ETP-KA/2020-3

Karlsruher Institut für Technologie

Investigation of the Top Quark Yukawa Coupling in Higgs Boson Production in Association with Top Quarks at 13 TeV with the CMS Experiment

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

genehmigte

DISSERTATION

von

M.Sc. Kevin Marcel Flöh
aus Karlsruhe

Mündliche Prüfung: 24. Januar 2020

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

Korreferent: Prof. Dr. Günter Quast
Institut für Experimentelle Teilchenphysik

Introduction

In 1964, Robert Brout, François Englert, and Peter Higgs proposed a mechanism that is now used to introduce gauge boson masses into the standard model of particle physics without violating local gauge invariance. This mechanism is associated with a particle, the Higgs boson. It couples to all massive fermions proportionally to their mass and to the massive bosons in proportion to their mass squared.

The properties of the elementary particles are studied with several types of experiments: astroparticle, fixed target or collider experiments. For the latter, man-made accelerators are built, which allow to perform experiments. The Large Hadron Collider at CERN is one of these accelerators. It is equipped with the ALICE, ATLAS, CMS and LHCb experiments. Following a long search, the Higgs boson was finally discovered by the ATLAS and CMS collaborations in 2012 as the last missing particle of the standard model. After its discovery, the properties of the Higgs boson are studied further to look for deviations from the standard model.

In this thesis, the coupling of the Higgs boson to the heaviest particle in the standard model, namely the top quark, is investigated. This coupling is called top quark Yukawa coupling. It can be studied in the production and decay of the Higgs boson. The Higgs boson cannot decay directly into top quarks due to their high mass. Hence, the decay only allows to access the coupling indirectly via quantum loops. The direct measurement of the coupling to the top quark is only possible in production processes. The Higgs boson production in association with a top quark pair gives direct access to the amplitude of the top quark Yukawa coupling. Furthermore, the Higgs boson production in association with a single top quark is sensitive to the amplitude and relative sign of the top quark Yukawa coupling and the coupling of the Higgs boson to vector bosons. Therefore, these two processes are studied within this thesis to constrain the Higgs boson couplings. Additionally, limits on a possible CP-mixing in the Higgs boson coupling are derived. Finally, upper limits on the signal strength of the standard-model-like single top quark production in association with a Higgs boson are set.

An analysis dedicated to single top quark production in association with a Higgs boson using the data taken in 2016 by the CMS experiment is presented. The analysis has been carried out in the $H \to b\bar{b}$ decay channel within the context of this thesis and is published in Ref. [1]. A combination with the multilepton and diphoton decay channels is presented in Section 6.9 and published in Ref. [2]. Since the publication, the $H \to b\bar{b}$ analysis is furthermore improved compared to the publication by simulating additional signal events, introducing a two-fold training of the machine learning methods, and introducing a DNN-based energy regression for jets originating from bottom quarks. Furthermore, this analysis provides upper limits on a possible CP-mixing in the Higgs boson coupling. In this thesis, the modification of the kinematic properties of the top quark pair production in association with a Higgs boson due to the possible CP-mixing is taken into account for the first time. Based on this improved analysis, a combined analysis of the Higgs boson production in association with single top quarks or in association with top quark pairs is performed.

The theoretical foundations of this thesis are presented in Chapter 1 and the experimental setup is explained in Chapter 2. The statistical and machine learning methods are introduced in Chapter 3. In Chapter 4, the event simulation and reconstruction is described. The common prerequisites for the two analyses are described in Chapter 5. This includes a description of signal and background processes and the jet assignment procedure, which was developed especially for the search for single top quark production in association with a Higgs boson. The two analyses are presented in Chapters 6 and 7. The thesis is summarized in Chapter 8, where also an outlook to possible future improvements is given.

Contents

1. Theory 1
   - 1.1. The Standard Model of Particle Physics . . . 1
     - 1.1.1. Fermions . . . 2
     - 1.1.2. Bosons . . . 3
   - 1.2. Theoretical Concepts . . . 4
     - 1.2.1. Electromagnetic Interaction . . . 4
     - 1.2.2. Weak Interaction . . . 5
     - 1.2.3. Electroweak Symmetry Breaking . . . 6
     - 1.2.4. Strong Interaction . . . 7
   - 1.3. Partons versus Particles . . . 8
   - 1.4. The Top Quark . . . 9
     - 1.4.1. Top Quark Production . . . 9
     - 1.4.2. Top Quark Decays . . . 11
   - 1.5. The Higgs Boson . . . 11
     - 1.5.1. Higgs Boson Production . . . 12
     - 1.5.2. Associated Production with a Top Quark Pair . . . 14
     - 1.5.3. Associated Production with a Single Top Quark . . . 15
     - 1.5.4. Higgs Boson Decays . . . 16

2. The Large Hadron Collider and the Compact Muon Solenoid Experiment 19
   - 2.1. The Large Hadron Collider . . . 19
   - 2.2. The Compact Muon Solenoid Experiment . . . 22
     - 2.2.1. Coordinate System . . . 22
     - 2.2.2. Tracking System . . . 24
     - 2.2.3. Electromagnetic Calorimeter . . . 26
     - 2.2.4. Hadron Calorimeter . . . 28
     - 2.2.5. Solenoid Magnet . . . 29
     - 2.2.6. Muon System . . . 29
     - 2.2.7. Trigger and Data Acquisition System . . . 30
     - 2.2.8. Computing . . . 31

3. Statistical Methods and Machine Learning 33
   - 3.1. Probability Density Function . . . 33
   - 3.2. Maximum Likelihood Method . . . 34
   - 3.3. Systematic Uncertainties . . . 34
   - 3.4. Exclusion Limit Calculation . . . 35
   - 3.5. Boosted Decision Trees . . . 36
   - 3.6. Artificial Neural Networks . . . 38
   - 3.7. Validation of Machine Learning Methods . . . 42

4. Simulation and Reconstruction of Events 43
   - 4.1. Event Simulation . . . 43
     - 4.1.1. Proton-Proton Scattering Process . . . 43
     - 4.1.2. Parton Shower and Hadronization . . . 44
     - 4.1.3. Underlying Event and Pileup . . . 45

4.1.4. Monte Carlo Event Generators . . . 45
4.1.5. Detector Simulation . . . 46
4.2. Event Reconstruction . . . 46
4.2.1. Particle Tracks and Primary Vertex Reconstruction . . . 47
4.2.2. Muon Reconstruction . . . 47
4.2.3. Electron Reconstruction . . . 48
4.2.4. Photon and Hadron Reconstruction . . . 48
4.2.5. Jet Reconstruction . . . 49
4.2.6. Missing Transverse Momentum . . . 54
4.2.7. W Boson Reconstruction . . . 54

