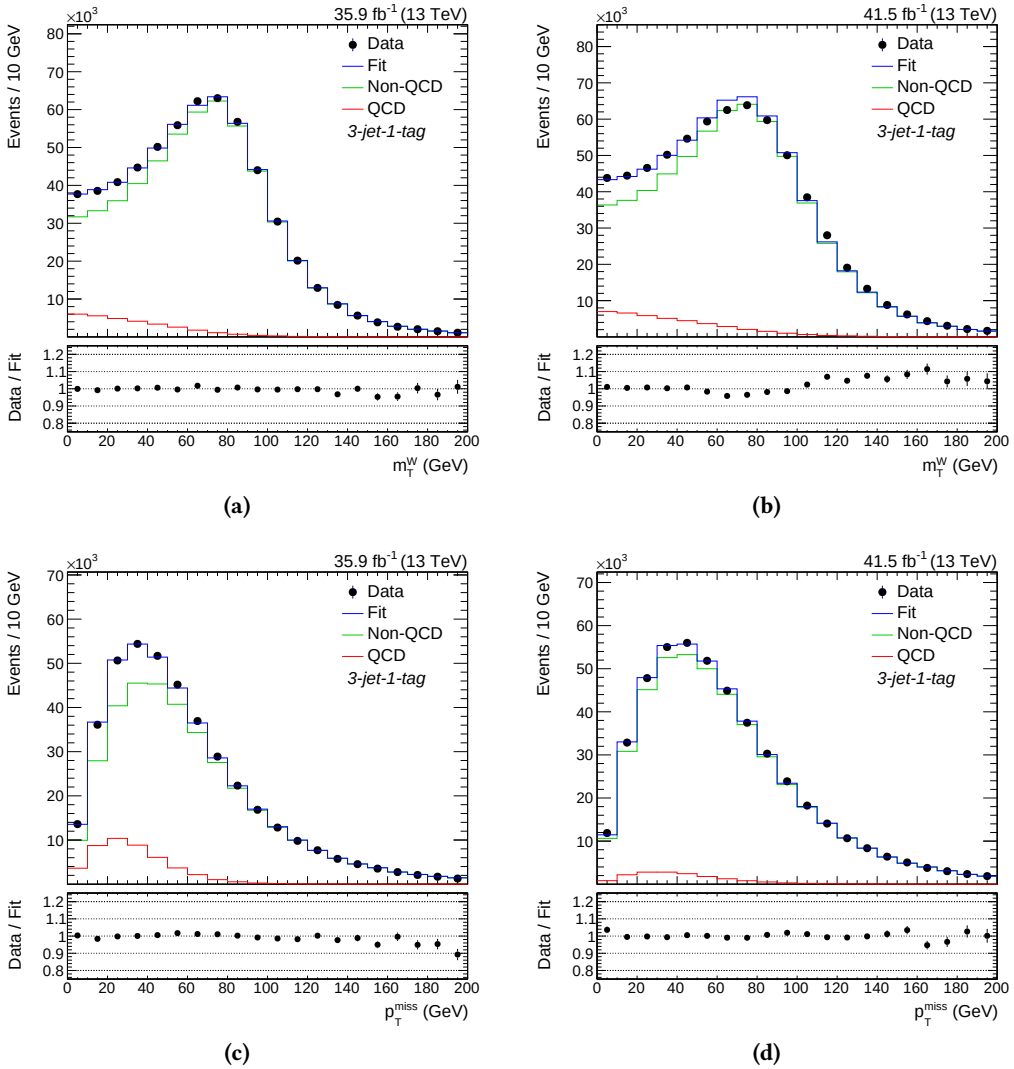


Figure B.6.: QCD template fit in the 3-jets–1-tag control region for the muon channel in 2016 (a), the muon channel in 2017 (c), the electron channel in 2016 (c) and the electron channel in 2017 (d). All fits show sufficient agreement between the prediction and data.

B.5. Input Variables for the DNN

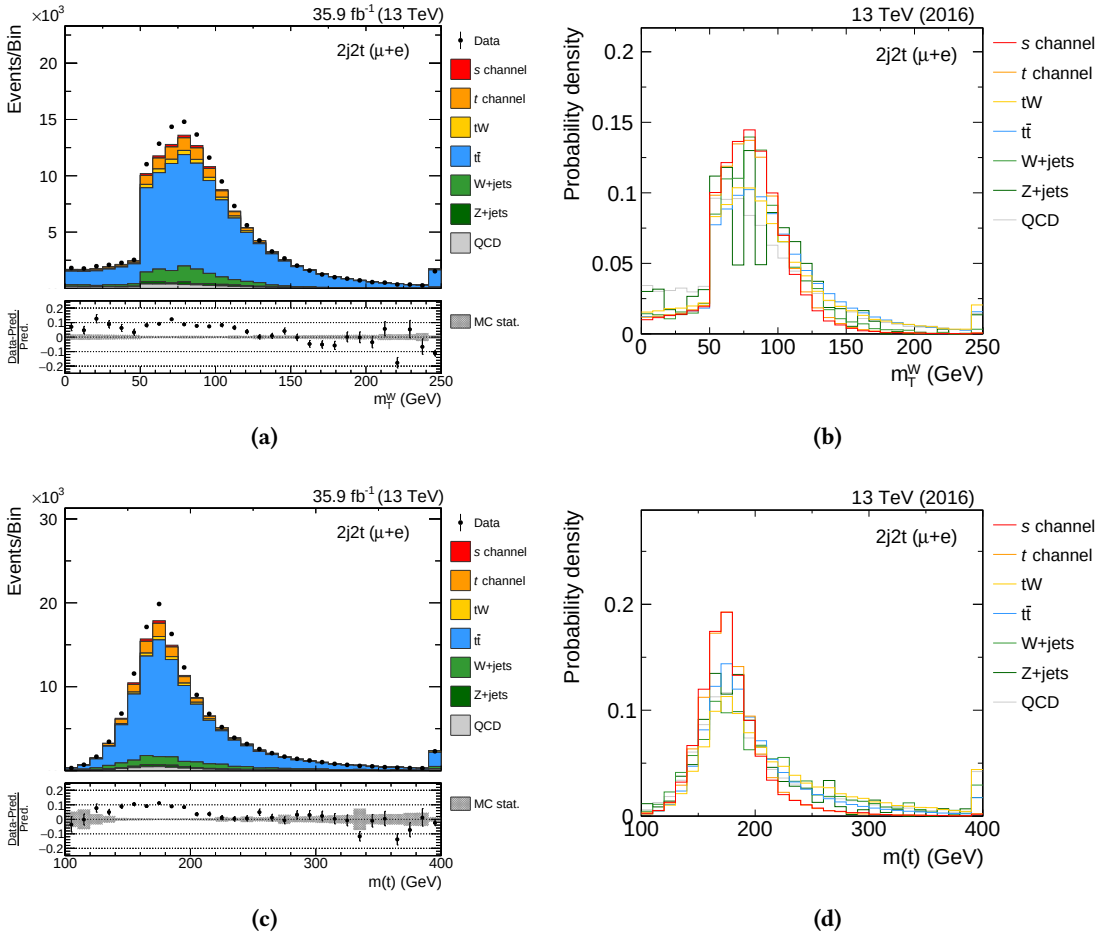


Figure B.7.: Distributions of input variables for the event classification in 2016, ranked 4^{th} and 5^{th} in the training of the BDT: the transverse mass of the reconstructed W boson ((a) and (b)) and the invariant mass of the reconstructed top quark ((c) and (d)). The distributions on the left are scaled to the expected number of events.

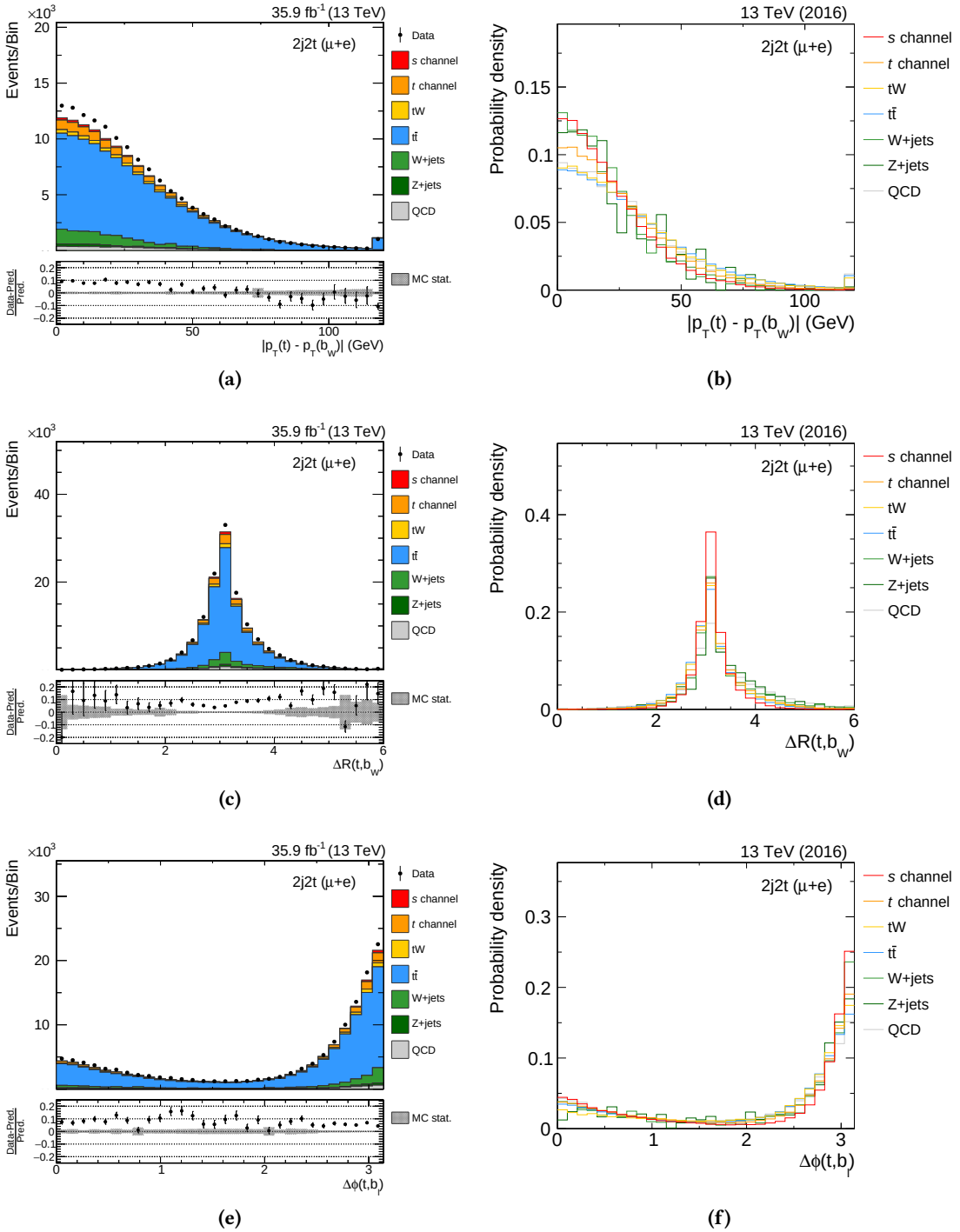


Figure B.8.: Distributions of input variables for the event classification in 2016, ranked 6$^{\text{th}}$ to 8$^{\text{th}}$ in the training of the BDT: the absolute difference of transverse momenta of the reconstructed top quark and the b jet from the time-like W boson ((a) and (b)), ΔR between the reconstructed top quark and the b jet from the time-like W boson ((c) and (d)) and $\Delta\phi$ between the reconstructed top quark and the leading b jet ((e) and (f)). The distributions on the left are scaled to the expected number of events.

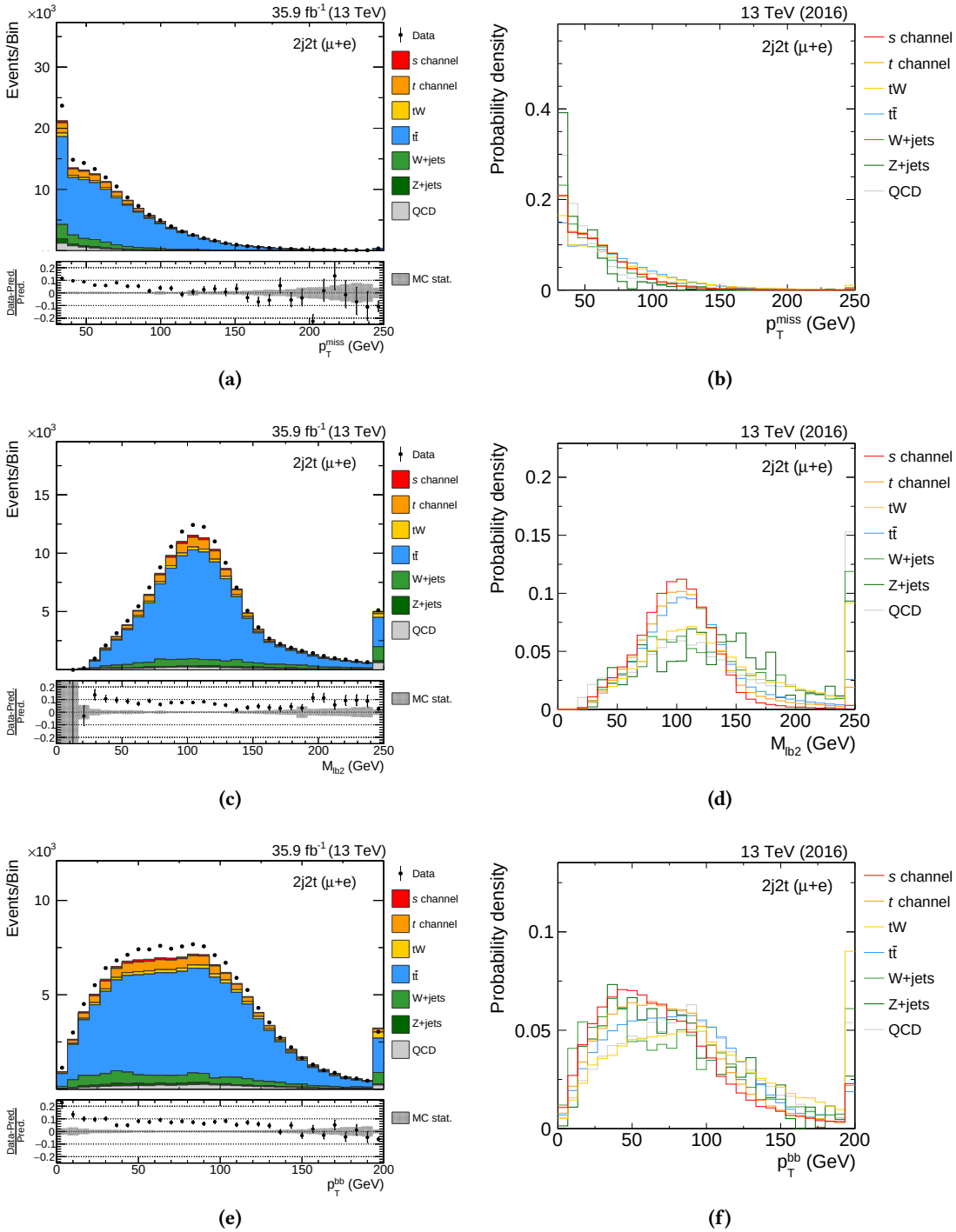


Figure B.9.: Distributions of input variables for the event classification in 2016, ranked 9th to 11th in the training of the BDT: the missing transverse momentum ((a) and (b)), the invariant mass of the lepton and the subleading b jet ((c) and (d)) and the vectorial sum of transverse momenta of the two b jets of the event ((e) and (f)). The distributions on the left are scaled to the expected number of events.

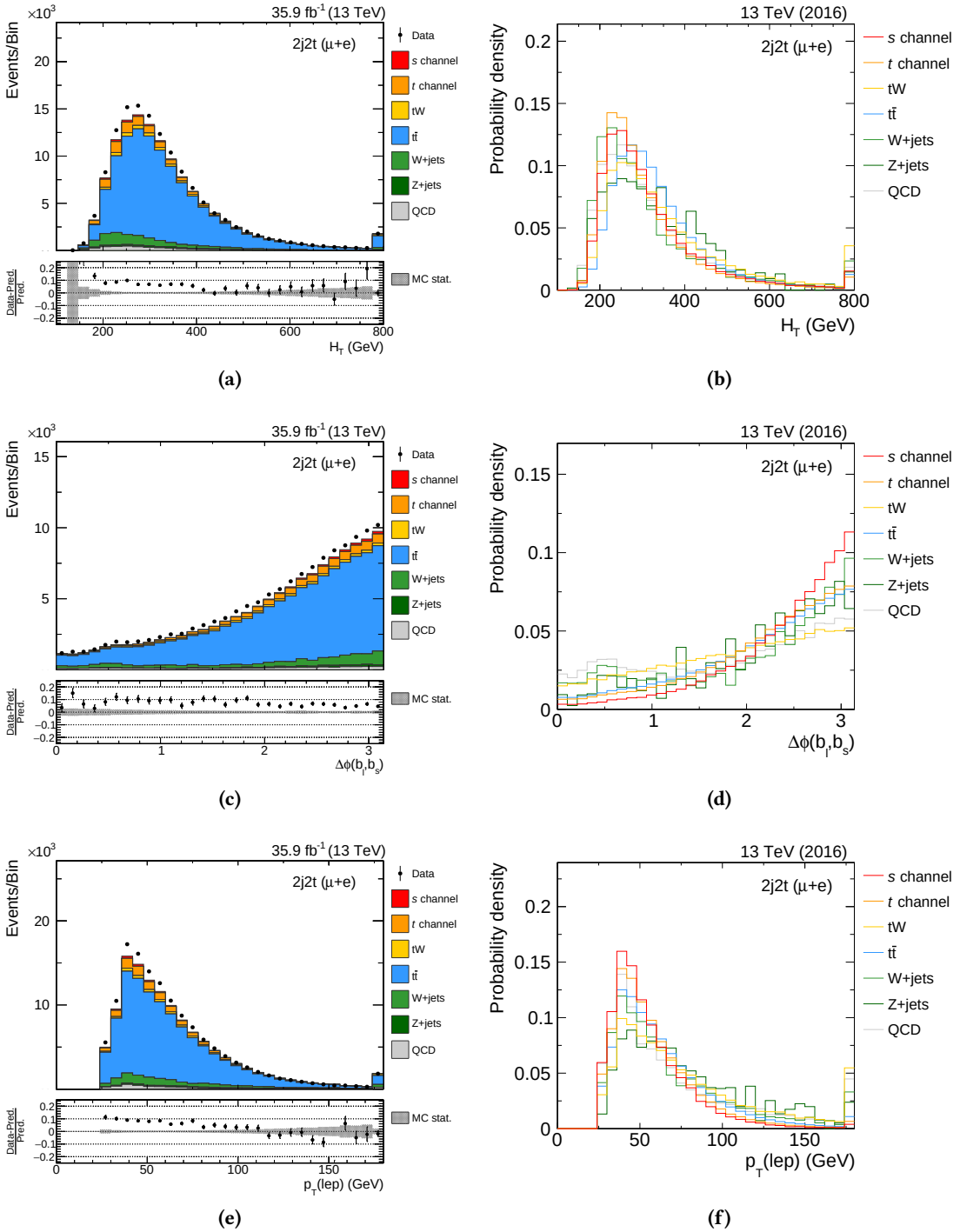


Figure B.10.: Distributions of input variables for the event classification in 2016, ranked 12^{th} to 14^{th} in the training of the BDT: the scalar p_{T} sum of all b jets, the isolated lepton and $p_{\text{T}}^{\text{miss}}$ in the event ((a) and (b)), $\Delta\phi$ between the two b jets of the event ((c) and (d)) and the transverse momentum of the lepton ((e) and (f)). The distributions on the left are scaled to the expected number of events.

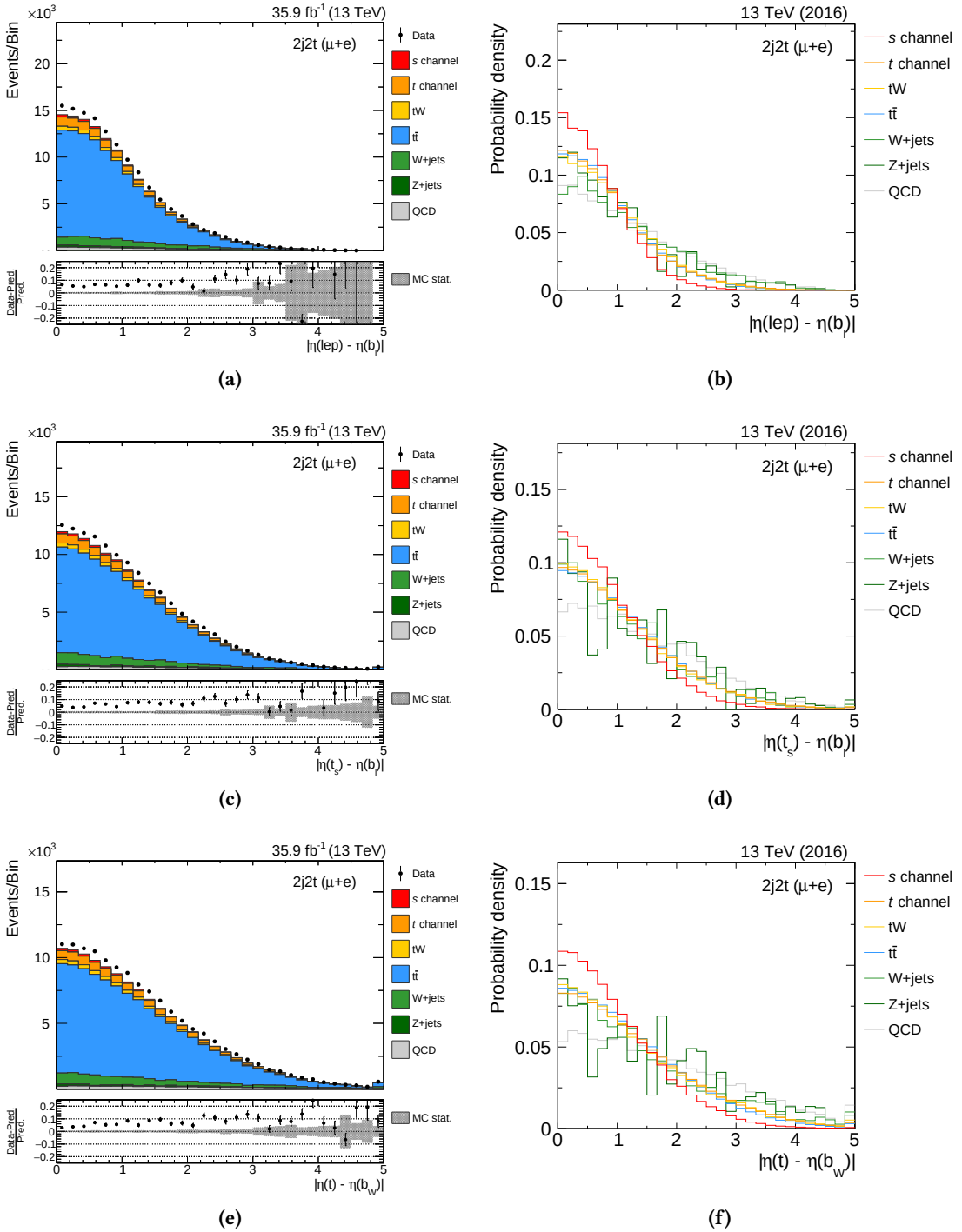


Figure B.11.: Distributions of input variables for the event classification in 2016, ranked 15$^{\text{th}}$ to 17$^{\text{th}}$ in the training of the BDT: the absolute difference of the pseudorapidity between the charged lepton and the leading b jet ((a) and (b)), the absolute difference of the pseudorapidity between the top quark reconstructed from the subleading b jet and the leading b jet ((c) and (d)) and the absolute difference of the pseudorapidity between the reconstructed top quark and the b jet from the time-like W boson ((e) and (f)). The distributions on the left are scaled to the expected number of events.

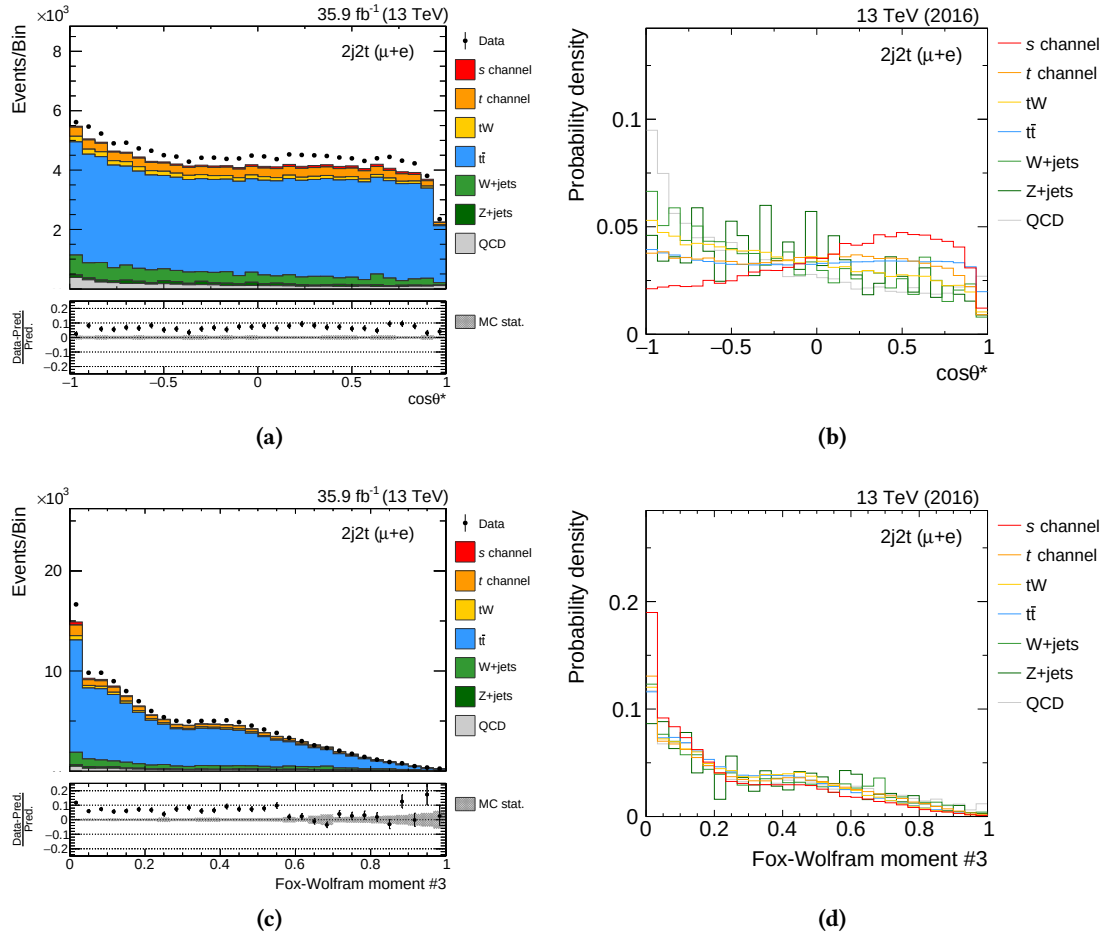


Figure B.12.: Distributions of input variables for the event classification in 2016, ranked 18th and 19th in the training of the BDT: the cosine of the angle between the lepton from the top quark decay and the b jet from the time-like W boson ((a) and (b)) and the third-order Fox-Wolfram moment of the event ((c) and (d)). The distributions on the left are scaled to the expected number of events.

B.6. Fit Result with 2016 Data

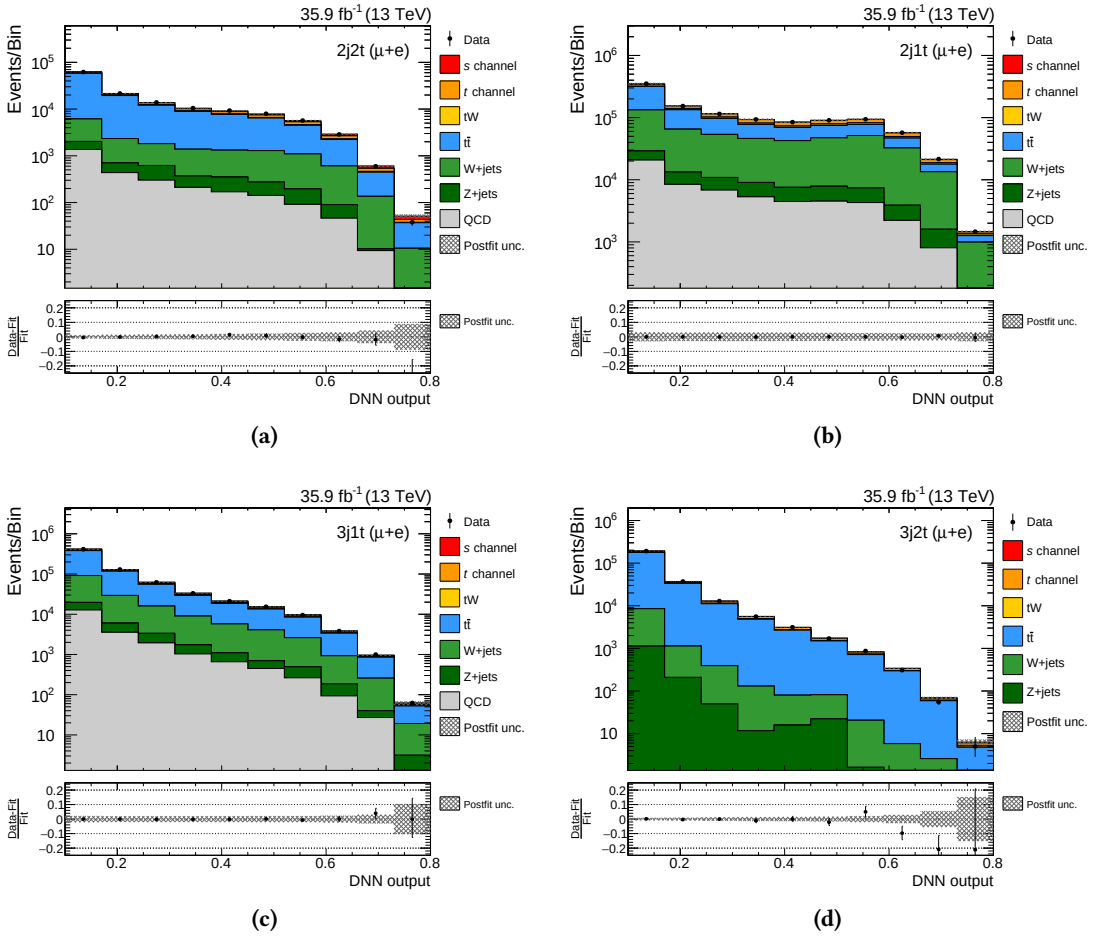


Figure B.13.: Distributions for the DNN classifier output in 2016 for the 2-jets–2-tags signal region (a), as well as the 2-jets–1-tag (b), 3-jets–1-tag (c) and 3-jets–2-tags (d) control regions, after the maximum-likelihood fit to the 2016 data. The simulation is scaled to the result of the fit. Good agreement is observed between simulation and data across all regions.

B.7. Fit Result with 2017 Data

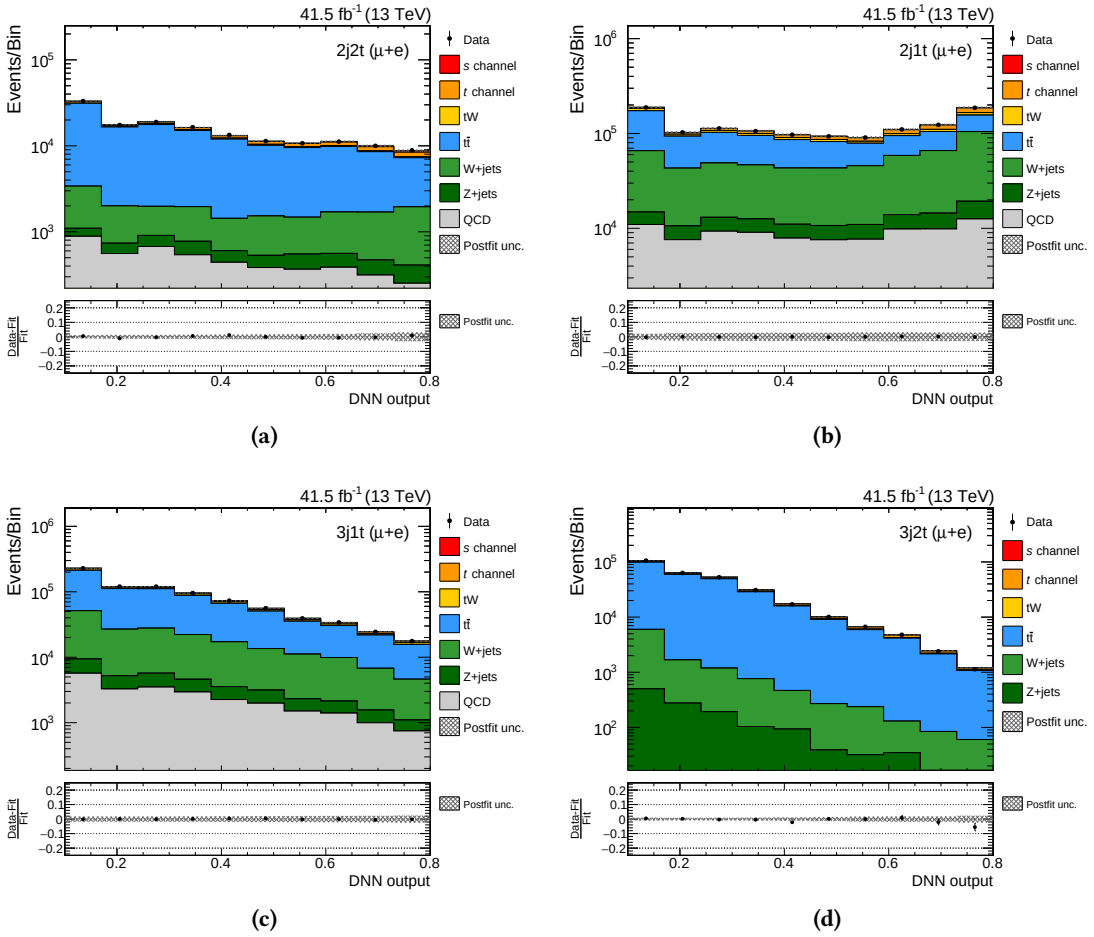


Figure B.14.: Distributions for the DNN classifier output in 2017 for the 2-jets–2-tags signal region (a), as well as the 2-jets–1-tag (b), 3-jets–1-tag (c) and 3-jets–2-tags (d) control regions, after the maximum-likelihood fit to the 2017 data. The simulation is scaled to the result of the fit. Good agreement is observed between simulation and data across all regions.

B.8. Pulls in the Combined Fit

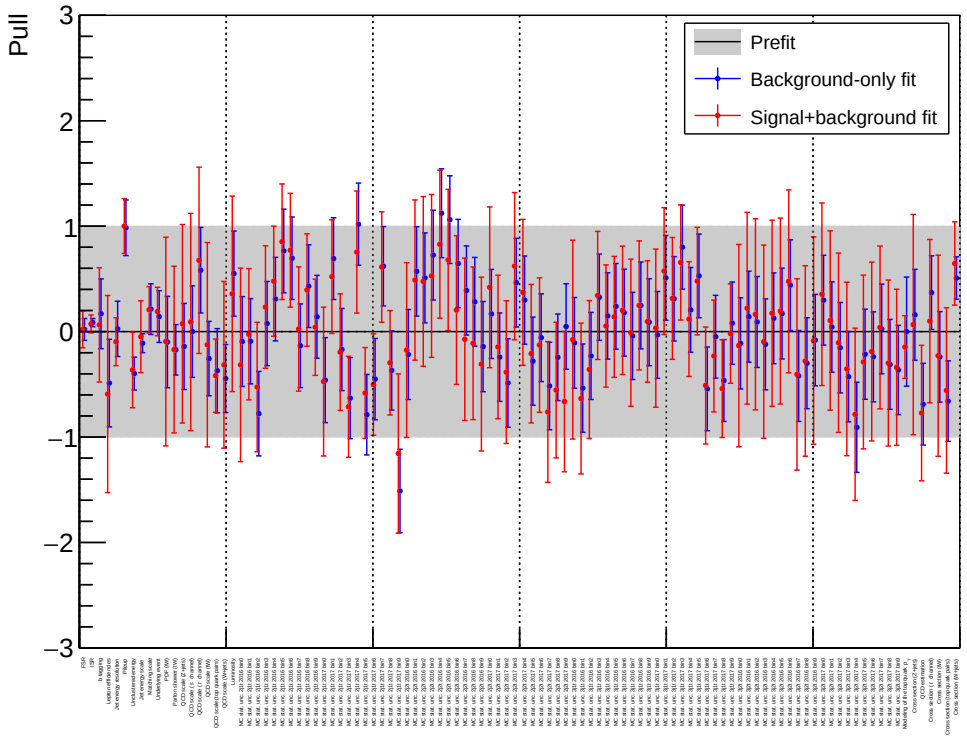


Figure B.15.: Pulls of all nuisance parameters in the combined fit for the background-only fit (blue) and the signal+background fit (red). The pull of a nuisance parameter shows how the value of the nuisance parameter changes during the maximum-likelihood fit with respect to its initial (prefit) value and uncertainty.

C. Appendix: Search for tH Production with $H \to b\bar{b}$ at $\sqrt{s} = 13$ TeV and Study of Higgs Boson Couplings

C.1. Data Sets and MC Samples

Table C.1.: The data sets of 2016 used for the analysis with the corresponding run ranges and the integrated luminosities.

Period	Run range	Integrated luminosity
/.../Run2016B-03Feb2017_ver2-v2/MINIAOD	272007–275376	$5.8\,\text{fb}^{-1}$
/.../Run2016C-03Feb2017-v1/MINIAOD	275657–276283	$2.6\,\text{fb}^{-1}$
/.../Run2016D-03Feb2017-v1/MINIAOD	276315–276811	$4.2\,\text{fb}^{-1}$
/.../Run2016E-03Feb2017-v1/MINIAOD	276831–277420	$4.0\,\text{fb}^{-1}$
/.../Run2016F-03Feb2017-v1/MINIAOD	277772–278808	$3.1\,\text{fb}^{-1}$
/.../Run2016G-03Feb2017-v1/MINIAOD	278820–280385	$7.5\,\text{fb}^{-1}$
/.../Run2016H-03Feb2017_ver{2,3}-v1/MINIAOD	280919–284044	$8.6\,\text{fb}^{-1}$
Total	272007–284044	$35.9\,\text{fb}^{-1}$

Table C.2.: List of all MC simulation samples used in the analysis of 2016 data with the number of generated events and the corresponding cross section times branching ratio. The data set name corresponds to the primary data set name, the fragments RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_-2016_TrancheIV_v6-v* and MINIAODSIM have been omitted for better readability.

Data set name	Events	Cross section × BR (pb)
THQ_ctcvcp_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1	9 776 136	see Section 8.1
THW_ctcvcp_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1	4 983 436	see Section 8.1
ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	67 240 808	136.02 (NLO [37])
ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	38 811 017	80.95 (NLO [37])
ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M2T4	992 024	35.85 (NNLO [37])
ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M2T4	998 276	35.85 (NNLO [37])
TTToSemilepton_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	152 720 952	831.76×0.4393 (NNLO [3, 34])
TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	79 092 400	831.76×0.1061 (NNLO [3, 34])
ttHTobb_M125_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	3 936 004	0.2934 (NLO [343])
TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	3 120 397	0.21 (NLO [344])
TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	833 298	0.435 (NLO [344])
TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	2 160 168	0.2629 (NLO [344])
TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	1 992 438	0.611 (NLO [344])
tZq_W_lept_Z_hadron_4f_ckm_NLO_13TeV_amcatnlo_pythia8	970 479	$0.824 \times 0.3258 \times 0.6991$ (NLO [3, 345])
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	28 968 252	5765 (NNLO [342])

C.2. Reweighting of Final-State Object Kinematics

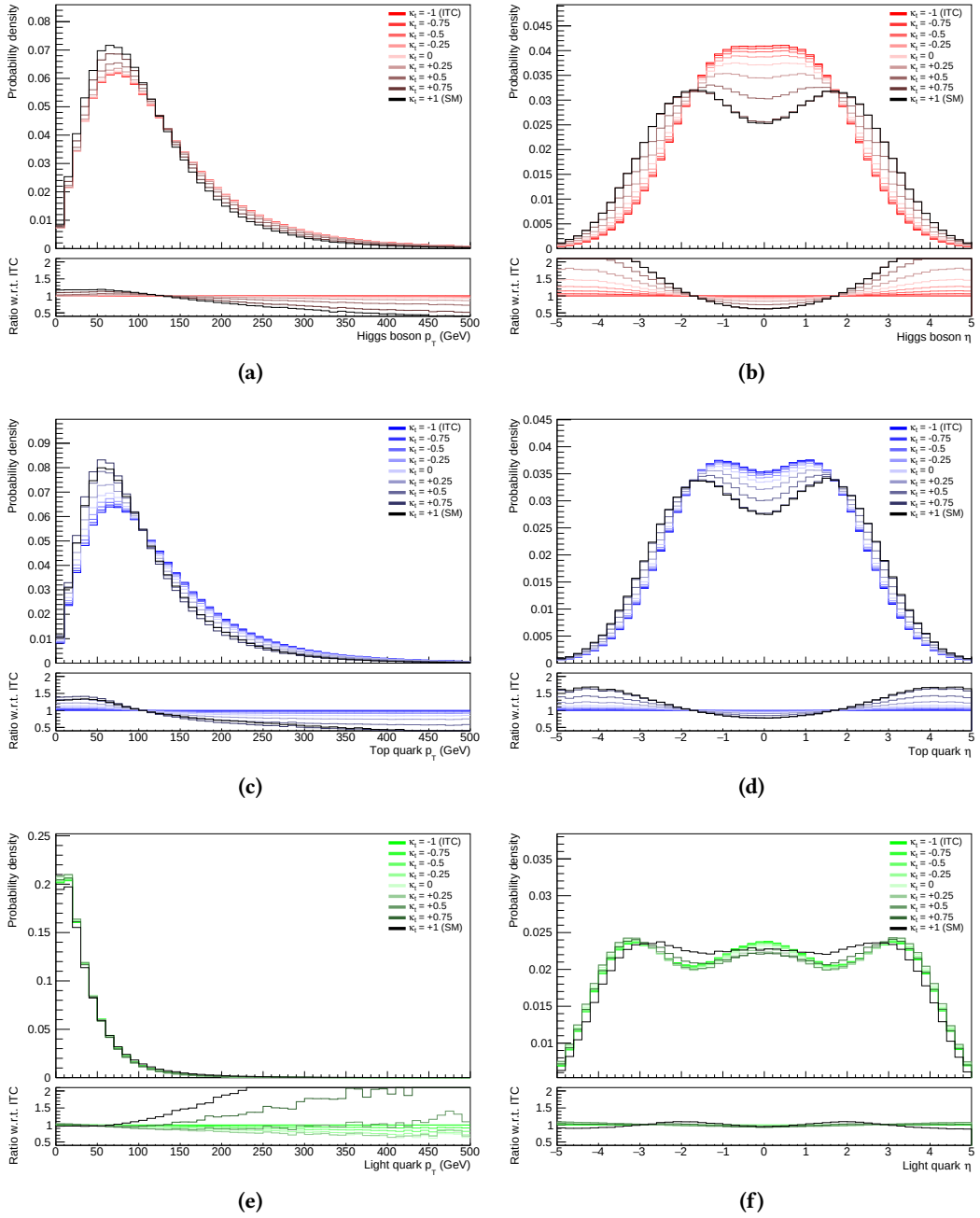


Figure C.1.: Effect of the event reweighting method on kinematic distributions of the final-state objects in the tHq process. The transverse momentum of the Higgs boson (a), the top quark (c) and the light-flavored quark (e) are shown for different values of the coupling modifier to the top-Yukawa coupling. In addition, also the pseudorapidity for each of the three objects is shown ((b), (d) and (f)).

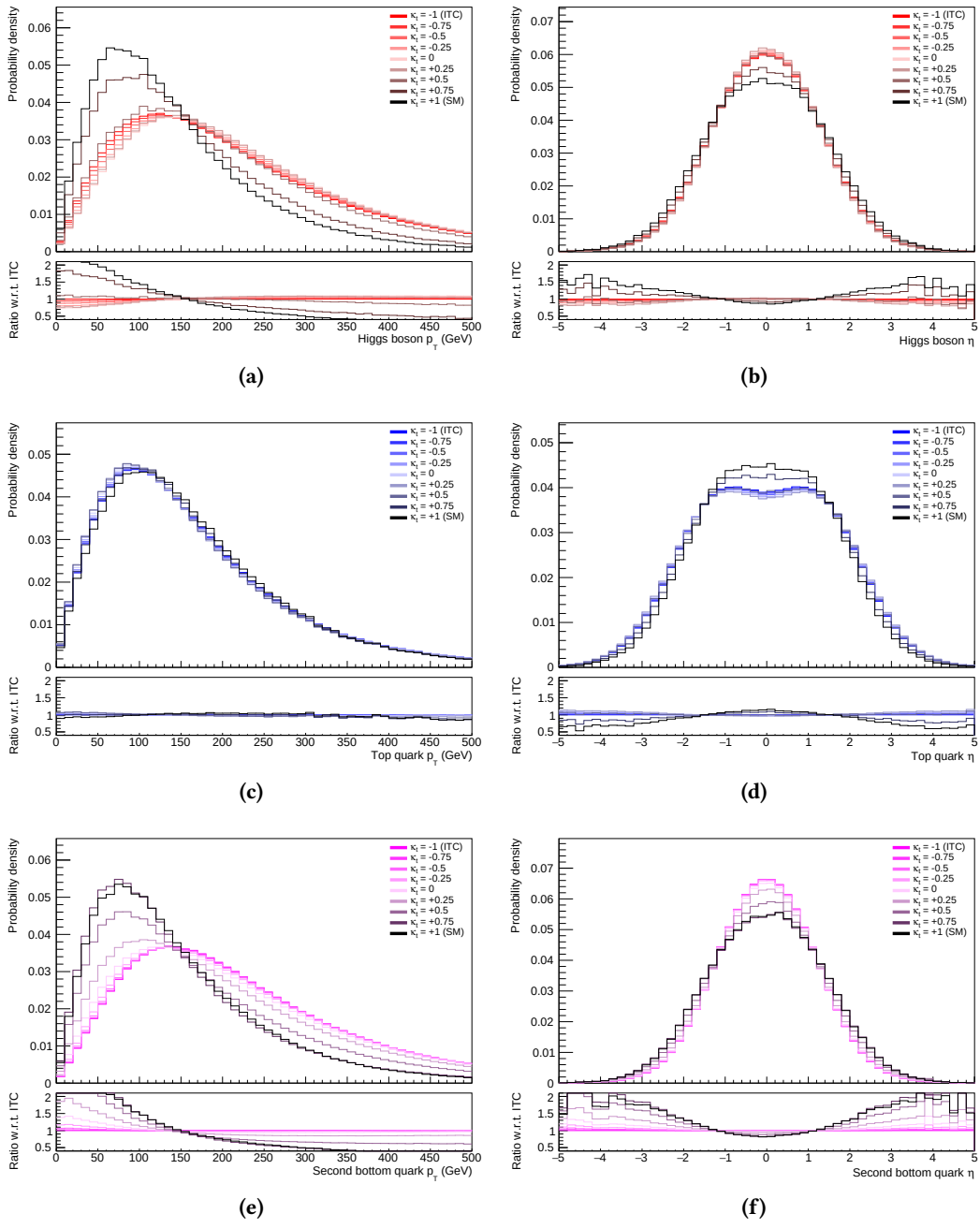


Figure C.2.: Effect of the event reweighting method on kinematic distributions of the final-state objects in the tHW process. The transverse momentum of the Higgs boson (a), the top quark (c) and the bottom quark of the weak top quark production vertex (e) are shown for different values of the coupling modifier to the top-Yukawa coupling. In addition, also the pseudorapidity for each of the three objects is shown ((b), (d) and (f)).