5. Event Topology and Jet Assignment 57
5.1. Signal Processes . . . 57
5.1.1. t-channel Single Top Quark Production in Association with a Higgs Boson . . . 58
5.1.2. tW Single Top Quark Production in Association with a Higgs Boson . . . 58
5.1.3. Top Quark Pair Production in Association with a Higgs Boson . . . 58
5.2. Background Processes . . . 60
5.2.1. Top Quark Pair Production . . . 60
5.2.2. Single Top Quark Production . . . 61
5.2.3. Minor Background Processes . . . 61
5.3. Jet Assignment . . . 62
5.3.1. tHq Hypothesis . . . 64
5.3.2. tHW Hypothesis . . . 64
5.3.3. $\mathrm{t\bar{t}H}$ Hypothesis . . . 66
5.3.4. $\mathrm{t\bar{t}}$ Hypothesis . . . 66
5.3.5. Evaluation . . . 66

6. Search for tH with $\mathrm{H} \to \mathrm{b\bar{b}}$ with the 2016 Data Set 71
6.1. Search Strategy . . . 71
6.2. b-jet Energy Regression . . . 72
6.3. Event Selection . . . 74
6.4. Corrections of Simulated Events . . . 79
6.5. Uncertainty Treatment . . . 81
6.6. Jet Assignment . . . 83
6.7. Classification of Events . . . 89
6.7.1. Signal Classification . . . 89
6.7.2. Flavor Classification . . . 94
6.8. Results of the Higgs Boson to $\mathrm{b\bar{b}}$ Decay Channel . . . 97
6.9. Combined Measurement of Different Higgs Boson Decay Channels . . . 101
6.9.1. $\mathrm{b\bar{b}}$ Channel . . . 101
6.9.2. Multilepton Channel . . . 103
6.9.3. Diphoton Channel . . . 103
6.9.4. Results . . . 104
6.10. Summary and Outlook . . . 106

7. Investigation of Higgs Boson Couplings in tH and $\mathrm{t\bar{t}H}$ Events with $\mathrm{H} \to \mathrm{b\bar{b}}$ with the 2016–2018 Data Set 107
7.1. Analysis Strategy . . . 107
7.2. Event Selection . . . 107
7.3. Systematic Uncertainties . . . 109
7.4. Jet Assignment . . . 109
7.5. Event Classification . . . 116

7.6. Expected Results 117

8. Conclusion and Outlook 119

A. Search for tH production 121
A.1. Cross Sections of Signal Processes 121
A.2. Kinematic Variations of Signal Processes 124
A.3. Input Variables of the Jet Assignment 127
A.4. Ranking of Input Variables of the Jet Assignment 131
A.5. Jet Assignment Efficiencies 133
A.6. Input Variables for the Flavor Classification BDT 134
A.7. Upper Limits on Cross Section Times Branching Ratio 136
A.8. Systematic Impacts 139

B. Combined tH and $t\bar{t}$H analysis 141
B.1. Input Variables of the Jet Assignment 141
B.2. Ranking of Input Variables to the Jet Assignment 150
B.3. Jet Assignment Efficiencies 154
B.4. Overtraining Check of JA BDTs 156
B.5. DNN Training 161
B.6. Systematic Impacts 162

1. Theory

The properties of elementary particles and their interactions can be described by a quantum field theory called the standard model of particle physics (SM). Built on the foundations laid in 1925, when Max Born, Werner Heisenberg and Pascual Jordan developed their theory of quantum mechanics [3], the SM was established in the 1960s. It is not only able to describe the particles known as of the time of its introduction but it was also able to predict particles that were not known by that time. These particle were discovered later on, which includes the charm quark (1974) [4, 5], the bottom quark (1977) [6], the top quark (1995) [7, 8], the tau neutrino (2000) [9], and as the last missing particle, the Higgs boson (2012) [10, 11].

With the discovery of the Higgs boson all particles described by the SM have been observed. Not only the Higgs boson itself but also its prevalent production and decay modes have been observed to be compatible to the prediction of the SM, e.g., the associated production with a pair of top quarks [12] or the decay into a pair of bottom quarks [13, 14]. Since the discovery, the focus shifted from the search for the Higgs boson to precision measurements and rare production and decay channels in the Higgs sector.

The electromagnetic, the weak and the strong force are compatible with SM prediction as well. Although the SM describes a lot of physical phenomena with high precision, there are several aspects of physics that cannot be addressed by the SM properly. In cosmological observations a deviation between the predicted and measured rotational velocity of stars far away from the center of their galaxy is observed [15]. This discrepancy can be solved by introducing dark matter. Since the SM does not provide any suitable dark matter candidates, this is a hint for physics beyond the SM (BSM). Theories like supersymmetry (SUSY) add additional particles to the SM, which could account for dark matter. Many searches have been performed but no evidence for SUSY has been found so far and a large fraction of the available parameter space has already been excluded [16].

Furthermore, the SM does not include a description of gravity. This is, however, not too severe since the interactions of particles happen on a subatomic level, where gravity is negligible due to its relative weakness compared to the other forces on this scale. Gravity can be described by Einstein's theory of general relativity [17] on large scales. In order to be able to describe states with large mass and high density, e.g., black holes, various approaches have been made to combine general relativity and the SM. Unfortunately, none of these theories lead to falsifiable predictions.

The interactions and particles of the SM are presented in this chapter. A focus lies on the properties of the top quark and the Higgs boson and on the interactions between them. All descriptions are made considering the conditions of the Large Hadron Collider (LHC) at CERN, Geneva, Switzerland. Natural units, i.e., $\hbar = c = 1$, are used in this chapter.

1.1. The Standard Model of Particle Physics

The SM describes six quarks, three charged leptons and their corresponding neutrinos, four gauge bosons and the Higgs boson. The quarks, charged leptons, and neutrinos all have half-integer spin, therefore they are grouped into one class of particles called fermions, see Table 1.1. The particles interact in several ways: all charged particles interact via the electromagnetic (EM) force mediated by the photon. The weak force acts on all left-handed and on the massive right-handed particles of the SM via the W and Z bosons. Furthermore, the gluons interact with the quarks, mediating the strong force. The mediators of these forces all have spin 1 and are referred to as gauge bosons. The particles of

Table 1.1.: Fermions of the SM: there are six quarks, three charged leptons and their corresponding neutrinos. The fermions are grouped into three generations each generation containing two quarks and two leptons. The electric charge is given in multiples of the elementary charge $e = 1.602 \times 10^{-19} C$ [18].

	Generation			Electric	Third Component
	1	2	3	Charge (e)	of Weak Isospin
quarks	up (u)	charm (c)	top (t)	$+\frac{2}{3}$	$+\frac{1}{2}$
	down (d)	strange (s)	bottom (b)	$-\frac{1}{3}$	$-\frac{1}{2}$
leptons	ν_e	ν_μ	ν_τ	0	$+\frac{1}{2}$
	electron (e)	muon (μ)	tauon (τ)	-1	$-\frac{1}{2}$

the SM gain their masses through the Higgs mechanism, which adds another particle, the Higgs boson. The Higgs boson has spin 0 and is the only scalar elementary particle in the SM. The properties of the bosons are summarized in Table 1.2.