C.3. Two-dimensional Distributions of Final-State Object Kinematics

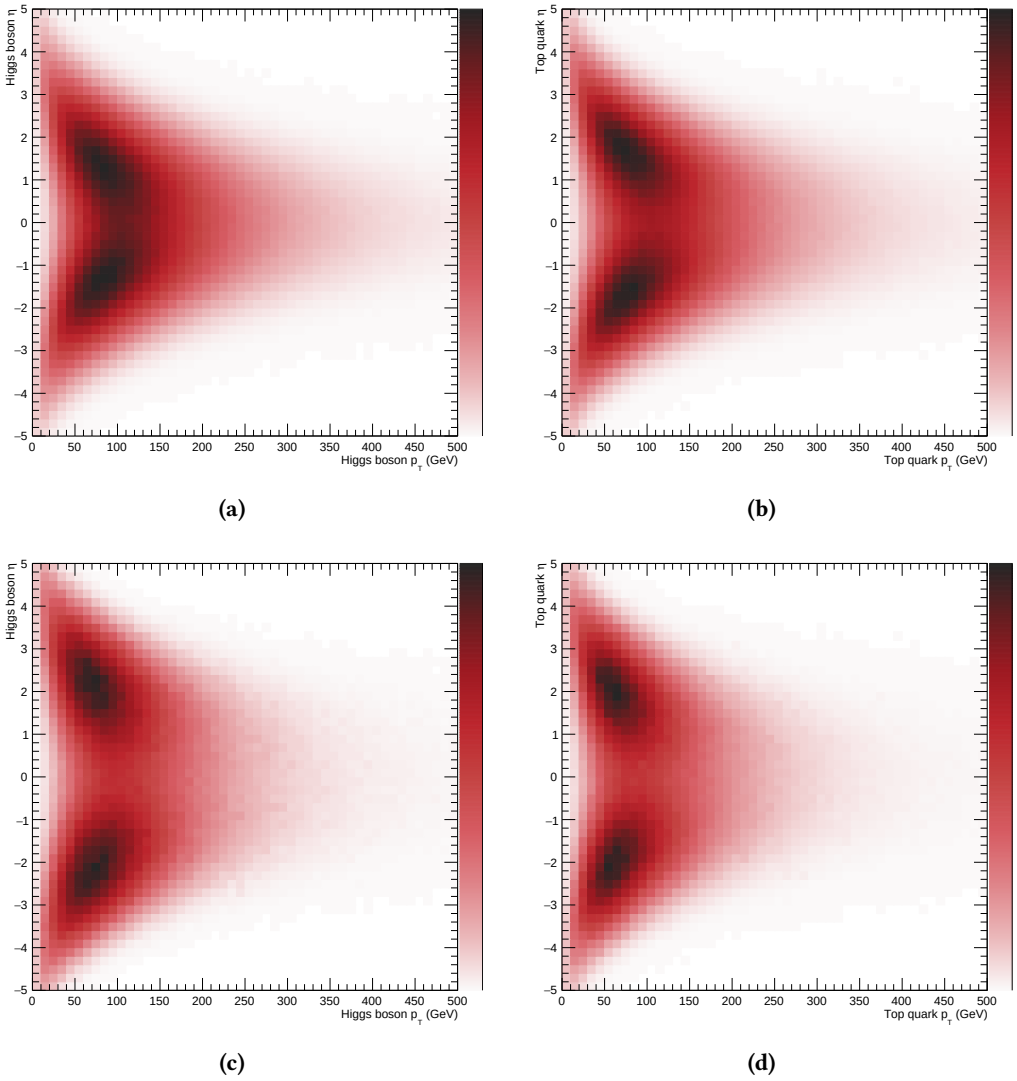


Figure C.3.: Two-dimensional distributions of the transverse momentum and pseudorapidity of the Higgs boson (a) and the top quark (b), both for the ITC scenario, in the simulation of the tHq process. The corresponding distributions for the SM scenario ((c) and (d)) are obtained by reweighting each generated event.

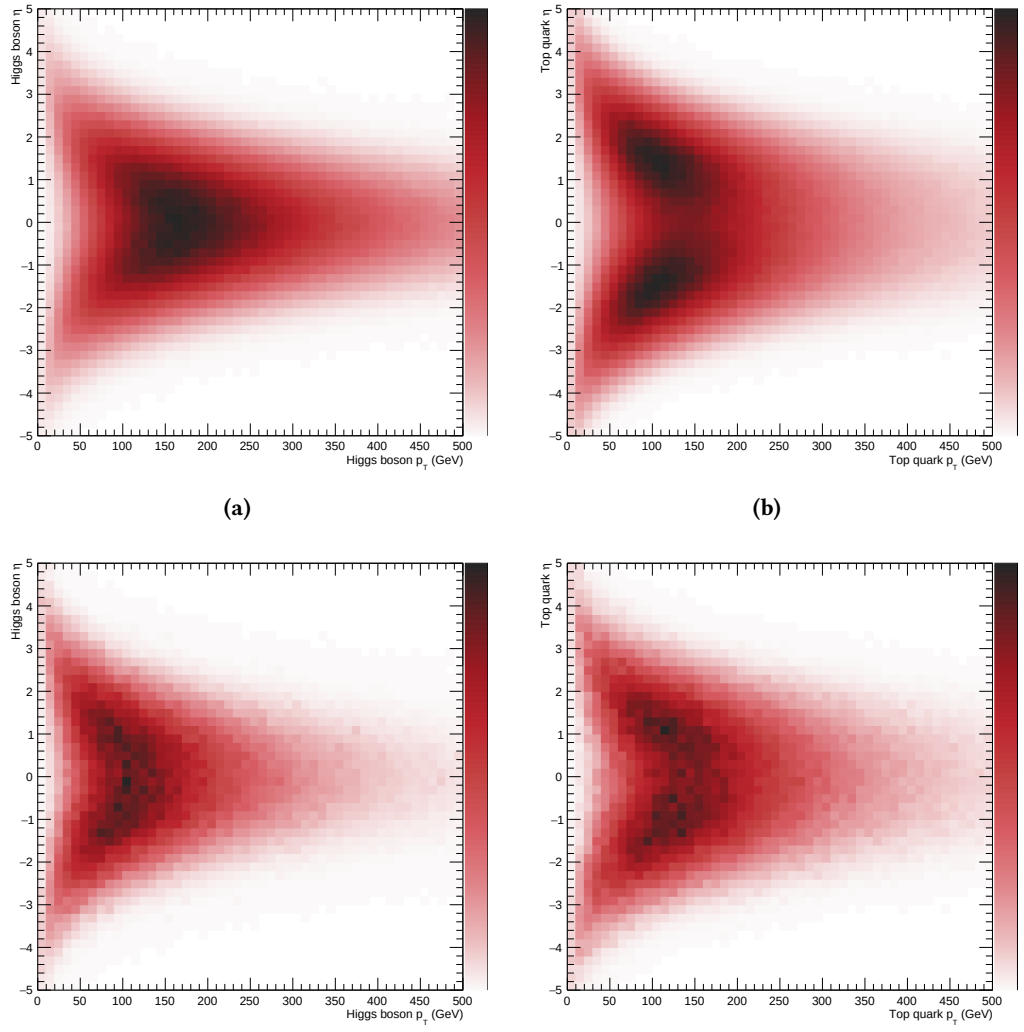


Figure C.4.: Two-dimensional distributions of the transverse momentum and pseudorapidity of the Higgs boson (a) and the top quark (b), both for the ITC scenario, in the simulation of the tHW process. The corresponding distributions for the SM scenario ((c) and (d)) are obtained by reweighting each generated event.

C.4. Settings for the Training of BDTs

Table C.3.: Different parameter settings used in the training of the reconstruction and classification BDTs.

Parameter	Reconstruction BDT	Classification BDT (single lepton)	Classification BDT (dilepton)
`NTrees`	400	400	150
`MinNodeSize`	1	1	2
`MaxDepth`	3	3	2
`BoosteType`	AdaBoost	AdaBoost	AdaBoost
`nCuts`	20	20	16
`AdaBoostBeta`	0.3	0.3	0.3
`SeparationType`	GiniIndex	GiniIndex	GiniIndex

C.5. Input Variables for the tHq Hypothesis

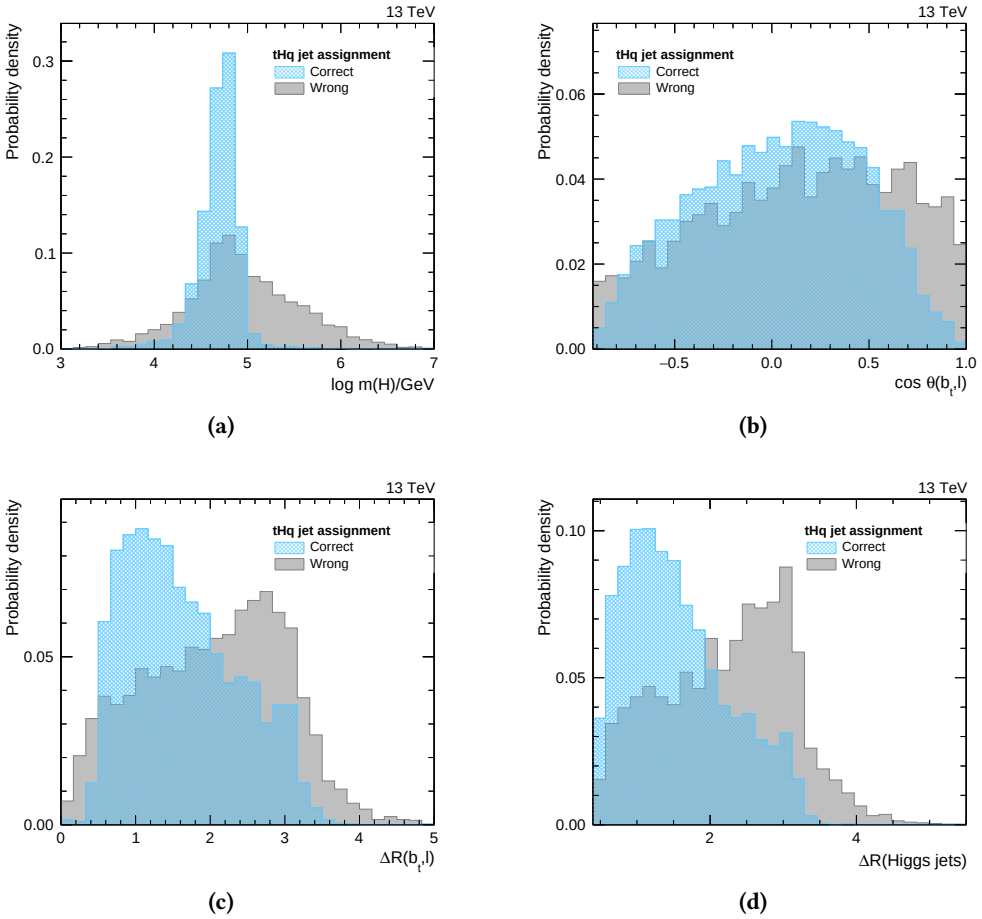


Figure C.5.: Distributions of input variables for the tHq event reconstruction, ranked 1^{st} to 4^{th} in the training of the BDT: the invariant mass of the reconstructed Higgs boson (a), the cosine of the angle between the jet assigned to the bottom quark of the top quark decay and the lepton (b), ΔR between the jet assigned to the bottom quark from the top quark decay and the lepton (c) and ΔR between the two jets assigned to the Higgs boson (d).

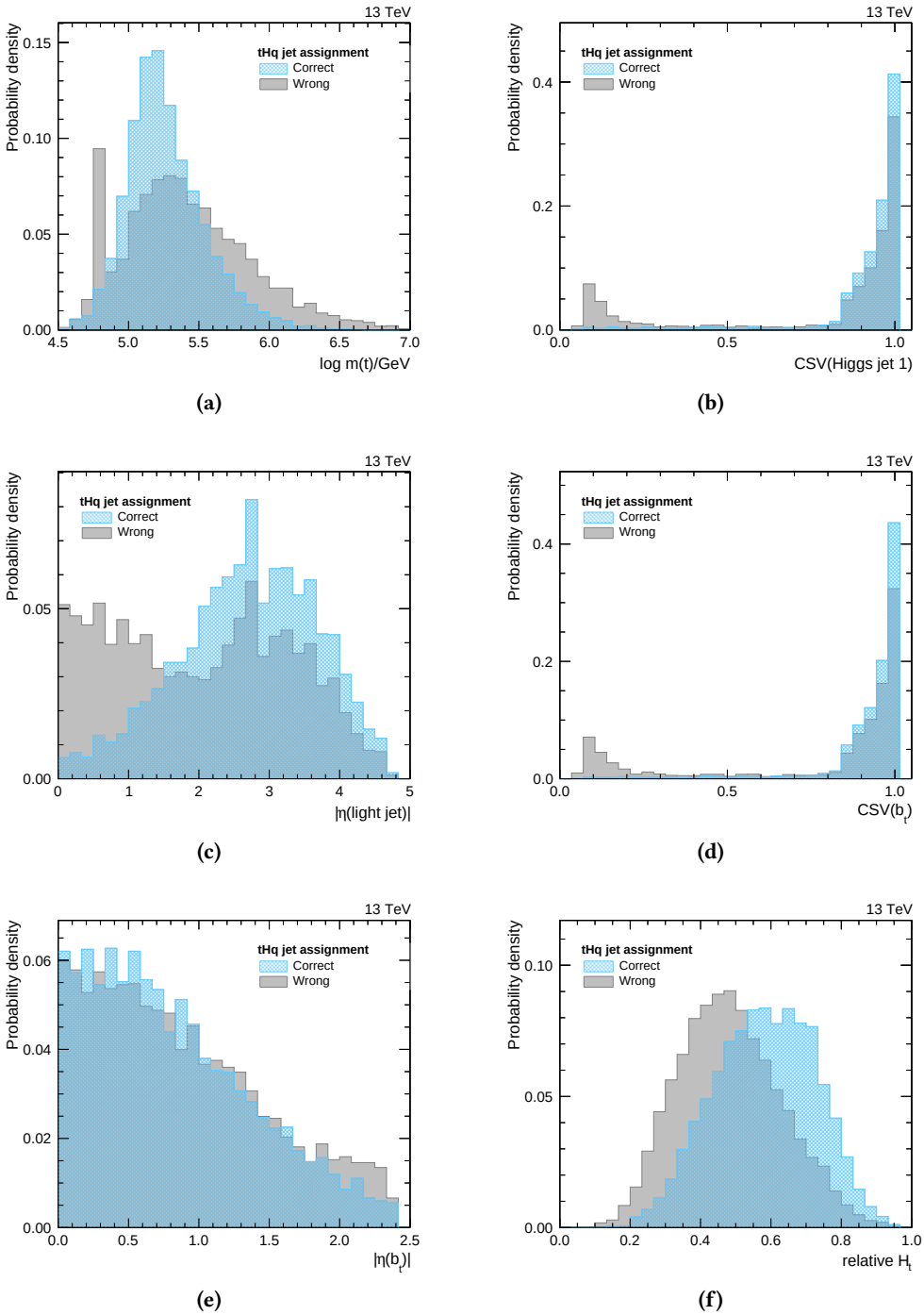


Figure C.6.: Distributions of input variables for the tHq event reconstruction, ranked 5^{th} to 10^{th} in the training of the BDT: the invariant mass of the reconstructed top quark (a), the output of the b tagging discriminant for the jet assigned to the Higgs boson with the highest transverse momentum (b), the absolute pseudorapidity of the light forward jet (c), the output of the b tagging discriminant for the jet assigned to the bottom quark from the top quark decay (d), the absolute pseudorapidity of the jet assigned to the bottom quark from the top quark decay (e) and the percentage of the total transverse momenta that falls to the b jet from the top quark, Higgs jets and the light forward jet (f).

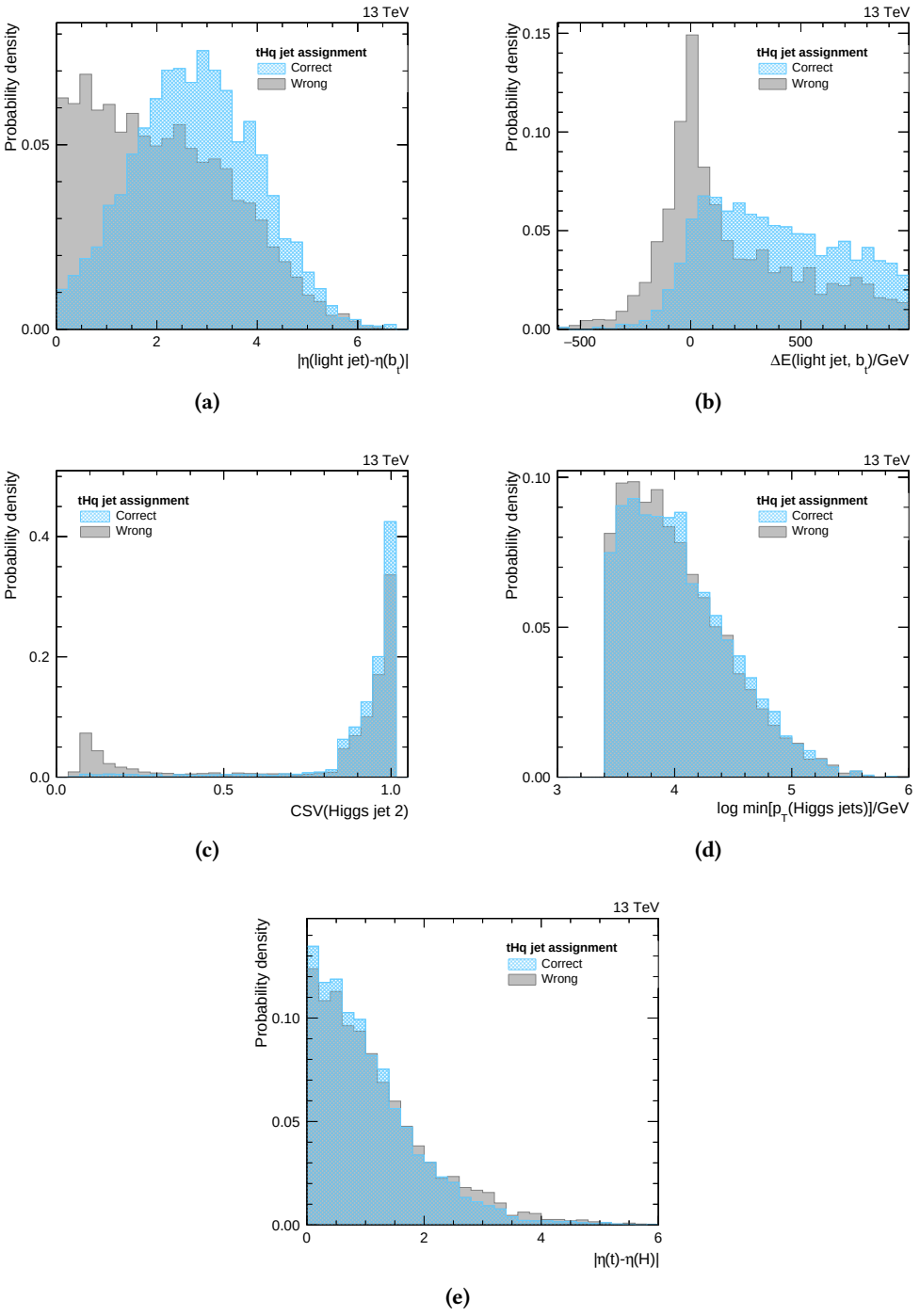


Figure C.7.: Distributions of input variables for the tHq event reconstruction, ranked 11^{th} to 15^{th} in the training of the BDT: the absolute difference of pseudorapidities of the light forward jet and the b jet fromt he top quark decay (a), the jet energy difference of the light forward jet and the jet assigned to the bottom quark from the top quark decay (b), the output of the b tagging discriminant for the jet assigned to the Higgs boson with the second highest transverse momentum (c), the lower transverse momentum of the two jets assigned to the Higgs boson (d) and the absolute difference in pseudorapidities of the reconstructed top quark and the reconstructed Higgs boson (e).

C.6. Input Variables for the tHW Hypothesis

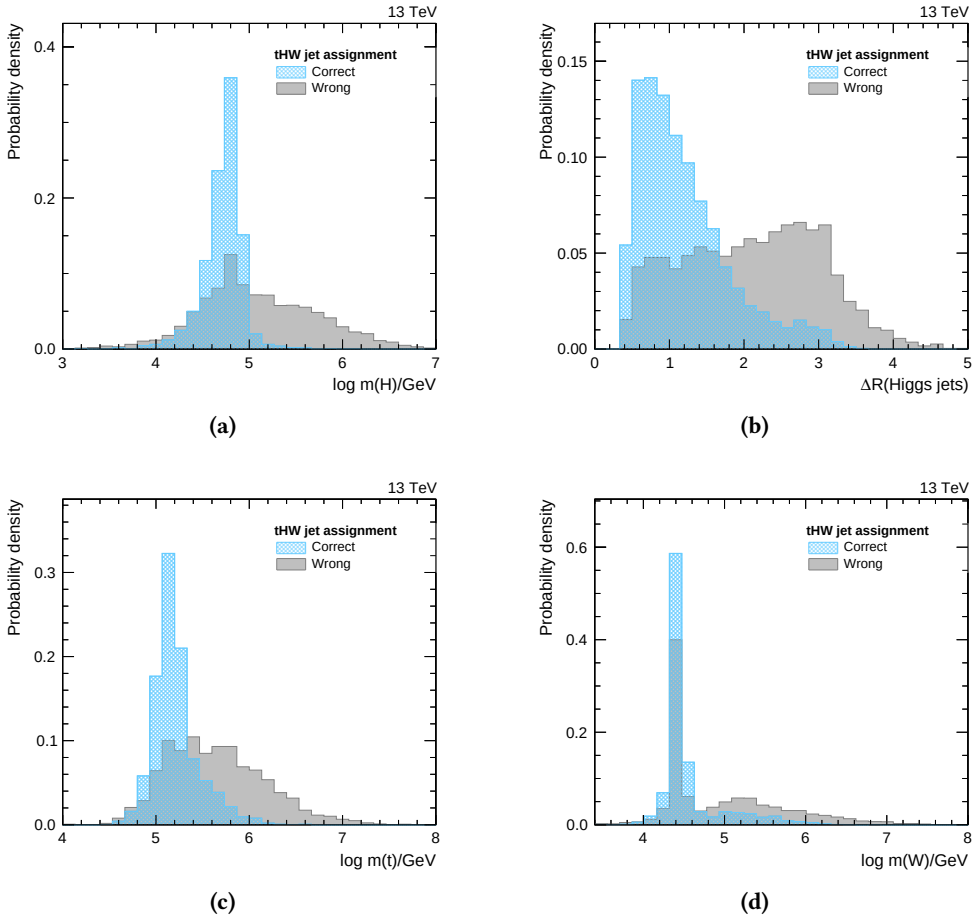


Figure C.8.: Distributions of input variables for the tHW event reconstruction, ranked 1^{st} to 4^{th} in the training of the BDT: the invariant mass of the reconstructed Higgs boson (a), ΔR between the two jets from the Higgs boson decay (b), the invariant mass of the reconstructed top quark (c) and the invariant mass of the W boson from the top quark production vertex (d).