1.1.1. Fermions

The SM introduces twelve fermions, which can be grouped into quarks and leptons. Furthermore, the quarks and leptons can be divided into three generations, where each generation consists of the left-handed isospin doublets

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \tag{1.1}$$

for the quarks and

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \tag{1.2}$$

for the leptons. The corresponding right-handed singlets are denoted as

$$\left(u\right)_R, \left(d\right)_R, \left(c\right)_R, \left(s\right)_R, \left(t\right)_R, \left(b\right)_R \tag{1.3}$$

for the quarks and

$$\left(e\right)_R, \left(\mu\right)_R, \left(\tau\right)_R \tag{1.4}$$

for the leptons. The SM does not contain right-handed neutrinos. The quarks can be further grouped by their electric charge into up-type (u, c, t) and down-type (d, s, b) quarks. The up-type quarks posses a charge of $+\frac{2}{3}e$, whereas the down-type quarks carry $-\frac{1}{3}e$. Similarly, the leptons can be grouped into charged leptons (e, μ, τ) and neutrinos (ν_e, ν_μ, ν_τ). For each of the fermions, an antifermion exists with opposite charges and the same mass. All visible matter that surrounds us consists of up and down quarks as well as electrons. All other particles can only be accessed via particle or astroparticle physics experiments.

The quarks are all massive with masses ranging from $2.2^{+0.5}_{-0.4}$ MeV (up quark) to 173.0 ± 0.4 GeV (top quark) [18]. These masses are related to the Yukawa coupling as described in Section 1.2.3. In contrast to the leptons, the quarks carry color charge, often referred to as red, green or blue. Quarks can only be observed in bound states due to confinement, which is explained in Section 1.2. These bound states have to be colorless. Quarks interact via all three forces as they carry color, weak and electromagnetic charge.

In the SM, only charged leptons are massive with masses reaching from 511 keV (electron) to 1.78 GeV (tauon) [18]. The neutrinos are considered as massless. Nonetheless, there are hints that this assumption is not correct. The observation of neutrino oscillations between different generations when propagating

Table 1.2.: Bosons of the SM: the mediators of the interactions of the SM are listed here. The electric charge is given in multiples of the elementary charge $e = 1.602 \times 10^{-19} C$, J is the spin and P is the eigenvalue of the parity operator [18].

	Force	Coupling to	Mass (GeV)	Electric Charge (e)	J^P
photon (γ)	EM	electric charge	0	0	1^-
$W^\pm$	weak	weak isospin	80.4	± 1	1^+
Z^0			91.2	0	1^+
gluons (g)	strong	color charge	0	0	1^-
Higgs (H)		mass	125	0	0^+

freely can only be explained if they have different masses. Therefore, at least two neutrinos have to be massive [19]. Due to the absence of electric and color charge, neutrinos only interact weakly, which makes them hardly detectable. The charged leptons interact electromagnetically and weakly.

1.1.2. Bosons

The interactions between fermions are mediated by bosons. In general, the effective range of each force is inversely proportional to the squared mass of the mediating gauge boson. However, self-interactions of the gauge bosons can limit the range of a force even further.

The electromagnetic force, which is the only force of the SM present at large distances, is mediated by the massless photon. The photon couples to the electric charge but does not carry electric charge itself. Since there is no self-interaction and the photon is massless, the range of the electromagnetic force is infinite. In order to describe the interactions with photons, the theory of quantum electrodynamics (QED) has been developed [20].

The weak interaction, mediated by the W and Z bosons, describes transitions between different lepton and quark flavors and scattering processes. All three gauge bosons of the weak force are massive. As the coupling is suppressed by the mass of respective the gauge boson m^{-2} at leading order, the range is limited to small length scales. This results in a relative weakness compared to other forces at low energy scales, which gives the force its name. The weak and the electromagnetic interactions are combined into the electroweak interaction, which describes both forces and includes interference effects between electromagnetic and weak processes [21–23]. The interference effects become relevant at energy scales larger than the Z boson mass. Since the W and Z bosons carry a weak charge themselves, self-interactions between them are possible.

The SM contains eight different gluons, where each gluon is a linear combination of a color-anticolor state, with the three colors blue, green, red and the corresponding anticolors. The gluons therefore carry (anti)color themselves, consequently, they interact not only with quarks but also with each other. This leads to the limitation of the range of the force to small distances. The separation of two color-charged particles, i.e., quarks, connected by the strong force require higher energies the further away they are from each other. If the binding energy becomes large enough, new particles are generated leading to new colorless bound states via a quantum effect called hadronization. Thus, free quarks cannot be observed, this effect is called confinement. For very small distances between two quarks, the strong force gets asymptotically small and the two particles can be considered as moving freely. This is named asymptotic freedom. The strong force is described by the theory of quantum chromodynamics (QCD) [24].

1.2. Theoretical Concepts

All particles in the SM are described as quantum states of fields ϕ. Only these states are observable by experiment. The dynamics of the quantum fields are derived by the principle of least interaction, which uses Lagrangian densities $\mathcal{L}$ to describe the particles and their interactions. The action S of a given system can be written as

$$S = \int \mathcal{L}(\phi, \partial_\mu \phi) \mathrm{d}^4 x, \tag{1.5}$$

with the four-vector x in spacetime and the spacetime derivative $\partial_\mu = \frac{\partial}{\partial x^\mu}$. In order to minimize the interaction, its variation has to vanish. Therefore, $\delta S = 0$ is required, which leads to the Euler-Lagrange equation:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \tag{1.6}$$

For a non-interacting fermion field ψ, the Langrangian density is given by

$$\mathcal{L}_{\text{Dirac}} = \overline{\psi}(\mathrm{i}\gamma^\mu \partial_\mu - m)\psi, \tag{1.7}$$

where ψ is the spinor, $\overline{\psi} = \psi^\dagger \gamma^0$ is the adjoint spinor, m is the mass of the particle and γ^μ are the gamma matrices. Inserting Equation (1.7) into Equation. (1.6) leads to the Dirac equation:

$$(\mathrm{i}\gamma^\mu \partial_\mu - m)\psi = 0. \tag{1.8}$$

The Lagrangian density for a scalar non-interacting boson field ϕ is given by:

$$\mathcal{L}_{\text{Klein-Gordon}} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2. \tag{1.9}$$

Interactions between particles are incorporated into the theory via local gauge invariance of the respective Langrangians.