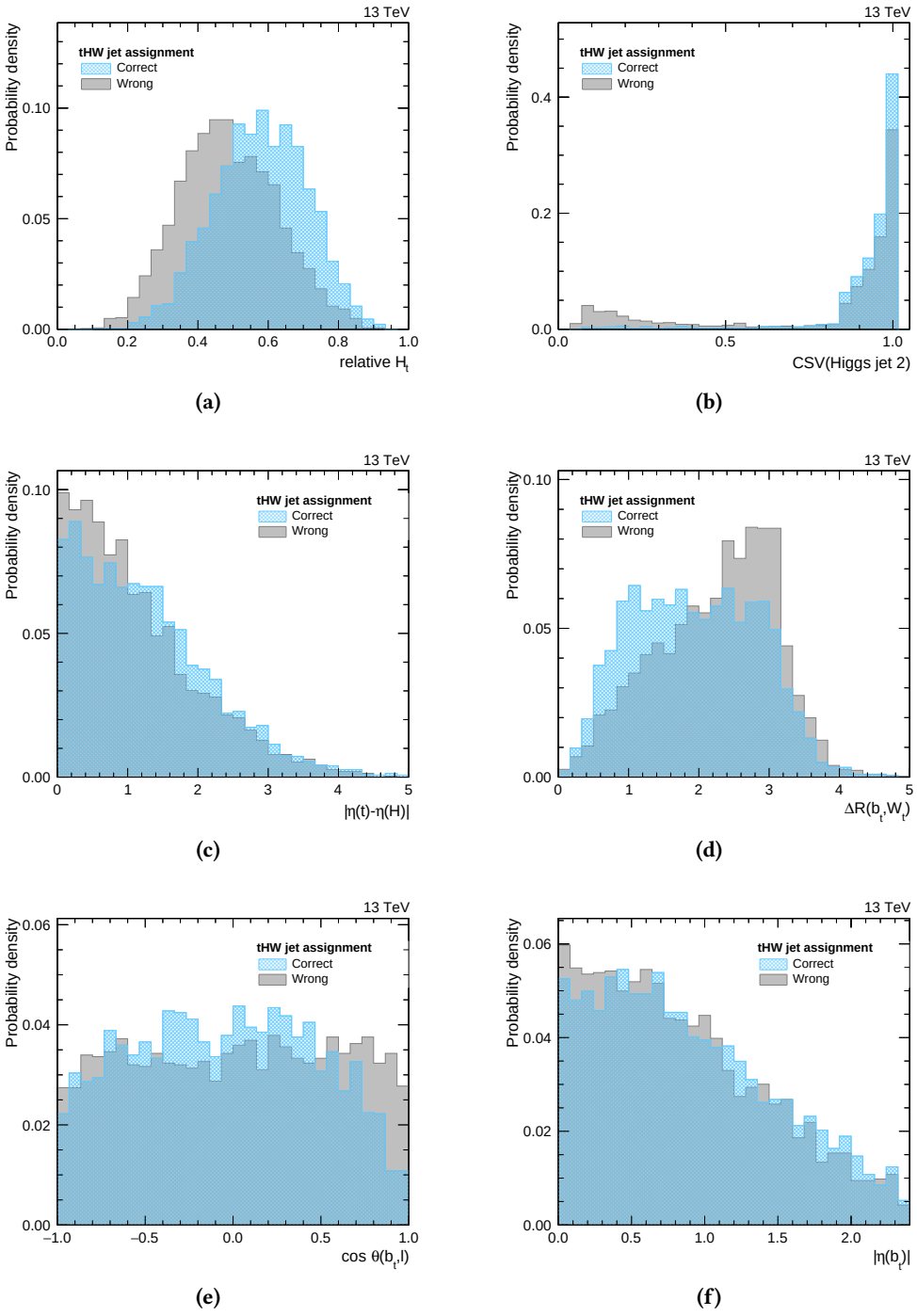


Figure C.9.: Distributions of input variables for the tHW event reconstruction, ranked 5^{th} to 10^{th} in the training of the BDT: the percentage of the total transverse momentum that falls to the b jet from the top quark decay, Higgs jets and the light forward jet (a), the output of the b tagging discriminant for the jet assigned to the Higgs boson with the second highest transverse momentum (b), the absolute difference of pseudorapidities of the reconstructed top quark and the reconstructed Higgs boson (c), ΔR between the jet assigned to the bottom quark from the top quark decay and the leptonically decaying W boson (d), cosine of the angle between the jet assigned to the bottom quark from the top quark decay and the lepton (e) and the absolute pseudorapidity of the jet assigned to the bottom quark of the top quark decay (f).

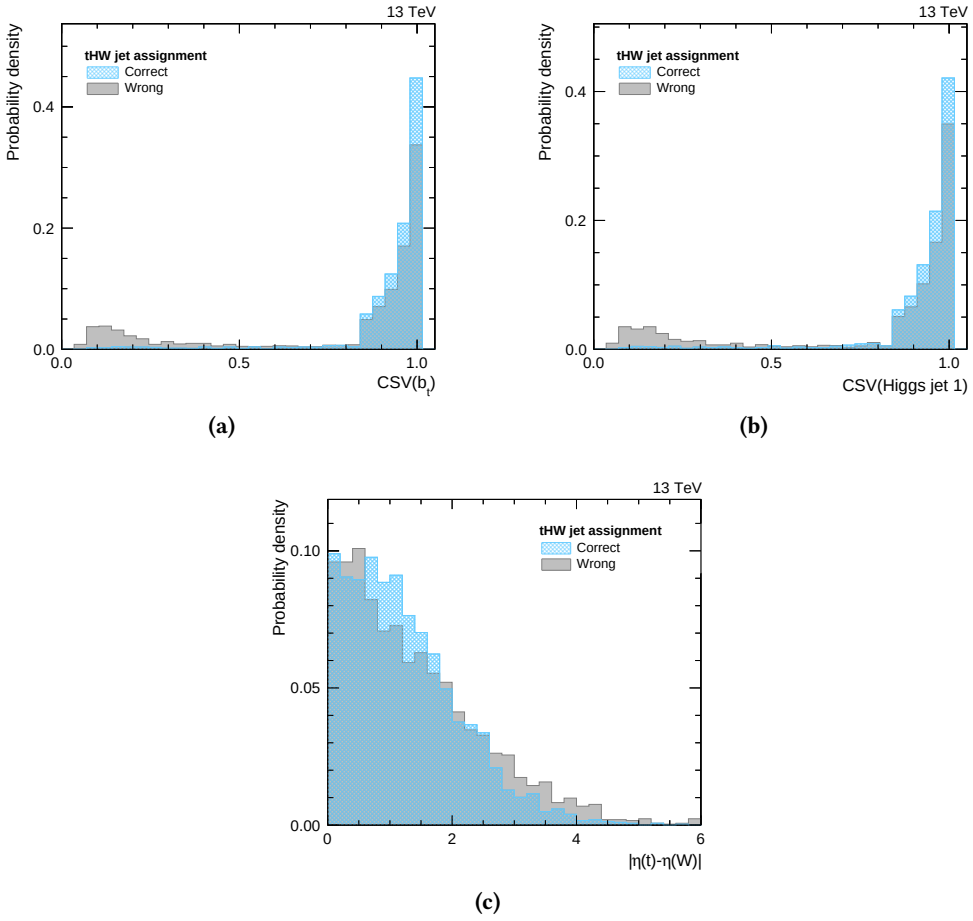


Figure C.10.: Distributions of input variables for the tHW event reconstruction, ranked 11^{th} to 13^{th} in the training of the BDT: the output of the b tagging discriminant for the jet assigned to the bottom quark from the top quark decay (a), the output of the b tagging discriminant for the jet assigned to the Higgs boson with the highest transverse momentum (b) and the absolute difference of pseudorapidities of the reconstructed top quark and the reconstructed W boson from the top quark production vertex (c).

C.7. Input Variables for the $t\bar{t}$ Hypothesis

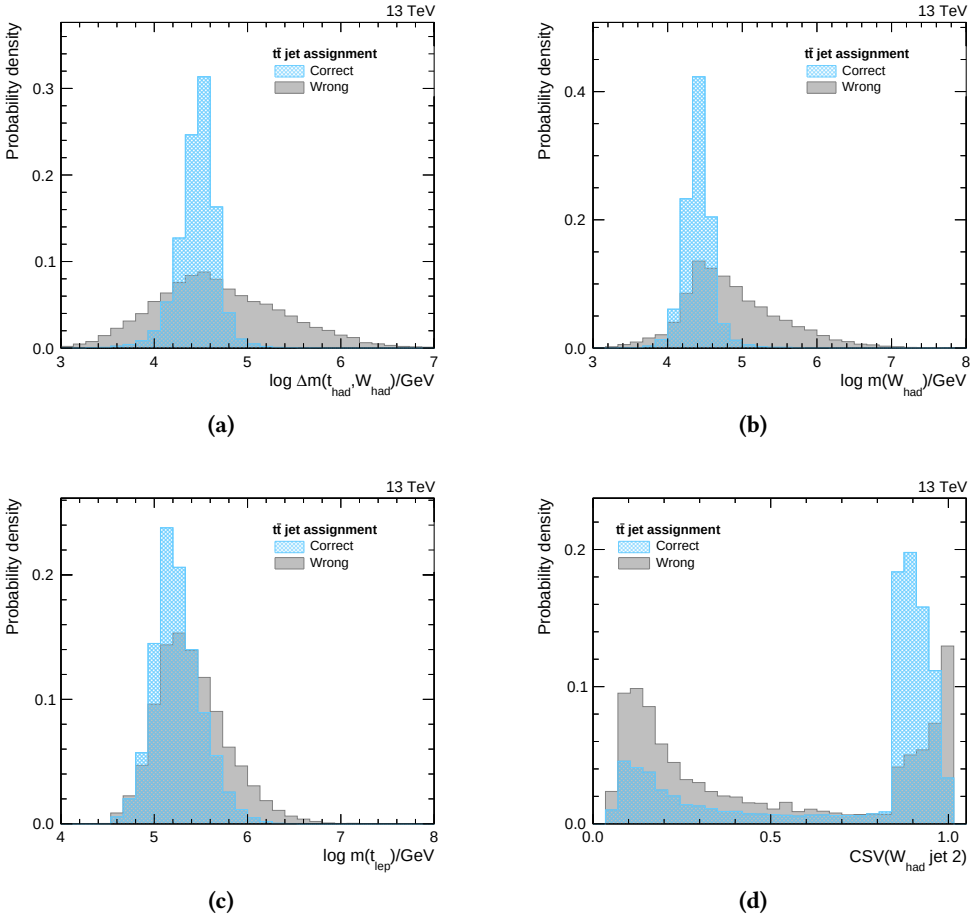


Figure C.11.: Distributions of input variables for the $t\bar{t}$ event reconstruction, ranked 1^{st} to 4^{th} in the training of the BDT: the difference between the invariant masses of the reconstructed hadronically decaying top quark and the hadronically decaying W boson (a), the invariant mass of the two jets assigned to the reconstructed W boson from the hadronically decaying top quark (b), the invariant mass of the reconstructed leptonically decaying top quark (c) and the output of the b tagging discriminant for the jet with the highest transverse momentum assigned to the hadronically decaying W boson (d).

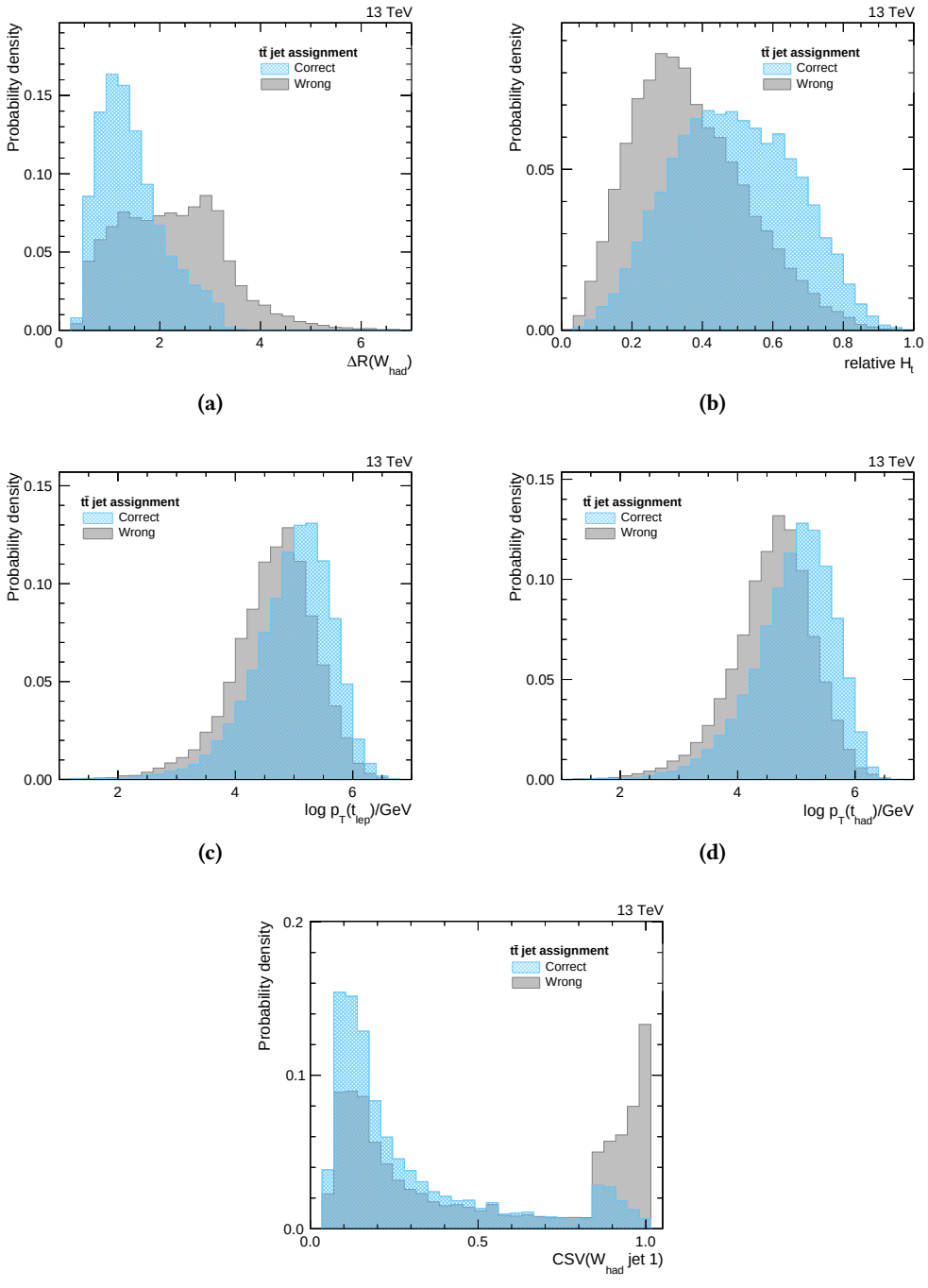


Figure C.12.: Distributions of input variables for the $t\bar{t}$ event reconstruction, ranked 5^{th} to 9^{th} in the training of the BDT: ΔR between the jet assigned to the bottom quark from the hadronically decaying top quark and the reconstructed W boson from the hadronically decaying top quark (a), the percentage of total transverse momentum that falls to the reconstructed hadronically and leptonically decaying top quarks (b), the transverse momentum of the reconstructed leptonically decaying top quark (c), the transverse momentum of the reconstructed hadronically decaying top quark (d) and the output of the b tagging discriminant for the jet with the highest transverse momentum assigned to the hadronically decaying W boson (e).

C.8. Input Variables for the Event Classification in the Single-Lepton Region

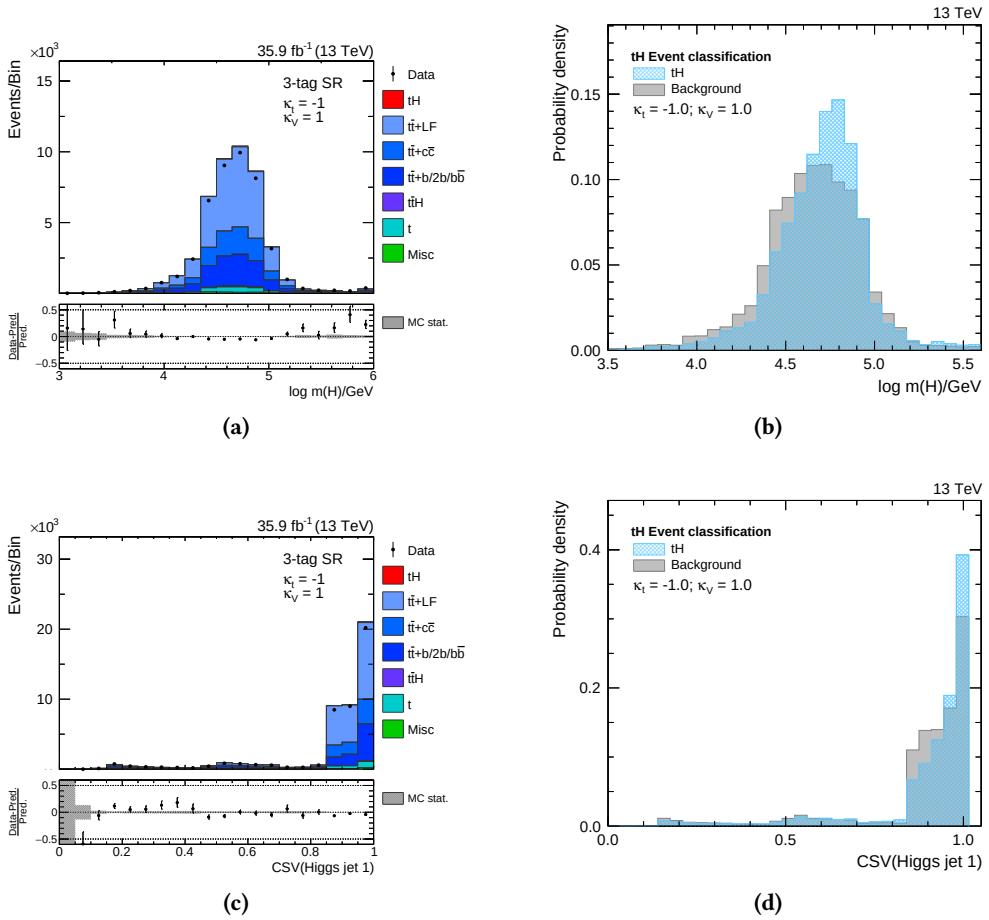


Figure C.13.: Distributions of input variables for the single-lepton event classification, ranked 4^{th} and 5^{th} in the training of the BDT: the invariant mass of the reconstructed Higgs boson candidate ((a) and (b)) and the output of the b tagging discriminant for the jet assigned to the Higgs boson with the highest transverse momentum ((c) and (d)). The distributions on the left are scaled to the expected number of events.

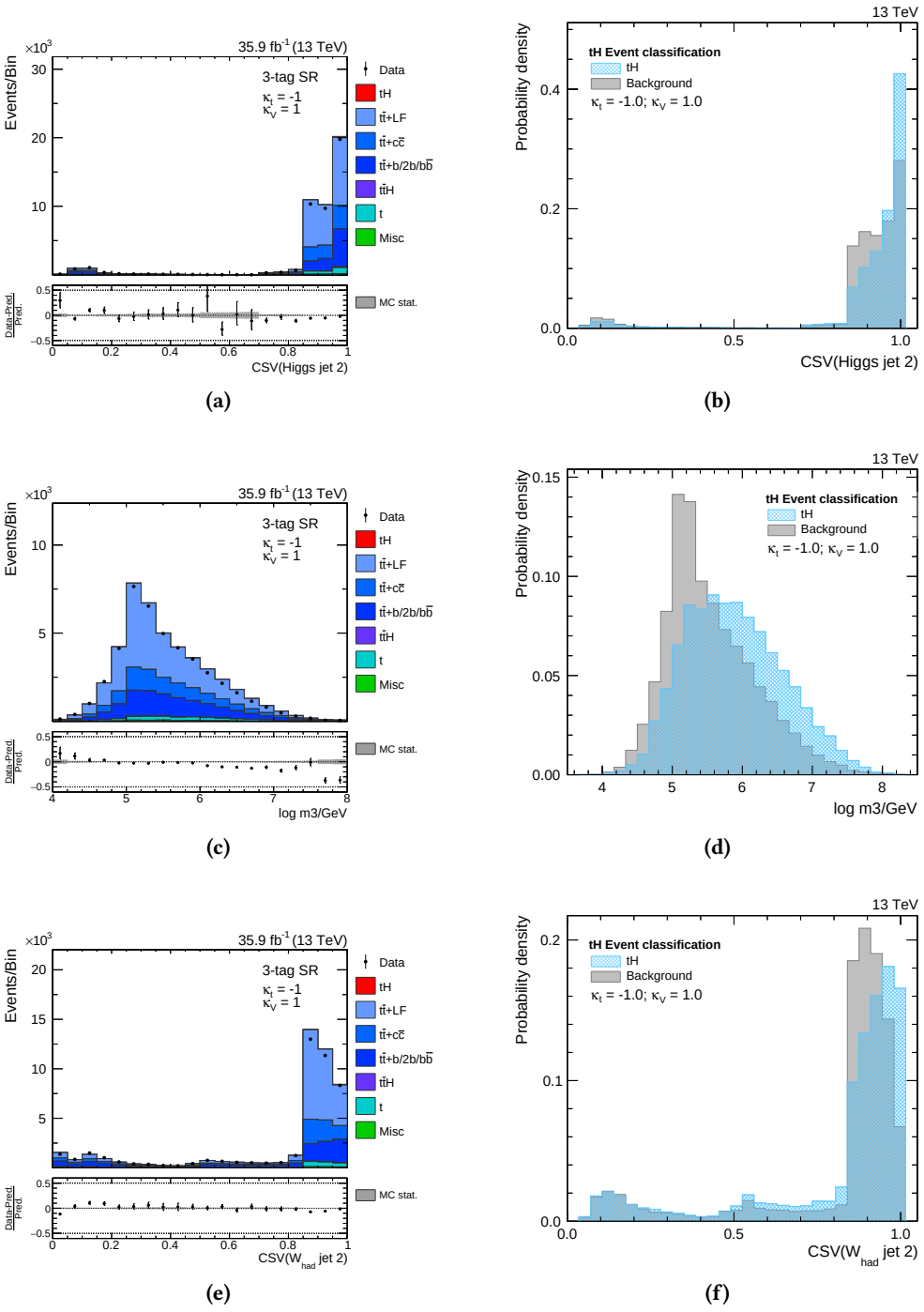


Figure C.14.: Distributions of input variables for the single-lepton event classification, ranked 6^{th} to 8^{th} in the training of the BDT: the output of the b tagging discriminant for the jet assigned to the Higgs boson candidate with the second highest transverse momentum ((a) and (b)), the invariant mass of the three jets that result in the highest transverse momentum when combined ((c) and (d)) and the output of the b tagging discriminant for the jet assigned to the hadronically decaying W boson with the second highest transverse momentum ((e) and (f)). The distributions on the left are scaled to the expected number of events.