1.2.1. Electromagnetic Interaction

At first, the electromagnetic interaction is described classically by two separate theories: one for magnetism and one for electrostatics [25]. The two theories are united into Maxwell's equations [26]. However, these neither include the quantization of the fields nor relativistic effects. A complete theory of electromagnetic interactions is given by QED [20]. It can be derived by requiring $\mathcal{L}$ to be symmetric under local U(1) gauge transformations which means that adding an arbitrary phase $\varphi(x)$ to the spinor does not change any physical observables

$$\psi(x) \mapsto \psi'(x) = \mathrm{e}^{\mathrm{i}q\varphi(x)}\psi(x), \tag{1.10}$$

where q is the electric charge. Applying this transformation, e.g., to $\mathcal{L}_{\text{Dirac}}$ leads to:

$$\mathcal{L}(\psi', \partial_\mu \psi') = \mathcal{L}(\psi, \partial_\mu \psi) + q\overline{\psi}\gamma^\mu \partial_\mu \varphi(x)\psi. \tag{1.11}$$

In order to recover local gauge invariance, the normal derivative is replaced with the covariant derivative

$$\partial_\mu \mapsto D_\mu = \partial_\mu - \mathrm{i}qA_\mu(x), \tag{1.12}$$

with the four-potential A_μ and its transformation:

$$A_\mu(x) \mapsto A'_\mu(x) = A_\mu(x) + \partial_\mu \varphi(x). \tag{1.13}$$

This leads to

$$\mathcal{L}_{\mathrm{QED}} = \overline{\psi}(\mathrm{i}\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.14}$$

$$= \overline{\psi}(\mathrm{i}\gamma^\mu \partial_\mu - m)\psi + q\overline{\psi}\gamma^\mu \psi A_\mu - \frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu), \tag{1.15}$$

where the first (last) term describes the propagation of a free fermion (photon) and the second term describes the interactions between fermions and photons. The field strength tensor is given by:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.16}$$

1.2.2. Weak Interaction

Fermi developed a theoretical description of the beta decay of the neutron [27]. It introduces a direct interaction between neutron, proton, electron and electron antineutrino. However, this approach violates unitarity for high energies. In order to avoid this, an additional particle is introduced, the W boson

$$W^\pm = \frac{1}{\sqrt{2}}(W^1 \mp \mathrm{i}W^2), \tag{1.17}$$

which is produced virtually in this case and mediates the weak interaction, with the gauge fields W^j_μ.

Similar to the electromagnetic interaction, a gauge group is associated to the weak interaction, namely SU(2). This enables the weak force to change the isospin of the particles it interacts with and therefore change their flavors. Hence, the particles are organized in isospin doublets, see Equations (1.1) and (1.2) and singlets, see Equations (1.3) and (1.4). As in QED, a covariant derivative is defined

$$D_\mu = \partial_\mu - \frac{1}{2}ig'B_\mu - \frac{1}{2}\mathrm{i}gW^j_\mu\sigma^j \tag{1.18}$$

$$= \partial_\mu - \frac{1}{2}ig'B_\mu - \frac{1}{2}\mathrm{i}g\begin{pmatrix} W^3_\mu & W^1_\mu - \mathrm{i}W^2_\mu \\ W^1_\mu + \mathrm{i}W^2_\mu & -W^3_\mu \end{pmatrix}, \tag{1.19}$$

where g is the weak charge and σ^j are the Pauli matrices. The gauge fields of the combined description of QED and weak theory are B_μ and W^j_μ. The charged current of the weak force is maximally parity violating, hence it only couples to one chiral eigenstate. In case of the weak force, these are the left-handed fermions and right-handed antifermions.

The weak force does not couple to the mass eigenstates but to rotations of them

$$\begin{pmatrix} \mathrm{u} \\ \mathrm{d}' \end{pmatrix}, \begin{pmatrix} \mathrm{c} \\ \mathrm{s}' \end{pmatrix}, \begin{pmatrix} \mathrm{t} \\ \mathrm{b}' \end{pmatrix}, \tag{1.20}$$

with the relation to the physical quarks given by

$$\begin{pmatrix} \mathrm{d}' \\ \mathrm{s}' \\ \mathrm{b}' \end{pmatrix} = \begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \\ \mathrm{b} \end{pmatrix}, \tag{1.21}$$

where the matrix linking the two is the Cabibbo-Kobayashi-Maskawa (CKM) matrix [28]. The magnitudes of the matrix elements are given by [18]:

$$|V_{\mathrm{CKM}}| = \begin{pmatrix} 0.97446 \pm 0.00010 & 0.22452 \pm 0.00044 & 0.00365 \pm 0.00012 \\ 0.22438 \pm 0.00044 & 0.97359^{+0.00010}_{-0.00011} & 0.04214 \pm 0.00076 \\ 0.00896^{+0.00024}_{-0.00023} & 0.04133 \pm 0.00074 & 0.999105 \pm 0.000032 \end{pmatrix}. \tag{1.22}$$

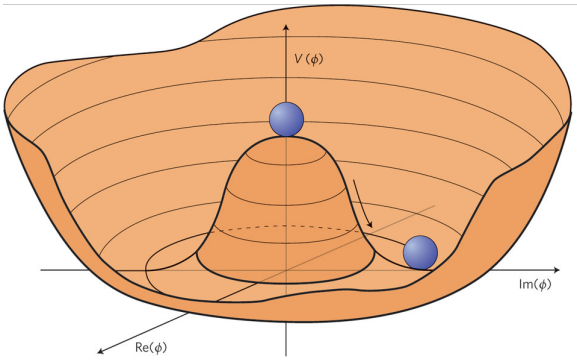

Figure 1.1.: Illustration of the Higgs potential split into real and imaginary components: the minima of the potential lie on a circle around the origin. Taken from Ref. [29].

In addition to the two W bosons given by Equation (1.17), there is also the Z boson mediating the neutral currents of the weak force. It cannot be simply identified with the field W_3 but has to be described as a linear combination between the W_3 and the B field

$$\begin{pmatrix} A \\ Z \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix}, \tag{1.23}$$

with the Weinberg angle $\sin^2\theta_W \approx 0.231$, the Z field representing the Z boson, and the field A representing the photon.

1.2.3. Electroweak Symmetry Breaking

The aforementioned theories consider the gauge bosons as well as the fermions as massless, although experimental observations prove these particles to be massive [18]. This issue can be solved by introducing the Higgs mechanism. Therefore, a Higgs field is added and the symmetry of its vacuum state is spontaneously broken, leading to massive gauge bosons of the weak interaction and massive fermions through the coupling of the Higgs field to the fermion fields. Furthermore, a new particle, the Higgs boson, is introduced. After the discovery of the Higgs boson in 2012, two of the authors, François Englert and Peter Higgs, were awarded the Nobel prize in physics.

Introducing the Higgs field

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{1.24}$$

and the corresponding term in the Lagrangian density

$$\mathcal{L}_H = (D^\mu\phi)^\dagger(D_\mu\phi) - V(\phi), \tag{1.25}$$

where

$$V(\phi) = -\mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2 \tag{1.26}$$

is the Higgs potential and D_μ is the covariant derivative as defined in Equation (1.18). The Higgs field has, depending on the value of μ^2, either one global minimum or a degenerate minimum on a circle of radius

$$v = \sqrt{\frac{-\mu^2}{\lambda}} \tag{1.27}$$

around the origin. The value of v can be determined from measurements of the Fermi coupling; it is $v \approx 246\,\text{GeV}$ [18]. The Higgs mechanism connects the masses of the W and Z bosons via the Weinberg angle θ_W:

$$m_Z = \frac{m_W}{\cos\theta_W}. \tag{1.28}$$

The masses of the fermions of the SM are incorporated via the Yukawa coupling. For the coupling to electrons, the interaction term in the Lagrangian density is

$$\mathcal{L} = -y_e(\overline{\psi}_L \phi \psi_R + \overline{\psi}_R \phi \psi_L), \tag{1.29}$$

with the Higgs field ϕ, the electron isospin doublet ψ, see Equation (1.2), and the electron Yukawa coupling y_e. The electron mass is then given by:

$$m_e = \frac{y_e v}{\sqrt{2}}. \tag{1.30}$$

Such a term can be added for any quark or charged lepton to introduce the respective masses to the Lagrangian density.