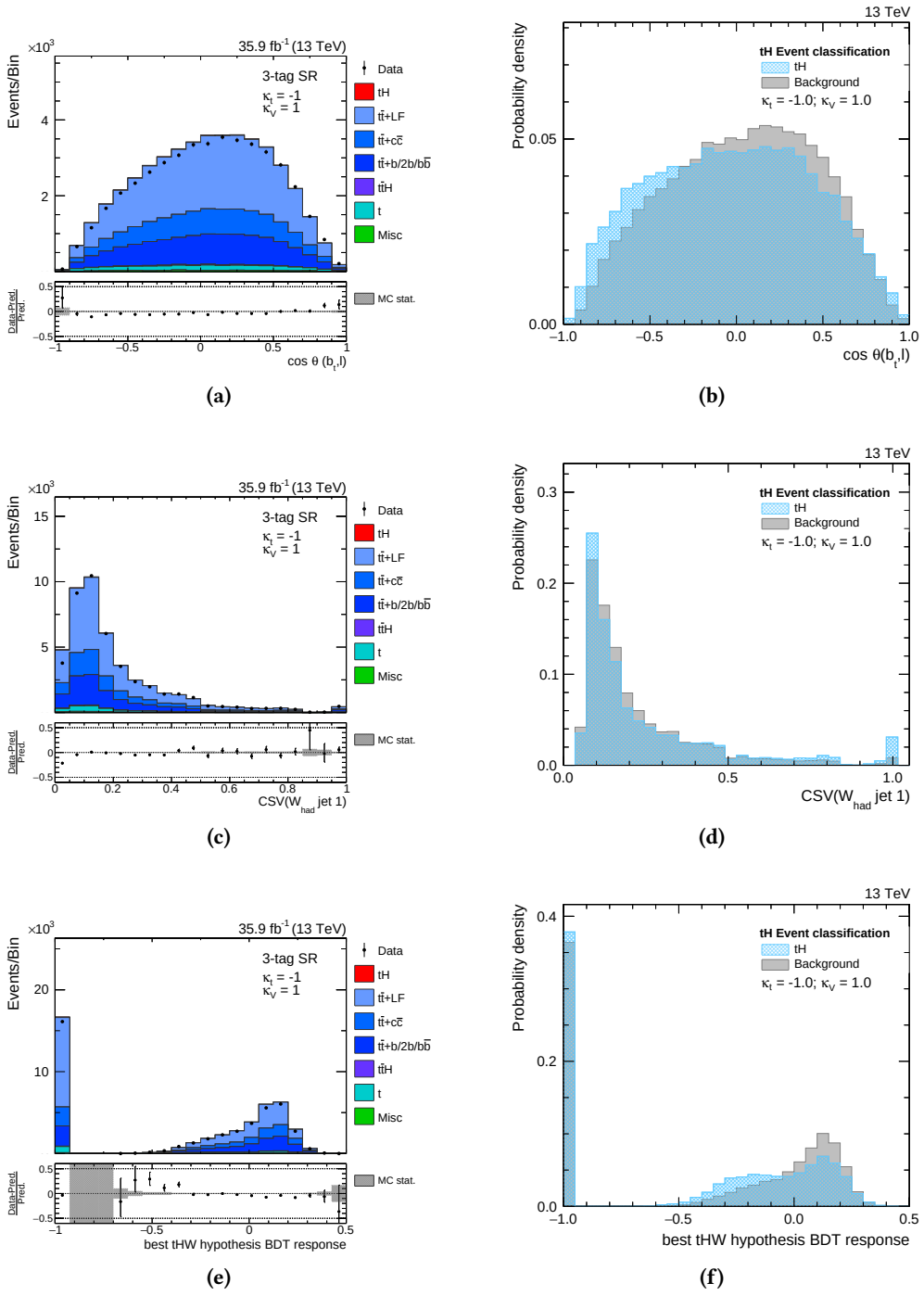
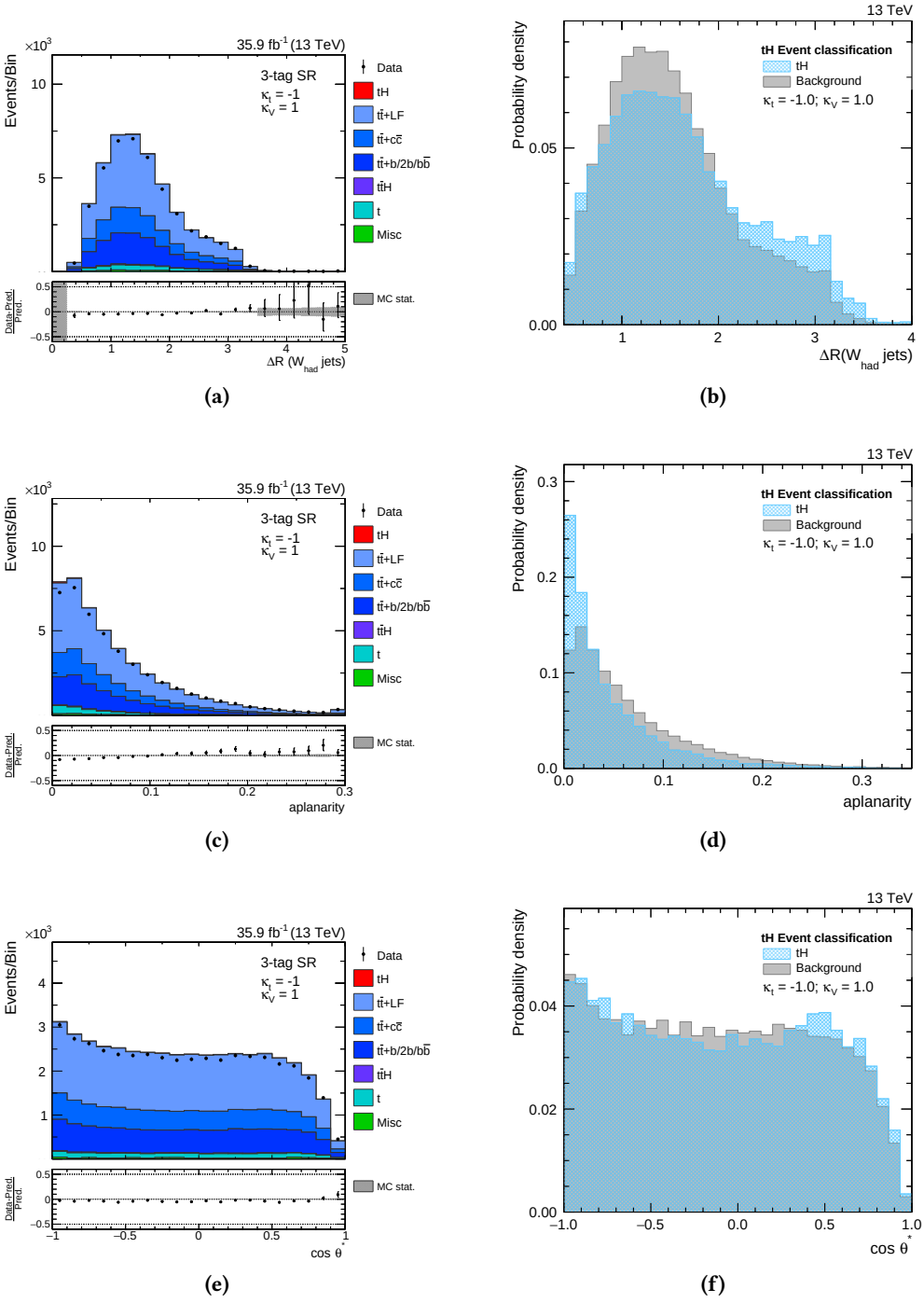


Figure C.15.: Distributions of input variables for the single-lepton event classification, ranked 9$^{\text{th}}$ to 11$^{\text{th}}$ in the training of the BDT: the cosine of the angle between the b-tagged jet from the top quark decay and the lepton ((a) and (b)), the output of the b tagging discriminant for the jet assigned to the hadronically decaying W boson with the highest transverse momentum ((c) and (d)) and the output of the tHW reconstruction BDT ((e) and (f)). The distributions on the left are scaled to the expected number of events.



(a) **(b)** **(c)** **(d)** **(e)** **(f)**

Figure C.16.: Distributions of input variables for the single-lepton event classification, ranked 12^{th} to 14^{th} in the training of the BDT: ΔR between the two light jets assigned to the hadronically decaying W boson ((a) and (b)), the aplanarity of the events ((c) and (d)) and the cosine of the angle between the light forward jet and the lepton in the top quark rest frame ((e) and (f)). The distributions on the left are scaled to the expected number of events.

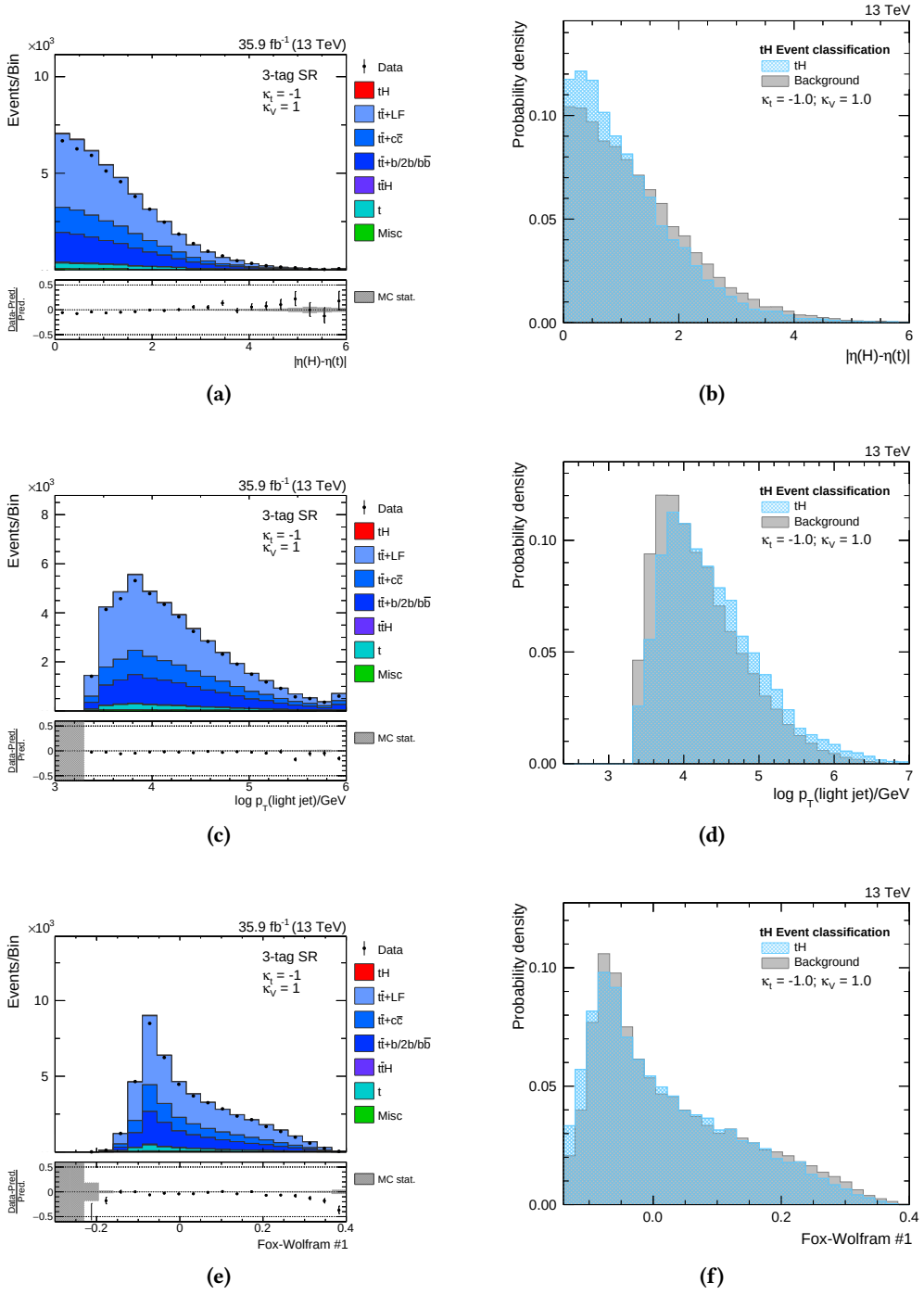
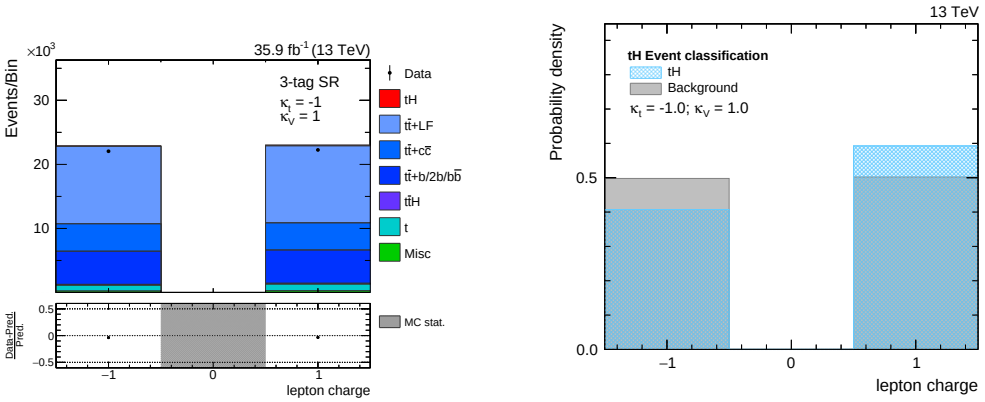


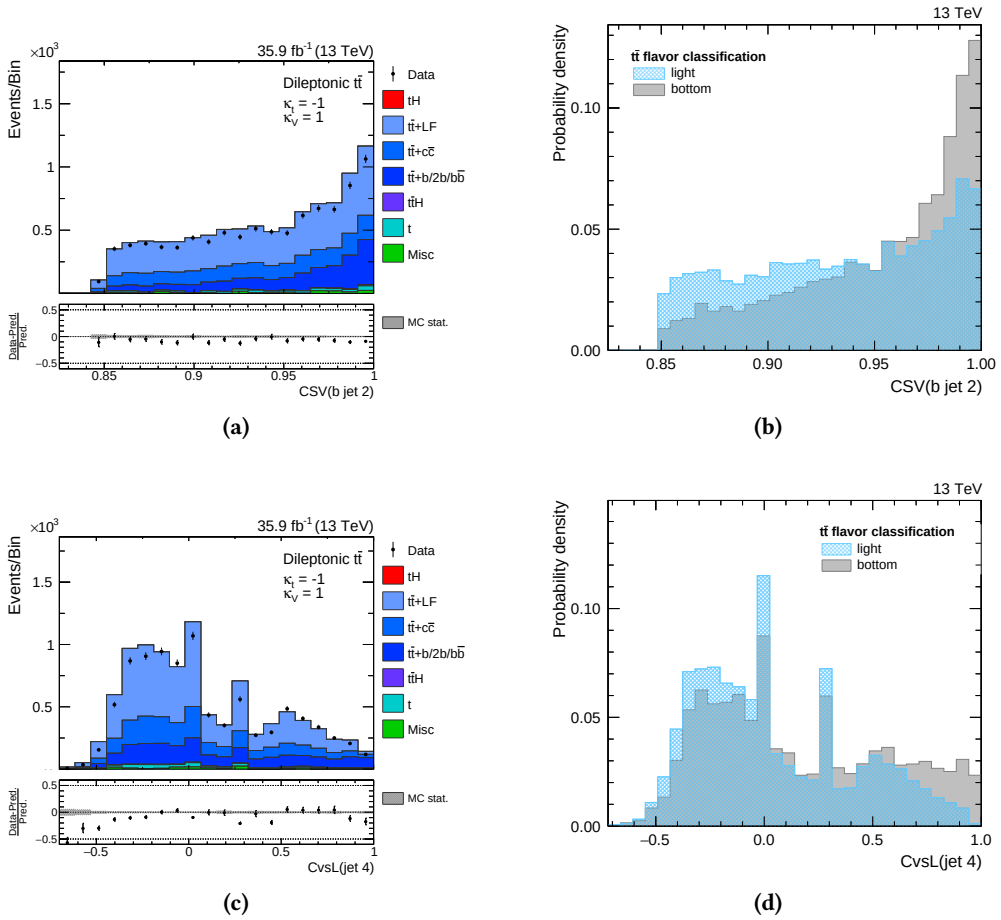
Figure C.17.: Distributions of input variables for the single-lepton event classification, ranked 15^{th} to 17^{th} in the training of the BDT: the absolute pseudorapidity difference of the reconstructed Higgs boson and the reconstructed top quark ((a) and (b)), the transverse momentum of the light forward jet ((c) and (d)) and the first Fox-Wolfram moment of the event ((e) and (f)). The distributions on the left are scaled to the expected number of events.



(a) **(b)**

Figure C.18.: Distributions of the input variable for the single-lepton event classification, ranked 18$^{\text{th}}$ in the training of the BDT: the charge of the lepton ((a) and (b)). The distribution on the left is scaled to the expected number of events.

C.9. Input Variables for the Event Classification in the Dileptonic Region



(a) (b) (c) (d)

Figure C.19.: Distributions of input variables for the dileptonic event classification, ranked 4^{th} and 5^{th} in the training of the BDT: the output of the b tagging discriminant for the b-tagged jet with the second highest b tagging value in the event ((a) and (b)) and the output of the charm vs. light flavor tagging algorithm for the jet with the fourth highest momentum in the event ((c) and (d)). The distributions on the left are scaled to the expected number of events.

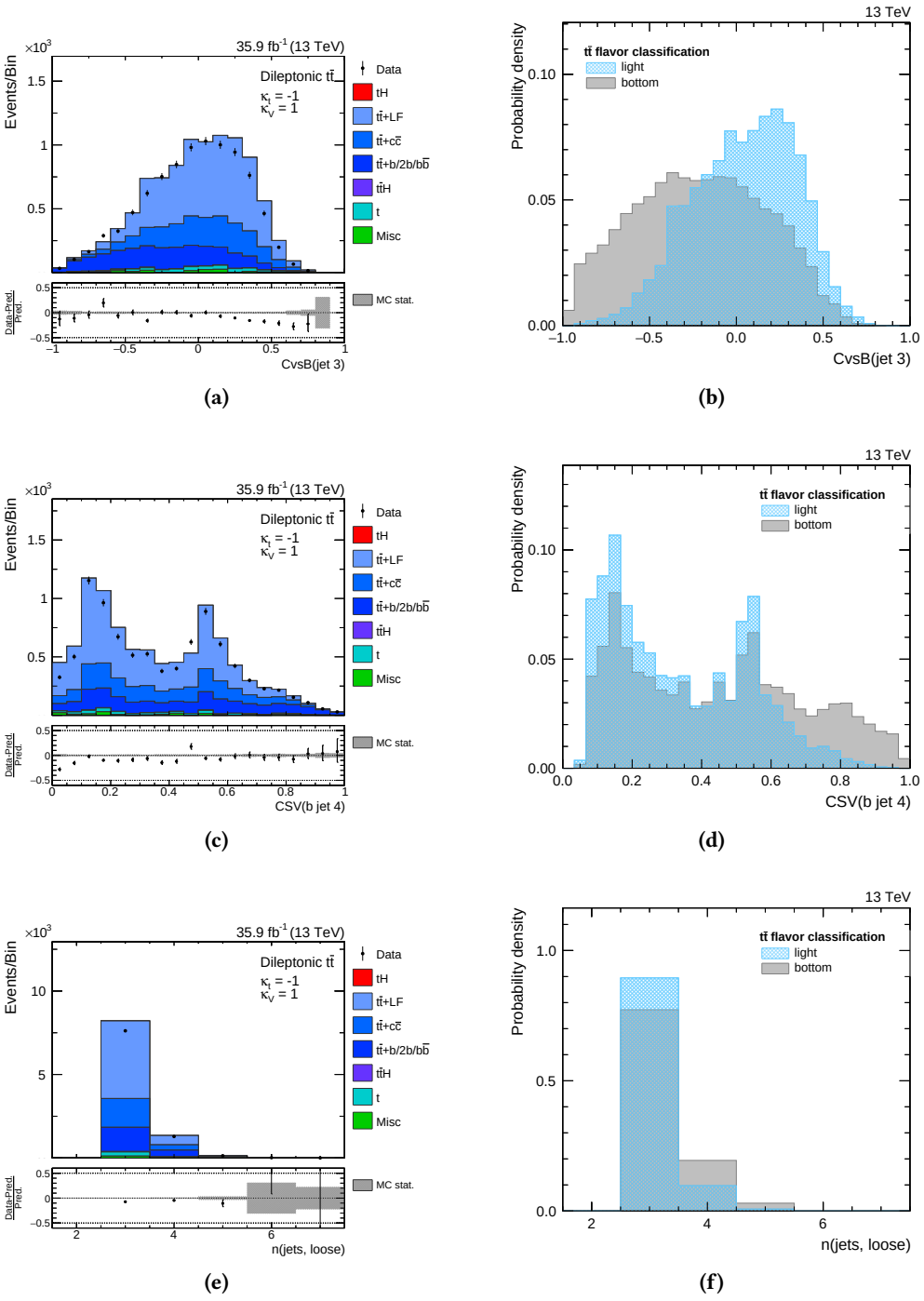


Figure C.20.: Distributions of input variables for the dileptonic event classification, ranked 6^{th} to 8^{th} in the training of the BDT: the output of the charm vs. bottom flavor tagging algorithm for the jet with the third highest transverse momentum in the event ((a) and (b)), the output of the b tagging discriminant for the b-tagged jet with the fourth highest b tagging value in the event ((c) and (c)ns) and the number of jets in the event passing the loose working point of the b tagging algorithm ((e) and (f)). The distributions on the left are scaled to the expected number of events.

List of Figures

1.1 The Higgs potential . . . 7
1.2 Electroweak interactions in the standard model . . . 7
1.3 Strong interactions in the standard model . . . 9
1.4 Transition from protons to stable collision products . . . 11
1.5 Feynman diagrams of top quark pair production at the LHC . . . 12
1.6 Experimental status of top quark-antiquark production at hadron-hadron colliders 13
1.7 Feynman diagrams of single top quark production at the LHC . . . 13
1.8 Experimental status of top quark-antiquark production at the LHC . . . 14
1.9 Predicted cross sections for various Higgs boson production modes at different center-of-mass energies . . . 16
1.10 Feynman diagrams of different Higgs boson production channels at the LHC . 17
1.11 Predicted Higgs boson branching ratios in dependence of the Higgs boson mass 18
1.12 Experimental status of the coupling modifiers of the Higgs boson to fermions and vector bosons. . . . 19
1.13 Feynman diagrams contributing to the associated procudion of single top quarks and a Higgs boson . . . 20
1.14 Predicted production cross section for a CP-violating Higgs boson dependent on the CP-mixture . . . 22

2.1 Structure of a decision tree . . . 26
2.2 Structure of a neural network . . . 28
2.3 Structure of a deep neural network . . . 29
2.4 Illustration of overtraining effect . . . 30
2.5 Illustration of the ROC curve . . . 31
2.6 Determination of p-values . . . 34
2.7 Conversion of p-values to standard deviations . . . 35

3.1 CERN accelerator complex . . . 38
3.2 Integrated luminosity recorded by the CMS experiment. . . . 40
3.3 The Compact Muon Solenoid . . . 41
3.4 Schematic overview of the tracker system . . . 42
3.5 The electromagnetic calorimeter of the CMS detector . . . 44
3.6 The hadron calorimeter of the CMS detector . . . 46
3.7 Illustration of the CMS magnet . . . 47
3.8 The muon system of the CMS experiment . . . 48

3.9 Illustration of the trigger and data acquisition system of the CMS experiment . 49
3.10 Illustration of the computing infrastructure of the CMS experiment 50

4.1 Simulation steps of proton-proton collisions 52
4.2 NNPDF parton distribution functions 53
4.3 Feynman diagrams for the single top quark production in the t channel for the four-flavor and five-flavor scheme 54
4.4 Illustration of the particle-flow reconstruction performance 58
4.5 Illustration of collinear and infrared safety violation 61
4.6 Example of different jet clustering algorithms 63
4.7 Illustration of the jet energy corrections in the CMS experiment 64
4.8 Principle of a secondary vertex 65
4.9 Distribution of b tagging algorithms 66
4.10 Comparison of different b tagging algorithms 67

5.1 Efficiency of muon identification and isolation 71
5.2 Jet identification variables 76

6.1 Integrated and peak luminosity of 2015 82
6.2 Feynman diagrams for single top quark and single top antiquark production in the t channel 83
6.3 Relevant signal and background processes for the measurement of the t-channel single top quark production cross section. 85
6.4 Corrections to the number of pileup interactions 90
6.5 Corrections for muon efficiencies 91
6.6 Corrections for b tagging efficiencies 92
6.7 QCD multijet background estimation principle 93
6.8 Pseudorapidity distribution of the light-flavored jet 94
6.9 QCD multijet background estimation in the early measurement 96
6.10 Distributions of the absolute pseudorapidity of the light-flavored jet in control regions 96
6.11 Result of the fit to the absolute pseudorapidity of the light-flavored jet 97
6.12 Reconstructed top quark mass in the early measurement 97
6.13 Comparison of the measured cross section from the early measurement with the prediction 98
6.14 Event display of a t channel single top quark candidate 99
6.15 QCD multijet background estimation in the 2-jets–0-tag control region 101
6.16 QCD multijet background estimation in the 2-jets–1-tag signal region 102
6.17 Distributions of the three most important input variables in the 2-jets–1-tag signal region 105
6.18 Neural network distribution obtained from training 106
6.19 Correlation of input variables 106
6.20 Illustration of the neural network 107
6.21 Check for overtraining of the neural network 107

6.22 Distributions of the neural network output in the 2-jets–1-tag signal region . . 110
6.23 Distributions of the neural network output in the 3-jets–1-tag and 3-jets–2-tags control region 111
6.24 Comparison of the measured ratio with different PDF sets 117
6.25 Comparison of the measured cross section from the full data set with the prediction 118
6.26 Preliminary result for the ratio of single top quark and antiquark production with the data set of 2016 119

7.1 Integrated and peak luminosity of 2016 and 2017 122
7.2 Feynman diagram of single top quark production in the *s* channel 123
7.3 Relevant background processes in the search single top production in the *s* channel 125
7.4 Comparison of b tagging efficiencies between different tagging algorithms and years 131
7.5 QCD fit and modeling in the muon channel 134
7.6 QCD fit and modeling in the electron channel 135
7.7 Illustration of top quark reconstruction 136
7.8 Distributions of the three most important input variables in the 2-jets–2-tag signal region in 2016 139
7.9 Distributions of the three most important input variables in the 2-jets–2-tag signal region in 2017 140
7.10 Check for overtraining of the DNN 141
7.11 Prefit distributions of the DNN output in 2016 142
7.12 Prefit distributions of the DNN output in 2017 143
7.13 Postfit distributions of the DNN output in 2016 after the fit to 2016 and 2017 data 149
7.14 Postfit distributions of the DNN output in 2017 after the fit to 2016 and 2017 data 150
7.15 Combined comparison of prediction and data 151

8.1 Integrated and peak luminosity of 2016 154
8.2 Predicted cross section for the tHq and tHW signal processes 155
8.3 Analysis workflow scheme 156
8.4 Signal processes for top-Higgs associated production 157
8.5 Dominant background processes in the search for tH production 160
8.6 Further background processes in the search for tH production 161
8.7 Kinematic distributions of objects in the 2 tag control region 163
8.8 Corrections to the number of pileup interactions 167
8.9 Corrections for lepton efficiencies 169
8.10 Corrections on the b tagging discriminant shape 170
8.11 Illustration of the event reconstruction 172
8.12 Overtraining test for the training of the tHq event reconstruction 173
8.13 Result of the tHq event reconstruction 175
8.14 Overtraining test for the training of the tHW event reconstruction 176
8.15 Result of the tHq event reconstruction 178
8.16 Overtraining test for the training of the top quark pair event reconstruction . . 179
8.17 Result of the top quark pair event reconstruction 181

8.18 Overtraining test for the training of the single-lepton event classification . . . 183
8.19 Most important variables in the single-lepton event classification 184
8.20 Result of the single-lepton event classification 185
8.21 Overtraining test for the training of the dileptonic event classification 188
8.22 Most important variables in the dileptonic event classification 189
8.23 Result of the dileptonic event classification 190
8.24 Postfit distributions for the tH search in the ITC scenario 195
8.25 Postfit distributions for the tH search in the SM scenario 196
8.26 Combined comparison of prediction and data 197
8.27 Exclusion limits for the κ_t/κ_V scan 198
8.28 Exclusion limits for the CP-mixing scenario 200
8.29 Combination of exclusion limits 202
8.30 Likelihood scan for κ_t 203

A.1 Distributions of input variables for the event classification in the 2-jets–1-tag region 211
A.2 Distributions of input variables for the event classification in the 2-jets–1-tag region 212
A.3 Distributions of input variables for the event classification in the 2-jets–1-tag region 213
A.4 Distributions of input variables for the event classification in the 3-jets–1-tag region 214
A.5 Distributions of input variables for the event classification in the 3-jets–1-tag region 215
A.6 Distributions of input variables for the event classification in the 3-jets–1-tag region 216
A.7 Distributions of input variables for the event classification in the 3-jets–1-tag region 217
A.8 Distributions of input variables for the event classification in the 3-jets–2-tags region 218
A.9 Distributions of input variables for the event classification in the 3-jets–2-tags region 219
A.10 Distributions of input variables for the event classification in the 3-jets–2-tags region 220
A.11 Distributions of input variables for the event classification in the 3-jets–2-tags region 221

B.1 Corrections to the number of pileup interactions 226
B.2 Corrections for lepton efficiencies in 2016 227
B.3 Corrections for lepton efficiencies in 2017 228
B.4 Corrections for b tagging efficiencies in 2016 and 2017 229
B.5 QCD fit and modeling in the 2-jets–1-tag control region 230
B.6 QCD fit and modeling in the 3-jets–1-tag control region 231
B.7 Distributions of input variables for the event classification 232

B.8 Distributions of input variables for the event classification 233
B.9 Distributions of input variables for the event classification 234
B.10 Distributions of input variables for the event classification 235
B.11 Distributions of input variables for the event classification 236
B.12 Distributions of input variables for the event classification 237
B.13 Postfit distributions of the DNN output in 2016 after the fit to 2016 data 238
B.14 Postfit distributions of the DNN output in 2017 after the fit to 2017 data 239
B.15 Pulls of nuisance parameters in the combined fit 240

C.1 Effect of the event reweighting on the final-state particles for the tHq process . 243
C.2 Effect of the event reweighting on the final-state particles for the tHQ process . 244
C.3 Two-dimensional distribution of top quark and Higgs boson properties in the tHq process simulation 245
C.4 Two-dimensional distribution of top quark and Higgs boson properties in the tHW process simulation 246
C.5 Distributions of input variables for the tHq event reconstruction 248
C.6 Distributions of input variables for the tHq event reconstruction 249
C.7 Distributions of input variables for the tHq event reconstruction 250
C.8 Distributions of input variables for the tHW event reconstruction 251
C.9 Distributions of input variables for the tHW event reconstruction 252
C.10 Distributions of input variables for the tHW event reconstruction 253
C.11 Distributions of input variables for the $t\bar{t}$ event reconstruction 254
C.12 Distributions of input variables for the $t\bar{t}$ event reconstruction 255
C.13 Distributions of input variables for the single-lepton event classification 256
C.14 Distributions of input variables for the single-lepton event classification 257
C.15 Distributions of input variables for the single-lepton event classification 258
C.16 Distributions of input variables for the single-lepton event classification 259
C.17 Distributions of input variables for the single-lepton event classification 260
C.18 Distributions of input variables for the single-lepton event classification 261
C.19 Distributions of input variables for the dileptonic event classification 262
C.20 Distributions of input variables for the dilepton event classification 263

List of Tables

1.1 Quarks of the standard model . . . 2
1.2 Leptons of the standard model . . . 2
1.3 Bosons in the standard model . . . 3

5.1 Selection criteria for the muon identification . . . 70
5.2 Selection criteria for the electron identification . . . 73
5.3 Jet requirements . . . 75
5.4 Working points of b tagging algorithms . . . 77

6.1 Event selection criteria for different regions . . . 88
6.2 Event yields for the early measurement . . . 95
6.3 Evaluation of uncertainties in the early measurement . . . 98
6.4 Event yields for the 2-jets–1-tag signal region . . . 100
6.5 Input variable ranking of the neural network . . . 103
6.6 Result of the fit parameters . . . 108
6.7 Impact of experimental uncertainties . . . 115
6.8 Impact of theoretical uncertainties . . . 116

7.1 List of high-level-triggers used in the analysis . . . 127
7.2 Event selection criteria for the different regions . . . 128
7.3 Event yields for the 2-jets–2-tags signal region . . . 129
7.4 Event yields for the 2-jets–1-tag, 3-jets–1-tag and 3-jets–2-tags control regions 130
7.5 QCD estimation in the muon channel . . . 134
7.6 QCD estimation in the electron channel . . . 135
7.7 List of input variables for the DNN . . . 138

8.1 Trigger requirements for the single-lepton selection . . . 162
8.2 Event selection requirements for the single-lepton region . . . 164
8.3 Trigger requirements for the dileptonic selection . . . 165
8.4 Event selection requirements for the dileptonic region . . . 165
8.5 Event yields after the event selection . . . 166
8.6 Input variables of the tHq event reconstruction . . . 174
8.7 Input variables of the tHq event reconstruction . . . 177
8.8 Input variables of the top quark pair event reconstruction . . . 180
8.9 Input variables of the single-lepton event classification . . . 182

8.10 Input variables of the dileptonic event classification 187
8.11 List of PDF uncertainties 193
8.12 Exclusion limits on tH production 194
8.13 List of all exclusion limits for the κ_t/κ_V scan 199
8.14 List of all exclusion limits for the CP-mixing scenario 201

A.1 List of 2015 data sets used in the analysis 209
A.2 List of main simulation samples in 2015 209
A.3 List of additional simulation samples for systematic variations in 2015 210

B.1 List of 2016 data sets used in the analysis 223
B.2 List of 2017 data sets used in the analysis 223
B.3 List of main simulation samples in 2016 224
B.4 List of additional simulation samples for systematic variations in 2016 224
B.5 List of main simulation samples in 2017 225
B.6 List of additional simulation samples for systematic variations in 2017 225

C.1 List of 2016 data sets used in the analysis 241
C.2 List of main simulation samples in 2016 242
C.3 Parameter settings for the training of the BDTs 247

Bibliography

[1] G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett. B*, vol. 716, pp. 1–29, 2012.

[2] S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys. Lett. B*, vol. 716, pp. 30–61, 2012.

[3] C. Patrignani *et al.*, "Review of Particle Physics," *Chin. Phys. C*, vol. 40, no. 10, p. 100001, 2016.

[4] S. Tomonaga, "On a relativistically invariant formulation of the quantum theory of wave fields," *Prog. Theor. Phys.*, vol. 1, pp. 27–42, 1946.

[5] J. S. Schwinger, "On Quantum electrodynamics and the magnetic moment of the electron," *Phys. Rev.*, vol. 73, pp. 416–417, 1948.

[6] R. P. Feynman, "Space - time approach to quantum electrodynamics," *Phys. Rev.*, vol. 76, pp. 769–789, 1949.

[7] K. Hagiwara, A. D. Martin, D. Nomura, and T. Teubner, "Improved predictions for g-2 of the muon and alpha(QED) (M**2(Z))," *Phys. Lett. B*, vol. 649, pp. 173–179, 2007.

[8] G. Gabrielse, D. Hanneke, T. Kinoshita, M. Nio, and B. C. Odom, "New Determination of the Fine Structure Constant from the Electron g Value and QED," *Phys. Rev. Lett.*, vol. 97, p. 030802, 2006. [Erratum: Phys. Rev. Lett.99,039902(2007)].

[9] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, "Experimental Test of Parity Conservation in Beta Decay," *Phys. Rev.*, vol. 105, pp. 1413–1414, 1957.

[10] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc.*, vol. C680519, pp. 367–377, 1968.

[11] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964.

[12] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.*, vol. 13, pp. 508–509, 1964.

[13] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles," *Phys. Rev. Lett.*, vol. 13, pp. 585–587, 1964.

[14] S. L. Glashow, "The renormalizability of vector meson interactions," *Nucl. Phys.*, vol. 10, pp. 107–117, 1959.

[15] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 1967.

[16] D. B. Chitwood *et al.*, "Improved measurement of the positive muon lifetime and determination of the Fermi constant," *Phys. Rev. Lett.*, vol. 99, p. 032001, 2007.

[17] L. Álvarez Gaumé and J. Ellis, "Eyes on a prize particle," *Nat Phys*, vol. 7, no. 1, pp. 2–3, 2011.

[18] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," *Phys. Rev. Lett.*, vol. 10, pp. 531–533, 1963. [,648(1963)].

[19] S. L. Glashow, J. Iliopoulos, and L. Maiani, "Weak Interactions with Lepton-Hadron Symmetry," *Phys. Rev. D*, vol. 2, pp. 1285–1292, 1970.

[20] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction," *Prog. Theor. Phys.*, vol. 49, pp. 652–657, 1973.

[21] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Nonabelian Gauge Theories," *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, 1973.

[22] H. D. Politzer, "Reliable Perturbative Results for Strong Interactions?," *Phys. Rev. Lett.*, vol. 30, pp. 1346–1349, 1973.

[23] K. G. Wilson, "Confinement of Quarks," *Phys. Rev. D*, vol. 10, pp. 2445–2459, 1974. [,45(1974)].

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

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

[26] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl. Phys. B*, vol. 126, pp. 298–318, 1977.

[27] G. Sterman and S. Weinberg, "Jets from quantum chromodynamics," *Phys. Rev. Lett.*, vol. 39, pp. 1436–1439, Dec 1977.

[28] M. Dobbs and J. B. Hansen, "The HepMC C++ Monte Carlo event record for High Energy Physics," *Comput. Phys. Commun.*, vol. 134, pp. 41–46, 2001.

[29] S. W. Herb *et al.*, "Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions," *Phys. Rev. Lett.*, vol. 39, pp. 252–255, 1977.

[30] F. Abe *et al.*, "Observation of top quark production in $\bar{p}p$ collisions," *Phys. Rev. Lett.*, vol. 74, pp. 2626–2631, 1995.

[31] S. Abachi *et al.*, "Observation of the top quark," *Phys. Rev. Lett.*, vol. 74, pp. 2632–2637, 1995.

[32] T. Behnke, J. E. Brau, B. Foster, J. Fuster, M. Harrison, J. M. Paterson, M. Peskin, M. Stanitzki, N. Walker, and H. Yamamoto, "The International Linear Collider Technical Design Report - Volume 1: Executive Summary," Technical Design Report, ILC, 2013.

[33] M. J. Boland *et al.*, "Updated baseline for a staged Compact Linear Collider," Technical Design Report, CLIC, 2016.

[34] ATLAS and C. collaborations, "NNLO+NNLL top-quark-pair cross sections." `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO`, Nov. 2017.

[35] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$," *Phys. Rev. Lett.*, vol. 110, p. 252004, 2013.

[36] LHC Top Physics Working Group, "LHCTopWG Summary Plots." `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots`, Nov. 2017.

[37] ATLAS and CMS collaborations, "NLO single-top channel cross sections." `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec`, Nov. 2017.

[38] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann, "HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR," *Comput. Phys. Commun.*, vol. 182, pp. 1034–1046, 2011.

[39] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz, P. Rieck, and P. Uwer, "HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions," *Comput. Phys. Commun.*, vol. 191, pp. 74–89, 2015.

[40] M. Botje *et al.*, "The PDF4LHC Working Group Interim Recommendations," *arXiv:1101.0538 [hep-ph]*, 2011.

[41] J. Butterworth *et al.*, "PDF4LHC recommendations for LHC Run II," *J. Phys.*, vol. G43, p. 023001, 2016.

[42] M. Brucherseifer, F. Caola, and K. Melnikov, "On the NNLO QCD corrections to single-top production at the LHC," *Phys. Lett. B*, vol. 736, pp. 58–63, 2014.

[43] E. L. Berger, J. Gao, C. P. Yuan, and H. X. Zhu, "NNLO QCD Corrections to t-channel Single Top-Quark Production and Decay," *Phys. Rev. D*, vol. 94, no. 7, p. 071501, 2016.

[44] I. I. Y. Bigi, Y. L. Dokshitzer, V. A. Khoze, J. H. Kuhn, and P. M. Zerwas, "Production and Decay Properties of Ultraheavy Quarks," *Phys. Lett. B*, vol. 181, pp. 157–163, 1986.

[45] J. R. Ellis, M. K. Gaillard, and D. V. Nanopoulos, "A Phenomenological Profile of the Higgs Boson," *Nucl. Phys. B*, vol. 106, p. 292, 1976.

[46] R. Barate *et al.*, "Search for the standard model Higgs boson at LEP," *Phys. Lett. B*, vol. 565, pp. 61–75, 2003.

[47] T. Aaltonen *et al.*, "Combination fo Searches for the Higgs Boson Using the Full CDF Data Set," *Phys. Rev. D*, vol. 88, no. 5, p. 052013, 2013.

[48] V. M. Abazov *et al.*, "Combined search for the Higgs boson with the D0 experiment," *Phys. Rev. D*, vol. 88, no. 5, p. 052011, 2013.

[49] G. Aad *et al.*, "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV," *JHEP*, vol. 08, p. 045, 2016.

[50] D. de Florian *et al.*, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector," CERN Yellow Reports CERN-2017-002-M, CERN, 2016.

[51] A. M. Sirunyan *et al.*, "Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector," *Phys. Lett. B*, vol. 779, pp. 283–316, 2018.

[52] M. Aaboud *et al.*, "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector," *arXiv:1808.08238 [hep-ex]*, 2018.

[53] A. M. Sirunyan *et al.*, "Observation of Higgs boson decay to bottom quarks," *arXiv:1808.08242 [hep-ex]*, 2018.

[54] LHC Higgs Cross Section Working Group, "Higgs decay width and branching ratios." `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWGCrossSectionsFigures`, Nov. 2013.

[55] J. R. Andersen *et al.*, "Handbook of LHC Higgs Cross Sections: 3. Higgs Properties," CERN Yellow Reports CERN-2013-004, CERN, 2013.

[56] M. Aaboud *et al.*, "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector," *Phys. Lett. B*, vol. 784, pp. 173–191, 2018.

[57] A. M. Sirunyan *et al.*, "Observation of $\mathrm{t\bar{t}H}$ production," *Phys. Rev. Lett.*, vol. 120, p. 231801, 2018.

[58] F. Maltoni, K. Paul, T. Stelzer, and S. Willenbrock, "Associated production of Higgs and single top at hadron colliders," *Phys. Rev. D*, vol. 64, p. 094023, 2001.

[59] S. Biswas, E. Gabrielli, and B. Mele, "Single top and Higgs associated production as a probe of the Htt coupling sign at the LHC," *JHEP*, vol. 01, p. 088, 2013.

[60] M. Farina, C. Grojean, F. Maltoni, E. Salvioni, and A. Thamm, "Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson," *JHEP*, vol. 05, p. 022, 2013.

[61] J. Chang, K. Cheung, J. S. Lee, and C.-T. Lu, "Probing the Top-Yukawa Coupling in Associated Higgs production with a Single Top Quark," *JHEP*, vol. 05, p. 062, 2014.

[62] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, "Higgs production in association with a single top quark at the LHC," *Eur. Phys. J. C*, vol. 75, no. 6, p. 267, 2015.

[63] CMS Collaboration, "Combined measurements of the Higgs boson's couplings at $\sqrt{s}$ = 13 TeV," CMS Physics Analysis Summary CMS-PAS-HIG-17-031, CERN, Geneva, 2018.

[64] V. Khachatryan *et al.*, "Search for the associated production of a Higgs boson with a single top quark in proton-proton collisions at $\sqrt{s}$ = 8 TeV," *JHEP*, vol. 06, p. 177, 2016.

[65] A. M. Sirunyan *et al.*, "Constraints on anomalous Higgs boson couplings using production and decay information in the four-lepton final state," *Phys. Lett. B*, vol. 775, pp. 1–24, 2017.

[66] M. Aaboud *et al.*, "Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *JHEP*, vol. 03, p. 095, 2018.

[67] A. Einstein, "Die Grundlage der allgemeinen Relativitätstheorie," *Annalen der Physik*, vol. 354, pp. 769–822, 1916.

[68] V. C. Rubin, N. Thonnard, and W. K. Ford, Jr., "Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R = 4 kpc) to UGC 2885 (R = 122 kpc)," *Astrophys. J.*, vol. 238, p. 471, 1980.

[69] E. Corbelli and P. Salucci, "The Extended Rotation Curve and the Dark Matter Halo of M33," *Mon. Not. Roy. Astron. Soc.*, vol. 311, pp. 441–447, 2000.

[70] E. Aubourg *et al.*, "Evidence for gravitational microlensing by dark objects in the galactic halo," *Nature*, vol. 365, pp. 623–625, 1993.

[71] C. Alcock *et al.*, "Possible Gravitational Microlensing of a Star in the Large Magellanic Cloud," *Nature*, vol. 365, pp. 621–623, 1993.

[72] G. Hinshaw *et al.*, "Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results," *Astrophys. J. Suppl.*, vol. 180, pp. 225–245, 2009.

[73] P. A. R. Ade *et al.*, "Planck 2015 results. XIII. Cosmological parameters," *Astron. Astrophys.*, vol. 594, p. A13, 2016.

[74] Y. Fukuda *et al.*, "Evidence for oscillation of atmospheric neutrinos," *Phys. Rev. Lett.*, vol. 81, pp. 1562–1567, 1998.

[75] Q. R. Ahmad *et al.*, "Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory," *Phys. Rev. Lett.*, vol. 87, p. 071301, 2001.

[76] C. Athanassopoulos *et al.*, "Candidate events in a search for anti-muon-neutrino —> anti-electron-neutrino oscillations," *Phys. Rev. Lett.*, vol. 75, pp. 2650–2653, 1995.

[77] C. Athanassopoulos *et al.*, "Evidence for neutrino oscillations from muon decay at rest," *Phys. Rev. C*, vol. 54, pp. 2685–2708, 1996.

[78] C. Athanassopoulos *et al.*, "Evidence for muon-neutrino —> electron-neutrino oscillations from pion decay in flight neutrinos," *Phys. Rev. C*, vol. 58, pp. 2489–2511, 1998.