In this thesis, the κ-framework is used to compare different coupling scenarios to the SM prediction. Therefore, the coupling modifiers

$$\kappa_t = \frac{y_t}{y_t^{SM}} \tag{1.31}$$

and

$$\kappa_V = \frac{g_{HVV}}{g_{HVV}^{SM}} \tag{1.32}$$

are introduced, where y_t^{SM} and g_V^{SM} denote the SM couplings at leading order (LO) to the top quark and the W boson, respectively. Hence, this leads to $\kappa_t = \kappa_V = 1$ for the SM case.

1.2.4. Strong Interaction

As stated above, the quarks of the SM cannot be observed freely. Consequently, there has to be another interaction, namely the strong interaction as described by QCD [24]. The gauge bosons of the strong interaction are the eight gluons, differing by their color content. They couple only to color-charged particles. An SU(3) symmetry group is used to describe QCD. The following Lagrangian density is introduced

$$\mathcal{L}_{QCD} = \overline{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu,a}, \tag{1.33}$$

where

$$D_\mu = \partial_\mu - i g A^a_\mu \lambda^a \tag{1.34}$$

is the covariant derivative with the gluon fields A^a_μ and the Gell-Mann matrices λ_a, and

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu \tag{1.35}$$

is the gluon field strength tensor, with the coupling strength g and the SU(3) structure constant f^{abc}. The potential of the strong interaction can be approximated as

$$V(r) = -\frac{4\alpha_S}{3r} + kr, \tag{1.36}$$

with the strong coupling constant

$$\alpha_S = \frac{g^2}{4\pi} \tag{1.37}$$

and a constant factor k for the linear term. The potential of the strong force increases with distance due to the linear term. This causes the aforementioned effects of confinement and asymptotic freedom. The strong coupling constant, as well as the electromagnetic and weak coupling constants, are actually not constant but dependent on the energy scale. This is caused by two effects: firstly, the color charge of a single quark can be shielded by the creation of virtual quark-antiquark pairs in gluon splitting, called screening. Secondly, virtual gluon loops strengthen the color charge of the quark as they carry color

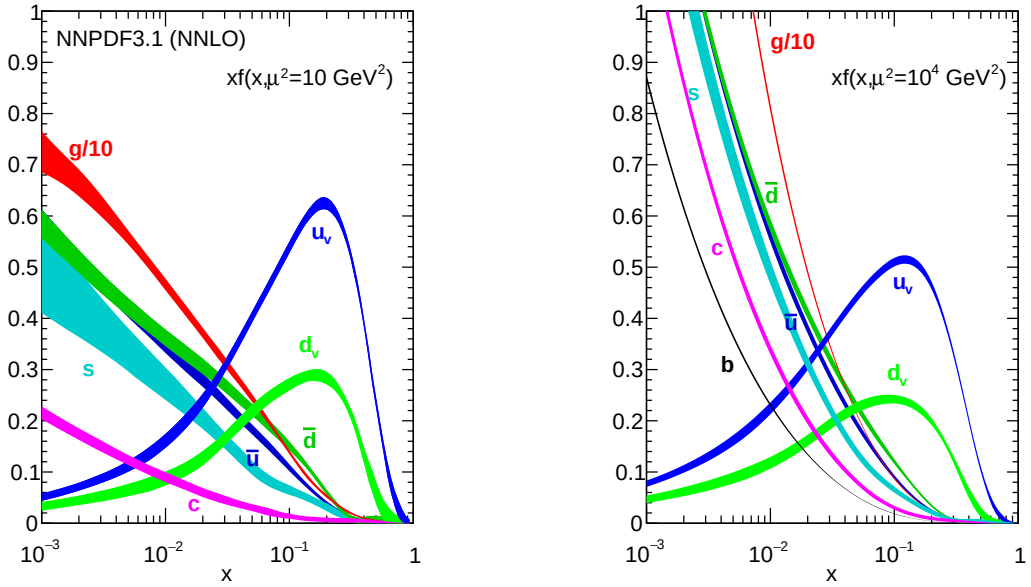

Figure 1.2.: Parton distribution functions for valence and sea quarks, as well as gluons: the PDFs are usually shown at an energy scale of $\mu^2 = 10\,\mathrm{GeV}^2$ (a). This distribution is transformed to $\mu^2 = 10^4\,\mathrm{GeV}^2$ (b) as this is the energy scale relevant at the LHC. Taken from Ref. [30].

charge themselves, called antiscreening. Both effects are taken into account, when defining the one-loop coupling as

$$\alpha_S(\mu^2) = \frac{\alpha_S}{1 - \alpha_S \beta \ln \frac{\mu^2}{\mu_R^2}}, \tag{1.38}$$

where μ_R^2 is the renormalization scale, μ^2 is the scale dependence and

$$\beta = \frac{1}{12\pi}(11 n_c - 2 n_f), \tag{1.39}$$

with the number of color charges n_c and quark flavors n_f. In the SM $n_c = 3$ and $n_f = 6$ holds. As the coupling is dependent on the scale, perturbation theory can only be applied starting from a certain scale Λ_{QCD}, where $\alpha_S \ll 1$.

1.3. Partons versus Particles

As quarks cannot propagate freely, neither the initial-state nor the final-state quarks of the processes searched for in this thesis can be observed directly. At the LHC protons are collided, which consist of three valence quarks, namely two up and one down quark. Additionally, there are gluons mediating the strong force between the quarks within the proton keeping them in a bound state. These gluons can form virtual quark-antiquark pairs, called sea quarks. All these components are referred commonly to as partons.