[79] A. Aguilar-Arevalo *et al.*, "Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam," *Phys. Rev. D*, vol. 64, p. 112007, 2001.

[80] A. A. Aguilar-Arevalo *et al.*, "A Search for electron neutrino appearance at the $\Delta m^2 \sim 1\text{eV}^2$ scale," *Phys. Rev. Lett.*, vol. 98, p. 231801, 2007.

[81] A. A. Aguilar-Arevalo *et al.*, "Unexplained Excess of Electron-Like Events From a 1-GeV Neutrino Beam," *Phys. Rev. Lett.*, vol. 102, p. 101802, 2009.

[82] A. A. Aguilar-Arevalo *et al.*, "Event Excess in the MiniBooNE Search for $\bar{\nu}_\mu \to \bar{\nu}_e$ Oscillations," *Phys. Rev. Lett.*, vol. 105, p. 181801, 2010.

[83] A. A. Aguilar-Arevalo *et al.*, "Observation of a Significant Excess of Electron-Like Events in the MiniBooNE Short-Baseline Neutrino Experiment," *arXiv:1805.12028 [hep-ex]*, 2018.

[84] A. D. Sakharov, "Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe," *Pisma Zh. Eksp. Teor. Fiz.*, vol. 5, pp. 32–35, 1967. [Usp. Fiz. Nauk161,61(1991)].

[85] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov, "On the Anomalous Electroweak Baryon Number Nonconservation in the Early Universe," *Phys. Lett. B*, vol. 155, p. 36, 1985.

[86] A. Hocker *et al.*, "TMVA - Toolkit for Multivariate Data Analysis," *PoS*, vol. ACAT, p. 040, 2007.

[87] F. Chollet *et al.*, "Keras." `https://github.com/fchollet/keras`, 2015.

[88] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems." `tensorflow.org`, 2015.

[89] C. Gini, "Variabilità e mutabilità: Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche,". Bologna: Tipografia di Paolo Cuppin, 1912.

[90] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, 1997.

[91] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.

[92] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms 1. general considerations," *IMA Journal of Applied Mathematics*, vol. 6, pp. 76–90, Mar. 1970.

[93] R. Fletcher, "A new approach to variable metric algorithms.," *Comput. J.*, vol. 13, no. 3, pp. 317–322, 1970.

[94] D. Goldfarb, "A family of variable-metric methods derived by variational means," *Mathematics of Computation*, vol. 24, no. 109, pp. pp. 23–26, 1970.

[95] D. F. Shanno, "Conditioning of quasi-newton methods for function minimization," *Mathematics of Computation*, vol. 24, no. 111, pp. pp. 647–656, 1970.

[96] "Overtraining." `https://commons.wikimedia.org/wiki/File:Overfitting.svg`, Feb. 2008.

[97] A. N. Kolmogorov, "Sulla Determinazione Empirica di una Legge di Distribuzione," *Giornale dell'Istituto Italiano degli Attuari*, vol. 4, pp. 83–91, 1933.

[98] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions," *Ann. Math. Statist.*, vol. 19, pp. 279–281, 06 1948.

[99] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *CoRR*, vol. abs/1207.0580, 2012.

[100] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014.

[101] "ROC curve." `https://en.wikipedia.org/wiki/File:ROC_curves.svg`, Oct. 2015.

[102] Th. Müller, J. Ott and J. Wagner-Kuhr, "theta - a framework for template-based modeling and inference," CMS Internal Note CMS-IN-2010-017, CMS, Jun 2010.

[103] Higgs PAG, "Documentation of the RooStats -based statistics tools for Higgs PAG." `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit`, Aug. 2017.

[104] G. Aad *et al.*, "Combined Measurement of the Higgs Boson Mass in *pp* Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments," *Phys. Rev. Lett.*, vol. 114, p. 191803, 2015.

[105] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling," *eConf*, vol. C0303241, p. MOLT007, 2003. [,186(2003)].

[106] A. L. Read, "Linear interpolation of histograms," *Nucl. Instrum. Meth. A*, vol. 425, pp. 357–360, 1999.

[107] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke, "Interpolation between multi-dimensional histograms using a new non-linear moment morphing method," *Nucl. Instrum. Meth. A*, vol. 771, pp. 39–48, 2015.

[108] J. Neyman and E. Pearson, "Ix. on the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 231, no. 694-706, pp. 289–337, 1933.

[109] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics," *Eur. Phys. J. C*, vol. 71, p. 1554, 2011. [Erratum: Eur. Phys. J.C73,2501(2013)].

[110] G. J. Feldman and R. D. Cousins, "A Unified approach to the classical statistical analysis of small signals," *Phys. Rev. D*, vol. 57, pp. 3873–3889, 1998.

[111] A. L. Read, "Modified frequentist analysis of search results (The CL(s) method)," CERN Yellow Reports CERN-2000-005, CERN, Geneva, 2000.

[112] A. L. Read, "Presentation of search results: The CL(s) technique," *J. Phys. G*, vol. 28, pp. 2693–2704, 2002. [,11(2002)].

[113] A. Wald, "Tests of statistical hypotheses concerning several parameters when the number of observations is large," *Transactions of the American Mathematical Society*, vol. 54, no. 3, pp. 426–482, 1943.

[114] S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *Annals Math. Statist.*, vol. 9, no. 1, pp. 60–62, 1938.

[115] L. Evans and P. Bryant, "LHC Machine," *JINST*, vol. 3, p. S08001, 2008.

[116] CERN, "The Large Hadron Collider." `https://home.cern/topics/large-hadron-collider`, Nov. 2017.

[117] CERN, "About CERN." `https://home.cern/about`, Nov. 2017.

[118] CERN, "Linear accelerator 2." `https://home.cern/about/accelerators/linear-accelerator-2`, Nov. 2017.

[119] CERN, "The Proton Synchrotron Booster." `https://home.cern/about/accelerators/proton-synchrotron-booster`, Nov. 2017.

[120] CERN, "The Proton Synchrotron." `https://home.cern/about/accelerators/proton-synchrotron`, Nov. 2017.

[121] CERN, "The Super Proton Synchrotron." `https://home.cern/about/accelerators/super-proton-synchrotron`, Nov. 2017.

[122] G. Arnison *et al.*, "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV," *Phys. Lett. B*, vol. 122, pp. 103–116, 1983. [,611(1983)].

[123] M. Banner *et al.*, "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider," *Phys. Lett. B*, vol. 122, pp. 476–485, 1983.

[124] E. Mobs, "The CERN accelerator complex. Complexe des accélérateurs du CERN," Jul 2016. General Photo.

[125] CMS Collaboration, "Public CMS Luminosity Information." https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults, June 2018.

[126] S. Chatrchyan *et al.*, "The CMS Experiment at the CERN LHC," *JINST*, vol. 3, p. S08004, 2008.

[127] S. Chatrchyan *et al.*, "Description and performance of track and primary-vertex reconstruction with the CMS tracker," *JINST*, vol. 9, no. 10, p. P10009, 2014.

[128] V. Karimäki, M. Mannelli, P. Siegrist, H. Breuker, A. Caner, R. Castaldi, K. Freudenreich, G. Hall, R. Horisberger, M. Huhtinen, and A. Cattai, "The CMS tracker system project: Technical Design Report," Technical Design Report CERN-LHCC-98-006, CMS, Geneva, 1997.

[129] D. A. Matzner Dominguez, D. Abbaneo, K. Arndt, N. Bacchetta, A. Ball, E. Bartz, W. Bertl, G. M. Bilei, G. Bolla, H. W. K. Cheung, *et al.*, "CMS Technical Design Report for the Pixel Detector Upgrade," Technical Design Report CERN-LHCC-2012-016, CMS, Geneva, 2012.

[130] K. Klein, "The Phase-2 Upgrade of the CMS Tracker," Technical Design Report CERN-LHCC-2017-009, CMS, Geneva, 2017.

[131] CMS Collaboration, "CMS: The electromagnetic calorimeter. Technical design report," Technical Design Report CERN-LHCC-97-33, CMS, Geneva, 1997.

[132] G. L. Bayatian *et al.*, "CMS Physics : Technical Design Report Volume 1: Detector Performance and Software," Technical Design Report CERN-LHCC-2006-001, CMS, Geneva, 2006.

[133] C. W. Fabjan and F. Gianotti, "Calorimetry for particle physics," *Rev. Mod. Phys.*, vol. 75, pp. 1243–1286, 2003.

[134] P. Adzic *et al.*, "Energy resolution of the barrel of the CMS electromagnetic calorimeter," *JINST*, vol. 2, p. P04004, 2007.

[135] CMS Collaboration, "CMS: The hadron calorimeter technical design report," Technical Design Report CERN-LHCC-97-31, CMS, Geneva, 1997.

[136] G. Acquistapace *et al.*, "CMS, the magnet project: Technical design report," Technical Design Report CERN-LHCC-97-10, CMS, Geneva, 1997.

[137] D. Campi *et al.*, "Commissioning of the CMS Magnet," *IEEE Trans. Appl. Supercond.*, vol. 17, no. 2, pp. 1185–1190, 2007.

[138] S. Chatrchyan *et al.*, "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays," *JINST*, vol. 5, p. T03021, 2010.

[139] C. G. L. E. Committee, "The CMS muon project," Technical Design Report CERN-LHCC-97-032, CMS, Geneva, 1997.

[140] S. Chatrchyan *et al.*, "Calibration of the CMS Drift Tube Chambers and Measurement of the Drift Velocity with Cosmic Rays," *JINST*, vol. 5, p. T03016, 2010.

[141] S. Chatrchyan *et al.*, "Performance of the CMS Cathode Strip Chambers with Cosmic Rays," *JINST*, vol. 5, p. T03018, 2010.

[142] S. Chatrchyan *et al.*, "Performance Study of the CMS Barrel Resistive Plate Chambers with Cosmic Rays," *JINST*, vol. 5, p. T03017, 2010.

[143] S. Chatrchyan *et al.*, "The performance of the CMS muon detector in proton-proton collisions at sqrt(s) = 7 TeV at the LHC," *JINST*, vol. 8, p. P11002, 2013.

[144] V. Khachatryan *et al.*, "The CMS trigger system," *JINST*, vol. 12, no. 01, p. P01020, 2017.

[145] S. Dasu *et al.*, "CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems," Technical Design Report CERN-LHCC-2000-038, CMS, Geneva, 2000.

[146] P. Sphicas, "CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger," Technical Design Report CERN-LHCC-2002-026, CMS, Geneva, 2002.

[147] I. Bird, K. Bos, N. Brook, D. Duellmann, C. Eck, I. Fisk, D. Foster, B. Gibbard, C. Grandi, F. Grey, *et al.*, "LHC computing Grid. Technical design report," Technical Design Report CERN-LHCC-2005-024, CERN, Geneva, 2005.

[148] I. Bird *et al.*, "Update of the Computing Models of the WLCG and the LHC Experiments," Technical Design Report CERN-LHCC-2014-014, CERN, Geneva, 2014.

[149] M. Aderholz *et al.*, "Models of networked analysis at regional centres for LHC experiments (MONARC). Phase 2 report," report, CERN, Geneva, 2000.

[150] E. Martelli and S. Stancu, "LHCOPN and LHCONE: Status and Future Evolution," *J. Phys. Conf. Ser.*, vol. 664, no. 5, p. 052025, 2015.

[151] F. James, "Monte Carlo Theory and Practice," *Rept. Prog. Phys.*, vol. 43, p. 1145, 1980.

[152] F. A. Berends, R. Kleiss, and S. Jadach, "Monte Carlo Simulation of Radiative Corrections to the Processes e+ e- —> mu+ mu- and e+ e- —> anti-q q in the Z0 Region," *Comput. Phys. Commun.*, vol. 29, pp. 185–200, 1983.

[153] S. Höche, "Introduction to parton-shower event generators," in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, pp. 235–295, 2015.

[154] S. Alekhin, J. Blumlein, and S. Moch, "The ABM parton distributions tuned to LHC data," *Phys. Rev. D*, vol. 89, no. 5, p. 054028, 2014.

[155] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, "New parton distribution functions from a global analysis of quantum chromodynamics," *Phys. Rev. D*, vol. 93, no. 3, p. 033006, 2016.

[156] F. D. Aaron *et al.*, "Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA," *JHEP*, vol. 01, p. 109, 2010.

[157] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs," *Eur. Phys. J. C*, vol. 75, no. 5, p. 204, 2015.

[158] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, "Reweighting and Unweighting of Parton Distributions and the LHC W lepton asymmetry data," *Nucl. Phys. B*, vol. 855, pp. 608–638, 2012.

[159] R. D. Ball *et al.*, "Parton distributions for the LHC Run II," *JHEP*, vol. 04, p. 040, 2015.

[160] R. D. Ball *et al.*, "Parton distributions from high-precision collider data," *Eur. Phys. J. C*, vol. 77, no. 10, p. 663, 2017.

[161] V. V. Sudakov, "Vertex parts at very high-energies in quantum electrodynamics," *Sov. Phys. JETP*, vol. 3, pp. 65–71, 1956. [Zh. Eksp. Teor. Fiz.30,87(1956)].

[162] A. Sen, "Asymptotic Behavior of the Sudakov Form-Factor in QCD," *Phys. Rev. D*, vol. 24, p. 3281, 1981.

[163] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, "QCD matrix elements + parton showers," *JHEP*, vol. 11, p. 063, 2001.

[164] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, "Matching matrix elements and shower evolution for top-quark production in hadronic collisions," *JHEP*, vol. 01, p. 013, 2007.

[165] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO," *JHEP*, vol. 12, p. 061, 2012.

[166] F. Maltoni, G. Ridolfi, and M. Ubiali, "b-initiated processes at the LHC: a reappraisal," *JHEP*, vol. 07, p. 022, 2012. [Erratum: JHEP04,095(2013)].

[167] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton Fragmentation and String Dynamics," *Phys. Rept.*, vol. 97, pp. 31–145, 1983.

[168] B. R. Webber, "A QCD Model for Jet Fragmentation Including Soft Gluon Interference," *Nucl. Phys. B*, vol. 238, pp. 492–528, 1984.

[169] E. Boos *et al.*, "Generic user process interface for event generators," in *Physics at TeV colliders. Proceedings, Euro Summer School, Les Houches, France, May 21-June 1, 2001*, 2001.

[170] J. Alwall *et al.*, "A Standard format for Les Houches event files," *Comput. Phys. Commun.*, vol. 176, pp. 300–304, 2007.

[171] A. Buckley *et al.*, "General-purpose event generators for LHC physics," *Phys. Rept.*, vol. 504, pp. 145–233, 2011.

[172] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP*, vol. 07, p. 079, 2014.

[173] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *JHEP*, vol. 06, p. 029, 2002.

[174] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, "MadGraph 5 : Going Beyond," *JHEP*, vol. 06, p. 128, 2011.

[175] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms," *JHEP*, vol. 11, p. 040, 2004.

[176] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," *JHEP*, vol. 11, p. 070, 2007.

[177] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP*, vol. 06, p. 043, 2010.

[178] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 05, p. 026, 2006.

[179] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.*, vol. 178, pp. 852–867, 2008.

[180] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An Introduction to PYTHIA 8.2," *Comput. Phys. Commun.*, vol. 191, pp. 159–177, 2015.

[181] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 Tune," *Eur. Phys. J. C*, vol. 74, no. 8, p. 3024, 2014.

[182] V. Khachatryan *et al.*, "Event generator tunes obtained from underlying event and multi-parton scattering measurements," *Eur. Phys. J. C*, vol. 76, no. 3, p. 155, 2016.

[183] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements," CMS Physics Analysis Summary CMS-PAS-GEN-17-001, CERN, Geneva, 2018.

[184] M. Bahr *et al.*, "Herwig++ Physics and Manual," *Eur. Phys. J. C*, vol. 58, pp. 639–707, 2008.

[185] S. Gieseke, C. Rohr, and A. Siodmok, "Colour reconnections in Herwig++," *Eur. Phys. J. C*, vol. 72, p. 2225, 2012.

[186] S. Agostinelli *et al.*, "GEANT4: A Simulation toolkit," *Nucl. Instrum. Meth. A*, vol. 506, pp. 250–303, 2003.

[187] J. Allison *et al.*, "Geant4 developments and applications," *IEEE Trans. Nucl. Sci.*, vol. 53, p. 270, 2006.

[188] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET," CMS Physics Analysis Summary CMS-PAS-PFT-09-001, CERN, Geneva, 2009.

[189] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector," CMS Physics Analysis Summary CMS-PAS-PFT-10-001, CERN, Geneva, 2010.

[190] A. M. Sirunyan *et al.*, "Particle-flow reconstruction and global event description with the CMS detector," *JINST*, vol. 12, no. 10, p. P10003, 2017.

[191] D. Buskulic *et al.*, "Performance of the ALEPH detector at LEP," *Nucl. Instrum. Meth. A*, vol. 360, pp. 481–506, 1995.

[192] S. Chatrchyan *et al.*, "Alignment of the CMS tracker with LHC and cosmic ray data," *JINST*, vol. 9, p. P06009, 2014.

[193] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting," *Nucl. Instrum. Meth. A*, vol. 262, pp. 444–450, 1987.

[194] P. Billoir, "Progressive track recognition with a Kalman like fitting procedure," *Comput. Phys. Commun.*, vol. 57, pp. 390–394, 1989.

[195] P. Billoir and S. Qian, "Simultaneous pattern recognition and track fitting by the Kalman filtering method," *Nucl. Instrum. Meth. A*, vol. 294, pp. 219–228, 1990.

[196] R. Mankel, "A Concurrent track evolution algorithm for pattern recognition in the HERA-B main tracking system," *Nucl. Instrum. Meth. A*, vol. 395, pp. 169–184, 1997.

[197] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems," *Proceedings of the IEEE*, vol. 86, pp. 2210–2239, Nov 1998.

[198] R. Frühwirth and A. Strandlie, "Track fitting with ambiguities and noise: A study of elastic tracking and nonlinear filters," *Comput. Phys. Commun.*, vol. 120, no. 2-3, pp. 197–214, 1999.

[199] T. Speer, K. Prokofiev, R. Frühwirth, W. Waltenberger, and P. Vanlaer, "Vertex fitting in the CMS tracker," CMS Note CERN-CMS-NOTE-2006-032, CMS, Geneva, 2006.

[200] R. Frühwirth, W. Waltenberger, and P. Vanlaer, "Adaptive vertex fitting," *J. Phys. G*, vol. 34, p. N343, 2007.

[201] CMS Collaboration, "Pileup Removal Algorithms," CMS Physics Analysis Summary CMS-PAS-JME-14-001, CERN, Geneva, 2014.

[202] S. Chatrchyan *et al.*, "Performance of CMS Muon Reconstruction in Cosmic-Ray Events," *JINST*, vol. 5, p. T03022, 2010.

[203] S. Chatrchyan *et al.*, "Performance of CMS muon reconstruction in *pp* collision events at $\sqrt{s}$ = 7 TeV," *JINST*, vol. 7, p. P10002, 2012.

[204] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, I. Puljak, C. Rovelli, R. Salerno, and Y. Sirois, "Electron reconstruction in CMS," *Eur. Phys. J. C*, vol. 49, pp. 1099–1116, 2007.

[205] V. Khachatryan *et al.*, "Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s}$ = 8 TeV," *JINST*, vol. 10, no. 06, p. P06005, 2015.

[206] R. Frühwirth, "Track fitting with nonGaussian noise," *Comput. Phys. Commun.*, vol. 100, pp. 1–16, 1997.

[207] W. Adam, R. Frühwirth, A. Strandlie, and T. Todor, "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC," CMS Note CERN-CMS-NOTE-2005-001, CMS, Geneva, 2005.

[208] V. Khachatryan *et al.*, "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at sqrt(s) = 8 TeV," *JINST*, vol. 10, no. 08, p. P08010, 2015.

[209] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP*, vol. 04, p. 063, 2008.

[210] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions," *Nucl. Phys. B*, vol. 406, pp. 187–224, 1993.

[211] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better jet clustering algorithms," *JHEP*, vol. 08, p. 001, 1997.

[212] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder," *Phys. Lett. B*, vol. 641, pp. 57–61, 2006.

[213] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual," *Eur. Phys. J. C*, vol. 72, p. 1896, 2012.

[214] CMS Collaboration, "Introduction to Jet Energy Corrections at CMS." `https://twiki.cern.ch/twiki/bin/view/CMS/IntroToJEC`, Feb. 2016.

[215] S. Chatrchyan *et al.*, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS," *JINST*, vol. 6, p. P11002, 2011.

[216] V. Khachatryan *et al.*, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV," *JINST*, vol. 12, no. 02, p. P02014, 2017.

[217] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas," *Phys. Lett. B*, vol. 659, pp. 119–126, 2008.

[218] S. Chatrchyan *et al.*, "Identification of b-quark jets with the CMS experiment," *JINST*, vol. 8, p. P04013, 2013.

[219] A. M. Sirunyan *et al.*, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV," *JINST*, vol. 13, no. 05, p. P05011, 2018.

[220] D. Guest, J. Collado, P. Baldi, S.-C. Hsu, G. Urban, and D. Whiteson, "Jet Flavor Classification in High-Energy Physics with Deep Neural Networks," *Phys. Rev. D*, vol. 94, no. 11, p. 112002, 2016.

[221] V. Khachatryan *et al.*, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s}$ = 8 TeV," *JINST*, vol. 10, no. 02, p. P02006, 2015.

[222] CMS Collaboration, "MET Analysis." `https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMetAnalysis`, June 2015.

[223] CMS Collaboration, "Muon Physics Object Group." `https://twiki.cern.ch/twiki/bin/view/CMS/MuonPOG`, Feb. 2018.

[224] CMS Collaboration, "Baseline muon selections for Run-II." `https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideMuonIdRun2`, Feb. 2018.

[225] S. Chatrchyan *et al.*, "Search for neutral Higgs bosons decaying to tau pairs in *pp* collisions at $\sqrt{s}$ = 7 TeV," *Phys. Lett. B*, vol. 713, pp. 68–90, 2012.

[226] CMS Collaboration, "E/gamma Physics Object Group." `https://twiki.cern.ch/twiki/bin/view/CMS/EgammaPOG`, Feb. 2018.

[227] CMS Collaboration, "Cut Based Electron ID for Run 2." `https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2`, Jan. 2018.

[228] CMS Collaboration, "Jet algorithms performance in 13 TeV data," CMS Physics Analysis Summary CMS-PAS-JME-16-003, CERN, Geneva, 2017.

[229] CMS Collaboration, "Jet Identification." `https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID`, Dec. 2016.

[230] CMS Collaboration, "Jet Identification for the 13 TeV data Run2016." `https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVRun2016`, Feb. 2018.

[231] CMS Collaboration, "JetID13TeVRun2017." https://twiki.cern.ch/twiki/bin/view/CMS/JetID13TeVRun2017, Feb. 2018.

[232] CMS Collaboration, "Jet Energy Resolution." https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution, July 2017.

[233] T. Chwalek, "Measurement of W -Boson Helicity-Fractions in Top-Quark Decays with the CDF II Experiment and Prospects for an Early $t\bar{t}$ Cross-Section Measurement with the CMS Experiment,". PhD thesis, CERN-THESIS-2010-255, Karlsruhe U., 2010.

[234] T. Aaltonen *et al.*, "First Observation of Electroweak Single Top Quark Production," *Phys. Rev. Lett.*, vol. 103, p. 092002, 2009.

[235] V. M. Abazov *et al.*, "Observation of Single Top Quark Production," *Phys. Rev. Lett.*, vol. 103, p. 092001, 2009.