All partons carry a part of the momentum of a proton. The distribution of the momenta follows a probability distribution called parton density function (PDF). The PDFs are a measure for the probability of observing a certain parton with a fraction x of the whole proton momentum. The PDFs cannot be derived by theory but have to be measured by experiments. The PDFs are dependent on the energy

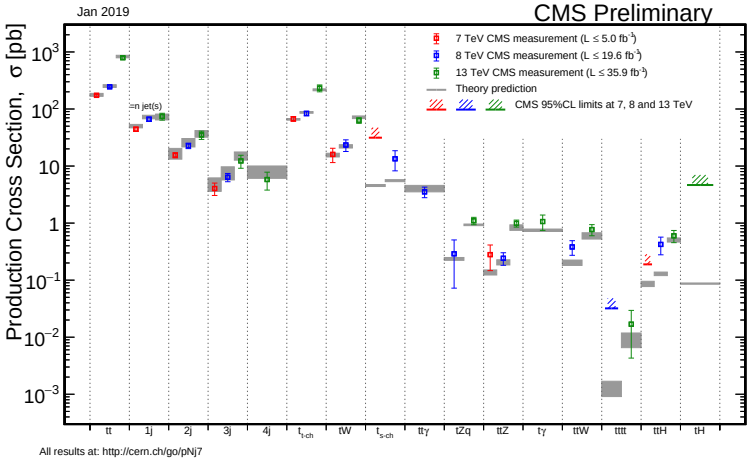

Figure 1.3.: Predicted and observed cross sections for processes containing top quarks. Top quark pair production ($t\bar{t}$) has the highest cross section, followed by single top quark production. The production cross section for tH is about two orders of magnitude smaller than for single top quark production. All measurements are compatible with the theory predictions. Figure taken from Ref. [34].

of the proton. Fortunately, the PDFs can be measured at a certain energy and transformed to other energies using the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equations [31–33]. PDFs for all partons are shown in Fig. 1.2.

1.4. The Top Quark

The top quark, discovered in 1995 [7, 8], is the heaviest particle of the SM with a mass of around 170 GeV. In turn, the Yukawa coupling of the top quark is stronger than for any other quark. The value of the top quark Yukawa coupling is $y_t = \sqrt{2}\frac{m_t}{v} \approx 0.995$, i.e., approximately one, which is still a puzzle, since it is not known whether this is by chance or because of an unknown underlying theory. Since the CKM matrix element V_{tb} is ≈ 1 and the mass of the bottom quark is far smaller than the top quark mass leading to a available large phase space for the decay, the top quark is the only quark not forming bound states as it decays too fast. The mass difference between the top quark and the bottom quark is large enough to enable the production of on-shell W bosons in the decay. The most important production and decay channels are described in the following.

1.4.1. Top Quark Production

Top quarks can be produced via the strong interaction in pairs as well as via the weak interaction as single top quarks. Both production channels give insight into parameters of the SM, e.g., the top quark mass or the V_{tb} CKM matrix element. Furthermore, they are background processes for other measurements, e.g., the search for the associated production of a single top quark and a Higgs boson. Therefore, these production modes are described in the following. All of these production modes have already been observed [7, 8, 35–39]. An overview of all production channels analyzed by the CMS experiment containing a top quark is shown in Fig. 1.3.

Top Quark Pair Production

The pair production is the dominating top quark production process at the LHC as it is mediated by the strong interaction. Two gluons or a quark and an antiquark in the initial state produce an intermediate gluon, which then decays into a top quark and a top antiquark in the *s*-channel production. In the

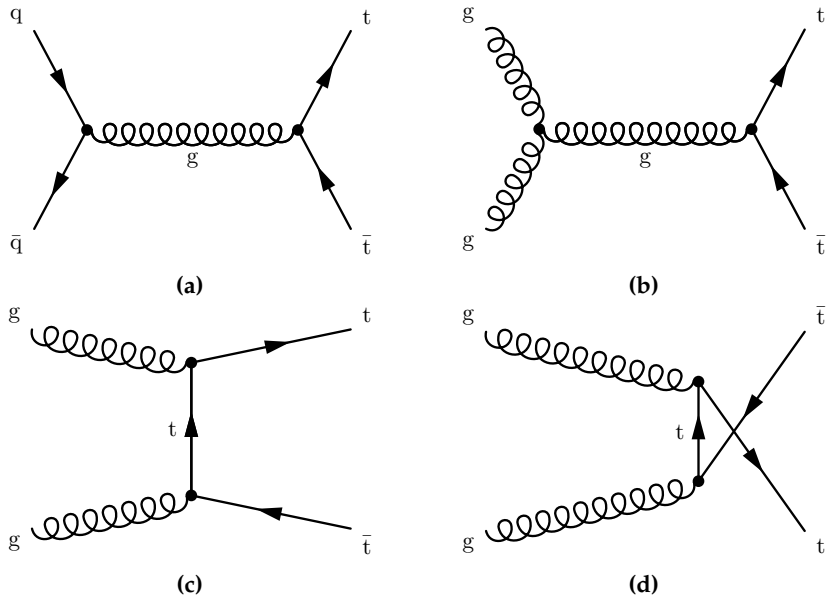

Figure 1.4.: Feynman diagrams for the top quark pair production process with quarks (a) or gluons (b-d) in the initial state.

t-channel production, two gluons in the initial state couple to top quarks, producing a top quark pair. The corresponding Feynman diagrams are shown in Fig. 1.4.

At the LHC, the diagrams with gluons in the initial state are more important, since there are no valence antiquarks in the proton and the PDF for gluons is large at the center-of-mass energy of the LHC. At a proton antiproton collider as the Tevatron at Fermilab with a center-of-mass energy of 1.96 TeV the diagram with a quark and an antiquark in the initial state dominates. In general, the gluon-gluon initial state becomes more important at higher center-of-mass energies also for proton antiproton colliders as gluons carry a larger fraction of the proton momentum at higher momentum transfers. The predicted cross section for top quark pair production at 13 TeV is

$$\sigma_{t\bar{t}} = 831.76^{+19.77}_{-29.20}\ (\text{scale}) \pm 35.06\ (\text{PDF} + \alpha_S)\ \text{pb}, \tag{1.40}$$

assuming a top quark mass of 172.5 GeV [40].

Single Top Quark Production

Single top quarks are produced in processes relying on the weak interaction. Hence, the cross section is smaller compared to the pair production. There are three channels, in which single top quarks can be produced: the *t* channel, the associated production with a W boson and the *s* channel, see Fig. 1.5.

In the *t*-channel production, a light-flavored quark and a bottom quark exchange a W boson and produce a top quark. The light-flavored quark is a characteristic feature of the single top *t*-channel process (tq), it is mostly emitted close to the beam axis. The production cross section

$$\sigma_{tq} = 217.0^{+6.6}_{-4.6}\ (\text{scale}) \pm 6.2\ (\text{PDF} + \alpha_S)\ \text{pb} \tag{1.41}$$

for a center-of-mass energy of 13 TeV is the largest of the three production modes [40].