[236] S. Chatrchyan *et al.*, "Measurement of the t-channel single top quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV," *Phys. Rev. Lett.*, vol. 107, p. 091802, 2011.

[237] S. Chatrchyan *et al.*, "Measurement of the single-top-quark t-channel cross section in pp collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 12, p. 035, 2012.

[238] V. Khachatryan *et al.*, "Measurement of the t-channel single-top-quark production cross section and of the $| V_{tb} |$ CKM matrix element in pp collisions at $\sqrt{s}$= 8 TeV," *JHEP*, vol. 06, p. 090, 2014.

[239] G. Aad *et al.*, "Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Phys. Lett. B*, vol. 717, pp. 330–350, 2012.

[240] G. Aad *et al.*, "Comprehensive measurements of t-channel single top-quark production cross sections at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Phys. Rev. D*, vol. 90, no. 11, p. 112006, 2014.

[241] M. Aaboud *et al.*, "Fiducial, total and differential cross-section measurements of t-channel single top-quark production in pp collisions at 8 TeV using data collected by the ATLAS detector," *Eur. Phys. J. C*, vol. 77, no. 8, p. 531, 2017.

[242] CMS Collaboration, "CMS Luminosity Measurement for the 2015 Data Taking Period," CMS Physics Analysis Summary CMS-PAS-LUM-15-001, CERN, Geneva, 2016.

[243] J. A. Aguilar-Saavedra, "A Minimal set of top anomalous couplings," *Nucl. Phys. B*, vol. 812, pp. 181–204, 2009.

[244] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method," *Eur. Phys. J. C*, vol. 71, p. 1547, 2011.

[245] CMS Collaboration, "Utilities for Accessing Pileup Information for Data." https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData, Mar. 2017.

[246] CMS Collaboration, "Pileup Reweighting Utilities." https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupMCReweightingUtilities, Apr. 2013.

[247] CMS Collaboration, "Reference muon id, isolation and trigger efficiencies for Run-II ." https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffsRun2, Feb. 2018.

[248] V. Khachatryan *et al.*, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 01, p. 080, 2011.

[249] CMS Collaboration, "Tag and Probe." https://twiki.cern.ch/twiki/bin/view/CMS/MuonTagAndProbe, Mar. 2014.

[250] CMS Collaboration, "Muon T&P Instructions for Run-II." https://twiki.cern.ch/twiki/bin/view/CMS/MuonTagAndProbeTreesRun2, Dec. 2017.

[251] CMS Collaboration, "Methods to apply b-tagging efficiency scale factors." https://twiki.cern.ch/twiki/bin/view/CMS/BTagSFMethods, Jan. 2018.

[252] CMS Collaboration, "b Tag & Vertexing Physics Object Group." https://twiki.cern.ch/twiki/bin/view/CMS/BtagPOG, Feb. 2018.

[253] CMS Collaboration, "Usage of b Tag Objects for 13 TeV Data with 25ns bunch spacing and 76X ReReco." https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation76X, Apr. 2016.

[254] CMS Collaboration, "First look at top-quark selections with 13 TeV data," CMS Detector Performance Note CMS-DP-2015-019, CERN, Geneva, Jul 2015.

[255] CMS Collaboration, "Measurement of the t-channel single top-quark cross section at 13 TeV," CMS Physics Analysis Summary CMS-PAS-TOP-15-004, CERN, Geneva, 2015.

[256] M. Feindt and U. Kerzel, "The NeuroBayes neural network package," *Nucl. Instrum. Meth. A*, vol. 559, p. 190, 2006.

[257] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2," CMS Physics Analysis Summary CMS-PAS-BTV-15-001, CERN, Geneva, 2016.

[258] R. J. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples," *Comput. Phys. Commun.*, vol. 77, pp. 219–228, 1993.

[259] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions," *JHEP*, vol. 09, p. 111, 2009. [Erratum: JHEP02,011(2010)].

[260] A. Kalogeropoulos and J. Alwall, "The SysCalc code: A tool to derive theoretical systematic uncertainties," *arXiv:1801.08401 [hep-ph]*, 2018.

[261] V. Khachatryan *et al.*, "Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV," *Eur. Phys. J. C*, vol. 75, no. 11, p. 542, 2015.

[262] R. Barlow, "Asymmetric systematic errors," *arXiv:0306138 [physics]*, 2003.

[263] A. M. Sirunyan *et al.*, "Cross section measurement of t-channel single top quark production in pp collisions at $\sqrt{s}$ = 13 TeV," *Phys. Lett. B*, vol. 772, pp. 752–776, 2017.

[264] M. Aaboud *et al.*, "Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *JHEP*, vol. 04, p. 086, 2017.

[265] CMS Collaboration, "Measurement of the single top quark and antiquark production cross sections in the t channel and their ratio in pp collisions at sqrt(s)=13 TeV," CMS Physics Analysis Summary CMS-PAS-TOP-17-011, CERN, Geneva, 2018.

[266] A. M. Sirunyan *et al.*, "Measurement of the top quark mass using single top quark events in proton-proton collisions at $\sqrt{s}$ = 8 TeV," *Eur. Phys. J. C*, vol. 77, no. 5, p. 354, 2017.

[267] G. Mahlon and S. J. Parke, "Improved spin basis for angular correlation studies in single top quark production at the Tevatron," *Phys. Rev. D*, vol. 55, pp. 7249–7254, 1997.

[268] J. A. Aguilar-Saavedra and J. Bernabeu, "Breaking down the entire W boson spin observables from its decay," *Phys. Rev. D*, vol. 93, no. 1, p. 011301, 2016.

[269] V. Khachatryan *et al.*, "Measurement of top quark polarisation in t-channel single top quark production," *JHEP*, vol. 04, p. 073, 2016.

[270] V. Khachatryan *et al.*, "Search for anomalous Wtb couplings and flavour-changing neutral currents in t-channel single top quark production in pp collisions at $\sqrt{s}$ = 7 and 8 TeV," *JHEP*, vol. 02, p. 028, 2017.

[271] M. Aaboud *et al.*, "Probing the W tb vertex structure in t-channel single-top-quark production and decay in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector," *JHEP*, vol. 04, p. 124, 2017.

[272] M. Aaboud *et al.*, "Analysis of the Wtb vertex from the measurement of triple-differential angular decay rates of single top quarks produced in the t-channel at $\sqrt{s}$ = 8 TeV with the ATLAS detector," *JHEP*, vol. 12, p. 017, 2017.

[273] K. Cheung, K. Mawatari, E. Senaha, P.-Y. Tseng, and T.-C. Yuan, "The Top Window for dark matter," *JHEP*, vol. 10, p. 081, 2010.

[274] T. Lin, E. W. Kolb, and L.-T. Wang, "Probing dark matter couplings to top and bottom quarks at the LHC," *Phys. Rev. D*, vol. 88, no. 6, p. 063510, 2013.

[275] C. Arina *et al.*, "A comprehensive approach to dark matter studies: exploration of simplified top-philic models," *JHEP*, vol. 11, p. 111, 2016.

[276] D. Pinna, A. Zucchetta, M. R. Buckley, and F. Canelli, "Single top quarks and dark matter," *Phys. Rev. D*, vol. 96, no. 3, p. 035031, 2017.

[277] T. Plehn, J. Thompson, and S. Westhoff, "Dark Matter from Electroweak Single Top Production," *arXiv:1712.08065 [hep-ph]*, 2017.

[278] A. V. Baskakov, E. E. Boos, L. V. Dudko, I. P. Lokhtin, and A. M. Snigirev, "Single top quark production in heavy ion collisions at energies available at the CERN Large Hadron Collider," *Phys. Rev. C*, vol. 92, no. 4, p. 044901, 2015.

[279] A. M. Sirunyan *et al.*, "Observation of top quark production in proton-nucleus collisions," *Phys. Rev. Lett.*, vol. 119, no. 24, p. 242001, 2017.

[280] M. Aaboud *et al.*, "Measurement of the cross-section for producing a W boson in association with a single top quark in pp collisions at $\sqrt{s}$ = 13 TeV with ATLAS," *JHEP*, vol. 01, p. 063, 2018.

[281] M. Aaboud *et al.*, "Measurement of differential cross-sections of a single top quark produced in association with a W boson at $\sqrt{s}$ = 13 TeV with ATLAS," *Eur. Phys. J. C*, vol. 78, no. 3, p. 186, 2018.

[282] A. M. Sirunyan *et al.*, "Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s}$ = 13 TeV," *arXiv:1805.07399 [hep-ex]*, 2018.

[283] V. M. Abazov *et al.*, "Evidence for s-channel single top quark production in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV," *Phys. Lett. B*, vol. 726, pp. 656–664, 2013.

[284] T. A. Aaltonen *et al.*, "Observation of s-channel production of single top quarks at the Tevatron," *Phys. Rev. Lett.*, vol. 112, p. 231803, 2014.

[285] V. Khachatryan *et al.*, "Search for s channel single top quark production in pp collisions at $\sqrt{s}$ = 7 and 8 TeV," *JHEP*, vol. 09, p. 027, 2016.

[286] G. Aad *et al.*, "Search for s-channel single top-quark production in proton–proton collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector," *Phys. Lett. B*, vol. 740, pp. 118–136, 2015.

[287] G. Aad *et al.*, "Evidence for single top-quark production in the s-channel in proton-proton collisions at $\sqrt{s}$ =8 TeV with the ATLAS detector using the Matrix Element Method," *Phys. Lett. B*, vol. 756, pp. 228–246, 2016.

[288] F. Schenkel, "Search for s-channel Single Top Quark Production at a Center-of-Mass Energy of 13 TeV with the CMS Experiment," Master's thesis, Karlsruhe Institute of Technology (KIT), 2017.

[289] D. Seith, "Search for the associated production of a Higgs boson with a single top quark and trigger development for the CMS experiment," Master's thesis, Karlsruhe Institute of Technology (KIT), 2018.

[290] CMS Collaboration, "Work in progress (evolving results) and dedicated studies from PAGs." https://twiki.cern.ch/twiki/bin/view/CMS/MuonWorkInProgressAndPagResults, Feb. 2018.

[291] CMS Collaboration, "Reference muon id, isolation and trigger efficiencies for 2017 data." https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2017, Apr. 2018.

[292] P. Ott, "Inclusive cross section measurement for the production of single top quarks in the t channel at 13 TeV with the CMS experiment," Master's thesis, Karlsruhe Institute of Technology (KIT), 2017.

[293] CMS Collaboration, "Electron Tag-and-Probe." https://twiki.cern.ch/twiki/bin/view/CMSPublic/ElectronTagAndProbe, Mar. 2018.

[294] CMS Collaboration, "Usage of b/c Tag Objects for 13 TeV Data in 2016 and 80X MC ." https://twiki.cern.ch/twiki/bin/view/CMS/BtagRecommendation80X, Aug. 2016.

[295] CMS Collaboration, "Heavy flavour tagging for 13 TeV data in 2017 and 94X MC ." https://twiki.cern.ch/twiki/bin/view/CMS/BtagRecommendation94X, Feb. 2018.

[296] G. C. Fox and S. Wolfram, "Observables for the analysis of event shapes in e^+e^- annihilation and other processes," *Phys. Rev. Lett.*, vol. 41, pp. 1581–1585, Dec 1978.

[297] R. D. Field, Y. Kanev, and M. Tayebnejad, "A Topological analysis of the top quark signal and background at hadron colliders," *Phys. Rev. D*, vol. 55, pp. 5685–5697, 1997.

[298] CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period," CMS Physics Analysis Summary CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

[299] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s}$ = 13 TeV," CMS Physics Analysis Summary CMS-PAS-LUM-17-004, CERN, Geneva, 2018.

[300] J. S. Conway, "Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra," in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, pp. 115–120, 2011.

[301] CMS Collaboration, "Measurement of inclusive W and Z boson production cross sections in pp collisions at $\sqrt{s}$ = 13 TeV," CMS Physics Analysis Summary CMS-PAS-SMP-15-004, CERN, Geneva, 2015.

[302] V. Khachatryan *et al.*, "Measurement of the $t\bar{t}$ production cross section using events in the eμ final state in pp collisions at $\sqrt{s}$ = 13 TeV," *Eur. Phys. J. C*, vol. 77, p. 172, 2017.

[303] CMS Collaboration, "CMS Top Quark Physics Group." https://twiki.cern.ch/twiki/bin/view/CMS/TWikiTopQuark, Mar. 2018.

[304] CMS Collaboration, "Investigations of the impact of the parton shower tuning in Pythia 8 in the modelling of $t\bar{t}$ at $\sqrt{s} = 8$ and 13 TeV," CMS Physics Analysis Summary CMS-PAS-TOP-16-021, CERN, Geneva, 2016.

[305] CMS Collaboration, "TOP Systematic Uncertainties (Run2)." `https://twiki.cern.ch/twiki/bin/view/CMS/TopSystematics`, Feb. 2018.

[306] G. Altarelli, B. Mele, and M. Ruiz-Altaba, "Searching for New Heavy Vector Bosons in $p\bar{p}$ Colliders," *Z. Phys. C*, vol. 45, p. 109, 1989. [Erratum: Z. Phys.C47,676(1990)].

[307] F. Huang, H.-L. Li, S.-Y. Li, Z.-G. Si, W. Su, and Z.-J. Yang, "Search for W′ signal in single top quark production at the LHC," *Chin. Phys. C*, vol. 42, no. 3, p. 033103, 2018.

[308] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, "Theory and phenomenology of two-Higgs-doublet models," *Phys. Rept.*, vol. 516, pp. 1–102, 2012.

[309] M. Hashemi, "Observability of Heavy Charged Higgs through s-channel Single Top Events at LHC," *JHEP*, vol. 11, p. 005, 2013.

[310] M. Hashemi, "Single Top Events as a Source of Light Charged Higgs in the Fully Hadronic Final State at LHC," *JHEP*, vol. 05, p. 112, 2013.

[311] I. Ahmed, M. Hashemi, and W. A. Tajuddin, "The *s*-channel Charged Higgs in the Fully Hadronic Final State at LHC," *Eur. Phys. J. C*, vol. 76, no. 4, p. 209, 2016.

[312] J. D'Hondt, A. Mariotti, K. Mawatari, S. Moortgat, P. Tziveloglou, and G. Van Onsem, "Signatures of top flavour-changing dark matter," *JHEP*, vol. 03, p. 060, 2016.

[313] CMS Collaboration, "Search for associated production of a single top quark and a Higgs boson in events where the Higgs boson decays to two photons at $\sqrt{s} = 8$ TeV," CMS Physics Analysis Summary CMS-PAS-HIG-14-001, CERN, Geneva, 2014.

[314] CMS Collaboration, "Search for H to bbbar in association with single top quarks as a test of Higgs couplings," CMS Physics Analysis Summary CMS-PAS-HIG-14-015, CERN, Geneva, 2014.

[315] CMS Collaboration, "Search for Associated Production of a Single Top Quark and a Higgs Boson in Leptonic Channels," CMS Physics Analysis Summary CMS-PAS-HIG-14-026, CERN, Geneva, 2015.

[316] CMS Collaboration, "Search for H to bbar in association with a single top quark as a test of Higgs boson couplings at 13 TeV," CMS Physics Analysis Summary CMS-PAS-HIG-16-019, CERN, Geneva, 2016.

[317] CMS Collaboration, "Modelling of the single top-quark production in association with the Higgs boson at 13 TeV." `https://twiki.cern.ch/twiki/bin/viewauth/CMS/SingleTopHiggsGeneration13TeV`, Feb. 2016.

[318] H. B. Hartanto, B. Jager, L. Reina, and D. Wackeroth, "Higgs boson production in association with top quarks in the POWHEG BOX," *Phys. Rev. D*, vol. 91, no. 9, p. 094003, 2015.

[319] CMS Collaboration, "Luminosity Physics Object Group." `https://twiki.cern.ch/twiki/bin/view/CMS/TWikiLUM`, Feb. 2018.

[320] CMS Collaboration, "Event reweighting using scale factors calculated with a tag and probe method." `https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagShapeCalibration`, Jan. 2018.

[321] CMS Collaboration, "Calibration of the Combined Secondary Vertex b-Tagging discriminant using dileptonic ttbar and Drell-Yan events," CMS Analysis Note CMS AN-2013/130, CERN, Geneva, 2013.

[322] J. Rauser, "Search for higgs boson production in association with a single top quark with the cms experiment," Master's thesis, Karlsruhe Institute of Technology (KIT), 2017.

[323] K. Flöh, "Search for single top quarks in association with a Higgs boson, $H - bbar$, at a center-of-mass energy of 13 TeV with the CMS detector," Master's thesis, Karlsruhe Institute of Technology (KIT), 2017.

[324] V. D. Barger, J. Ohnemus, and R. J. N. Phillips, "Event shape criteria for single lepton top signals," *Phys. Rev. D*, vol. 48, pp. R3953–R3956, 1993.

[325] D. Müller, "Search for Higgs Boson Production in Association with a Single Top Quark as a Test of Higgs Boson Couplings at 13 TeV with the CMS Experiment," Master's thesis, Karlsruhe Institute of Technology (KIT), 2017.

[326] CMS Collaboration, "Jet energy scale uncertainty sources." `https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources`, Nov. 2017.

[327] CMS Collaboration, "Search for production of a Higgs boson and a single top quark in multilepton final states in proton collisions at $\sqrt{s} = 13$ TeV," CMS Physics Analysis Summary CMS-PAS-HIG-17-005, CERN, Geneva, 2017.

[328] A. M. Sirunyan *et al.*, "Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV," *arXiv:1804.02716 [hep-ex]*, 2018.

[329] CMS Collaboration, "Search for the tH(H$\rightarrow$ b$\bar{\text{b}}$) process in pp collisions at $\sqrt{s} = 13$ TeV and study of Higgs boson couplings," CMS Physics Analysis Summary CMS-PAS-HIG-17-016, CERN, Geneva, 2018.

[330] G. Aad *et al.*, "Search for $H \rightarrow \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector," *Phys. Lett. B*, vol. 740, pp. 222–242, 2015.

[331] CMS Collaboration, "b-jet energy regression at 13 TeV." `https://twiki.cern.ch/twiki/bin/view/Main/BJetRegression`, May 2018.

[332] M. Erdmann, B. Fischer, and M. Rieger, "Jet-parton assignment in $t\bar{t}$H events using deep learning," *JINST*, vol. 12, no. 08, p. P08020, 2017.

[333] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.

[334] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[335] G. Louppe, K. Cho, C. Becot, and K. Cranmer, "QCD-Aware Recursive Neural Networks for Jet Physics," *arXiv:1702.00748 [hep-ph]*, 2017.

[336] CERN, "The High-Luminosity LHC." `https://home.cern/topics/high-luminosity-lhc`, May 2018.

[337] CERN, "LHC Commissioning with beam." `https://lhc-commissioning.web.cern.ch/lhc-commissioning/`, May 2018.

[338] G. Apollinari, I. Béjar Alonso, O. Brüning, M. Lamont, and L. Rossi, "High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report," Technical Design Report, CERN, Geneva, 2015.

[339] E. Boos, Y. Kurihara, Y. Shimizu, M. Sachwitz, H. J. Schreiber, and S. Shichanin, "Top quark production in the reaction $e^+e^- \to e$ neutrino tb at linear collider energies," *Z. Phys. C*, vol. 70, pp. 255–262, 1996.

[340] A. Escamilla, A. O. Bouzas, and F. Larios, "Single top production at linear e^-e^+ colliders," *Phys. Rev. D*, vol. 97, no. 3, p. 033004, 2018.

[341] V. Barger, E. Basso, Y. Gao, and W.-Y. Keung, "Neutrino signals in IceCube from weak production of top and charm quarks," *Phys. Rev. D*, vol. 95, no. 9, p. 093002, 2017.

[342] CMS Collaboration, "Summary table of samples produced for the 1 billion campaign, with 25ns bunch-crossing." `https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns`, Apr. 2017.

[343] LHC Higgs Cross Section Working Group, "Higgs cross sections and decay branching ratios." `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG`, Nov. 2017.

[344] F. Maltoni, D. Pagani, and I. Tsinikos, "Associated production of a top-quark pair with vector bosons at NLO in QCD: impact on tt̄H searches at the LHC," *Journal of High Energy Physics*, vol. 02, p. 113, 2016.

[345] J. Campbell, R. K. Ellis, and R. Röntsch, "Single top production in association with a Z boson at the LHC," *Phys. Rev. D*, vol. 87, p. 114006, 2013.

Danksagung

Zunächst möchte ich Prof. Dr. Thomas Müller dafür danken, dass er mich zum Ende meines Studiums in seine Arbeitsgruppe aufgenommen hat und mir ermöglicht hat, in diesem Forschungsbereich zu promovieren. Ich konnte jederzeit auf seine Unterstützung zählen bei der Präsentation meiner Arbeit auf Konferenzen sowie bei der Weiterbildung auf Schulen. Darüber hinaus ließ er mir stets genug Freiheiten um meinen eigenen Forschungsinteressen nachzugehen.

Ich möchte außerdem Prof. Dr. Ulrich Husemann danken für die Übernahme des Korreferats, sowie für die zahlreichen Kommentare und Verbesserungsvorschläge über die vergangenen Jahre. Der regelmäßige Austausch der verschiedenen Analysegruppen innerhalb des Instituts sorgte für konstantes, konstruktives Feedback.

Mein besonderer Dank gilt meinem Betreuer Dr. Thorsten Chwalek, welcher mich ebenfalls seit dem ersten Tag am Institut bzw. der Arbeitsgruppe begleitet hat. Er hatte stets ein offenes Ohr für meine Sorgen und fand stets die richtigen Worte, sei es zur Motivation oder zum Inhalt.

Den aktuellen Mitgliedern meiner Arbeitsgruppe, Darius Bühler, Kevin Flöh, Denise Müller, Johann Rauser und David Seith möchte ich ganz herzlich für die entspannte, aber dennoch konstruktive Arbeitsatmosphäre danken, sowie dafür dass sie mir in den letzten Monaten den Rücken frei gehalten haben.

Ebenfalls danken möchte ich Dr. Matthias Mozer für hilfreiche Anregungen zu den Analysen in den letzten Jahren und insbesondere für physikalische Fragestellungen im Hinblick auf die Prüfung. Darüber hinaus möchte ich noch allen zuvor genannten Personen zusätzlich danken für das intensive Korrekturlesen dieser Doktorarbeit.

Ein großer Dank geht auch an meine ehemaligen Kollegen Dr. Chrstian Böser, Dr. Simon Fink, Dr. Benedikt Maier und Dr. Frank Roscher, die maßgeblich zu meiner Ausbildung beigetragen haben und von deren Wissen ich enorm profitiert habe. Danken möchte ich auch Philipp Ott und Fabian Schenkel, die ich als Masterstudenten betreuen durfte, für ihren Enthusiasmus und Eifer, den sie an den Tag gelegt haben.

In addition, I would also like to thank all people in the CMS Collaboration that contributed to this work, especially the people in the TOP subgroup, involving former and current conveners and colleagues alike.

Generell möchte ich weiterhin allen Mitgliedern des Instituts für Experimentelle Teilchenphysik (ETP, ehemals IEKP, bzw. EKP) danken für die angenehme Atmosphäre. Hervorzuheben ist hier

im Speziellen Frau Bräunling, ohne deren Hilfe der bürokratische und organisatorische Aufwand für jeden von uns unweit höher wäre. Ebenso möchte ich mich bei dem Admin-Team bedanken für einen reibungslosen Ablauf der täglichen Arbeit bezüglich unserer IT-Infrastruktur.

Zuletzt möchte ich mich bei meinen Eltern bedanken für ihre Unterstützung und das Vertrauen, welches sie in mich gesetzt haben. Ohne sie wäre das alles nicht möglich gewesen.