Following in cross section is the tW associated production, where a bottom quark absorbs a gluon and decays into a top quark and a W boson. The cross section of this process is

$$\sigma_{tW} = 71.7 \pm 1.8\ (\text{scale}) \pm 3.4\ (\text{PDF} + \alpha_S)\ \text{pb} \tag{1.42}$$

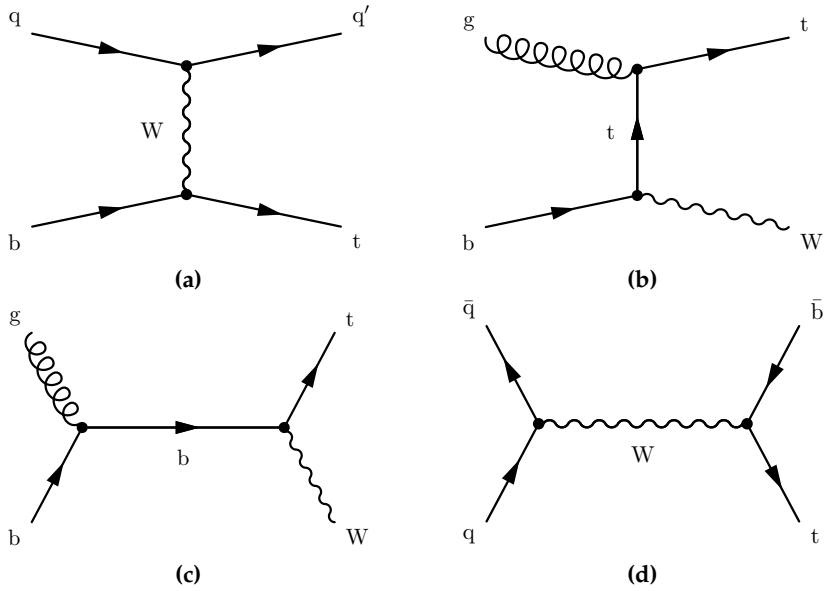

Figure 1.5.: Feynman diagrams ordered by the production cross section: the diagrams show the *t*-channel (a), tW associated (b,c) and *s*-channel (d) single top quark production processes.

for a center-of-mass energy of 13 TeV [40]. The smallest cross section falls to the *s*-channel production (tb), where a quark and an antiquark in the initial state form a virtual W boson which produces a top and a bottom quark. The cross section of this process is

$$\sigma_{\mathrm{tb}} = 10.32^{+0.29}_{-0.24}\ (\mathrm{scale}) \pm 0.27\ (\mathrm{PDF} + \alpha_{\mathrm{S}})\ \mathrm{pb} \tag{1.43}$$

for a center-of-mass energy of 13 TeV [40].

1.4.2. Top Quark Decays

The top quark decays almost exclusively into a W boson and a bottom quark, which corresponds to the V_{tb} matrix element being approximately equal to one. Hence, the different decay channels of the top quark are characterized by the W boson decay.

In 66% of the cases, the W boson decays into quarks, which form hadrons. Therefore, these decays are called hadronic top quark decays. In the rest of the cases, the W boson decays into a charged lepton and the corresponding neutrino. This is called a leptonic top quark decay.

For top quark pairs, this yields three different categories: fully-leptonic, semileptonic and fully-hadronic. In the fully-leptonic decay channel both of the W bosons decay into leptons, whereas in the semileptonic decay channel one of the W bosons decays leptonically and the other one hadronically. In the fully-hadronic decay channel both W bosons decay into hadrons.

1.5. The Higgs Boson

The Higgs boson was the last missing particle of the SM until its discovery in 2012. Its properties are still to be investigated to search for deviations from the prediction. The Higgs boson couples to all massive particles of the SM, but as its coupling is proportional to the mass of the fermions or to the square of the mass of the bosons, respectively, the dominating channels contain heavy particles, such as top and bottom quarks, as well as W and Z bosons and tauons. The best fit values for the coupling strength as a function of the particle mass is shown in Fig. 1.6. After the discovery of the Higgs boson,

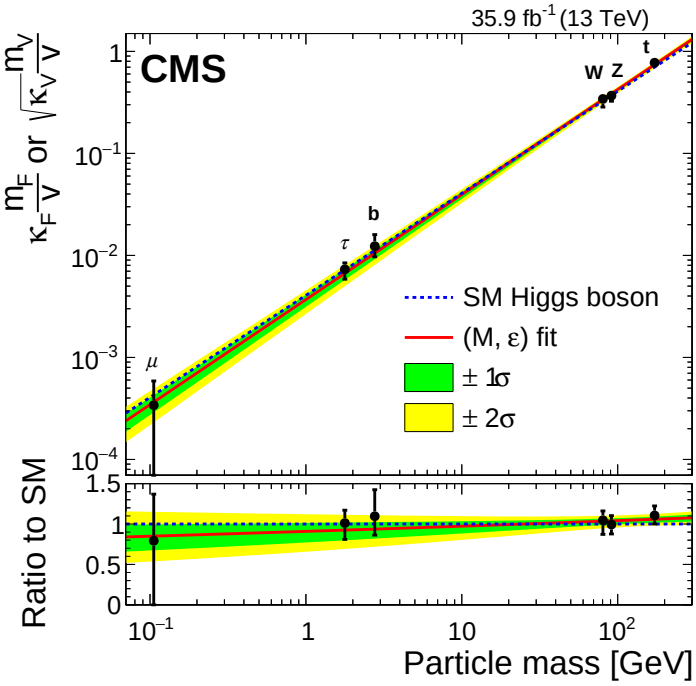

Figure 1.6.: Best fit values for the coupling strength as a function of the particle mass: the measurements for the top and bottom quarks, the W and Z boson, the muon, and the tauon are shown. The SM prediction is given by the dashed blue line, the result of the phenomenological fit is shown as a red line and the uncertainties are shown as green and yellow bands. Taken from Ref. [41].

it is now studied in detail to determine whether deviations from the SM predictions occur. A 2D scan of the Higgs boson couplings to fermions and bosons is shown in Fig. 1.7. In the following, the main production and decay channels are described.

1.5.1. Higgs Boson Production

The Higgs boson can be produced in several different production channels, which are described below. The corresponding Feynman diagrams can be found in Fig. 1.8. The production processes are ordered by their production cross section at the LHC at $\sqrt{s} = 13\,\text{TeV}$.

Gluon-Gluon Fusion (ggF)

At the LHC, processes with gluons in the initial state are very common, as the relevant PDFs favor such processes. However, the Higgs boson cannot couple directly to gluons as they are massless. Instead, it couples via a virtual quark loop, which has dominating contributions by top quarks as shown in Fig. 1.8a. With a predicted cross section of

$$\sigma_{\text{ggF}} = 48.6^{+2.2}_{-3.3}\,(\text{scale}) \pm 1.6\,(\text{PDF} + \alpha_S)\,\text{pb} \tag{1.44}$$

at 13 TeV, it has the highest cross section at the LHC [40].

Vector Boson Fusion (VBF)

Two quarks in the initial state radiate off a vector boson each, which form a Higgs boson. The Feynman diagram for this process can be found in Fig. 1.8b. The predicted cross section for this process is

$$\sigma_{\text{VBF}} = 3.78^{+0.02}_{-0.01}\,(\text{scale}) \pm 0.08\,(\text{PDF} + \alpha_S)\,\text{pb} \tag{1.45}$$

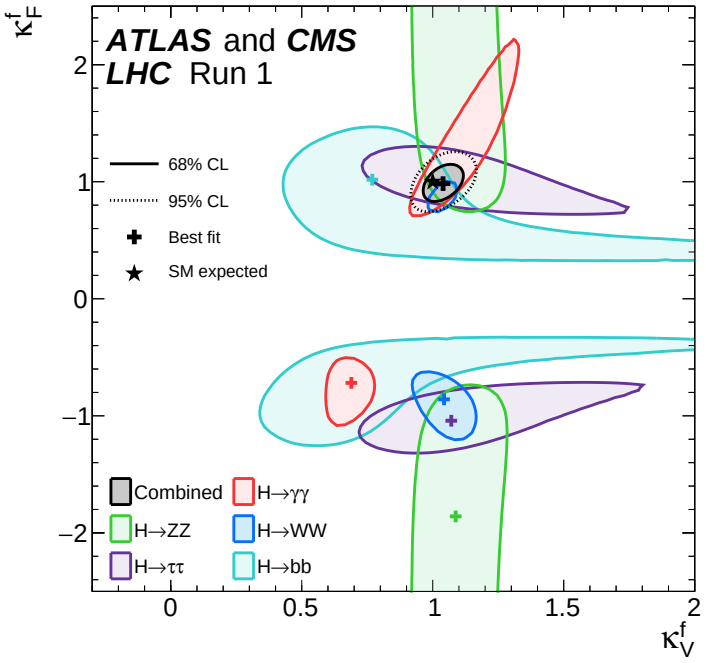

Figure 1.7.: Exclusion limits on the Higgs boson couplings to fermions (κ_F^f) and vector bosons (κ_V^f): most of the processes are not sensitive to the relative sign between the two couplings, which results in symmetric limits. The Higgs boson decay into a pair of photons is sensitive to the relative sign through loop contributions. The difference in the contours for the Higgs boson decay into a pair of W bosons is present as different likelihood fits are used to obtain the minima. Taken from Ref. [42].

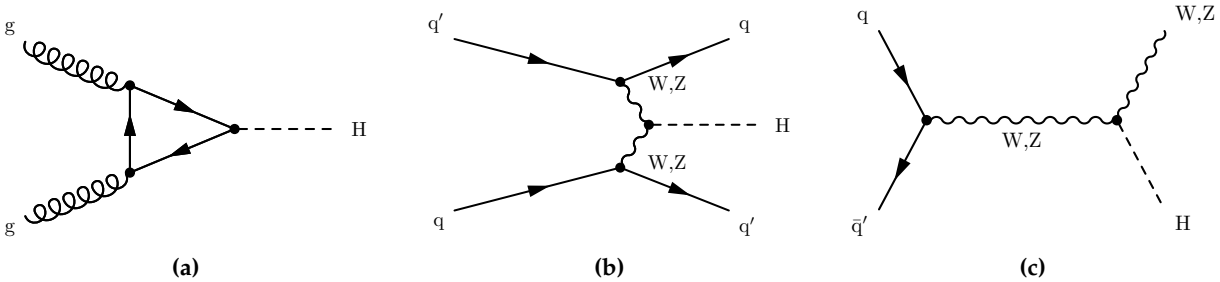

Figure 1.8.: Feynman diagrams for the gluon fusion (a), vector boson fusion (b) and associated vector boson production (c) processes.

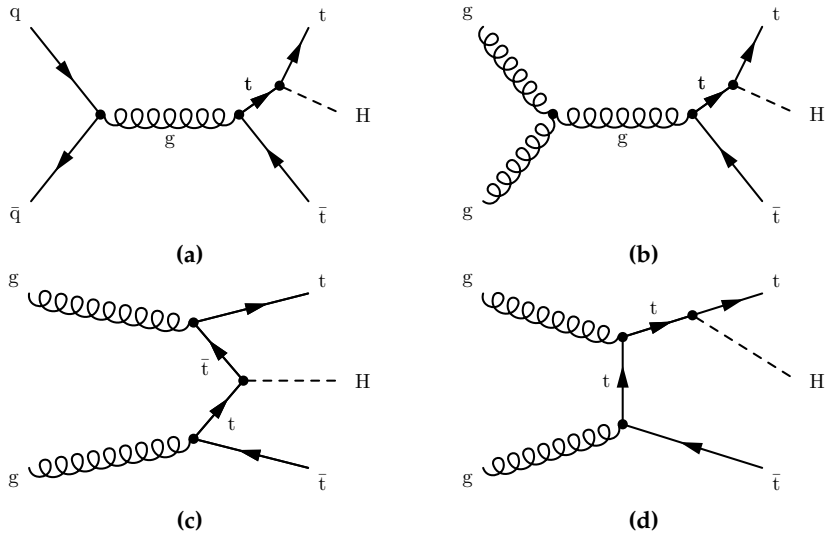

Figure 1.9.: Feynman diagrams for the $t\bar{t}H$ process with quarks (a) and gluons (b-d) in the initial state: the Higgs boson can couple to either of the two top quarks.

at a center-of-mass energy of 13 TeV [40]. It was observed in 2016 by the ATLAS and CMS collaborations [42].

Associated Vector Boson Production (VH)

In a process called Higgsstrahlung or associated vector boson production (VH), two initial-state quarks form a vector boson, which radiates off a Higgs boson, see Fig. 1.8c. With a predicted cross section [40] of

$$\sigma_{VH} = 2.25^{+0.04}_{-0.05}\ (\text{scale}) \pm 0.05\ (\text{PDF} + \alpha_S)\ \text{pb}, \tag{1.46}$$

the VH production was observed in 2018 at the LHC [14].

1.5.2. Associated Production with a Top Quark Pair

The top quark pair production in association with a Higgs boson ($t\bar{t}H$) as shown in Fig. 1.9 is of special interest as it is directly sensitive to the absolute value of the top Yukawa coupling y_t. Direct access to y_t is not possible in measurements of the Higgs boson decays, as the top quark is too heavy to allow the Higgs boson to decay into a pair of top quarks. Before the measurement of the $t\bar{t}H$ process, y_t was only accessible via quantum loops as present in the gluon-gluon fusion process or the Higgs boson decay into photons. As these loops may also contain non-SM particles, which could compensate deviations in the top quark Yukawa coupling, a direct measurement is needed.

With a predicted SM cross section [40] of

$$\sigma^{SM}_{t\bar{t}H} = 507.1^{+11.1}_{-46.6}\ (\text{scale}) \pm 18.3\ (\text{PDF} + \alpha_S)\ \text{fb}, \tag{1.47}$$

the cross section is clearly smaller than for the $t\bar{t}$ process, which is one of the main backgrounds for this analysis. This overwhelming background contribution makes this analysis especially challenging. The $t\bar{t}H$ process was observed in 2018 at the LHC [43].

As the cross section of the $t\bar{t}H$ process

$$\sigma_{t\bar{t}H}(\kappa_t) = \kappa_t^2 \cdot \sigma^{SM}_{t\bar{t}H}, \tag{1.48}$$

depends on the square of the top quark Yukawa coupling, the measurement of this process allows to set limits on the magnitude of κ_t and probe whether this coupling is compatible with the SM prediction.

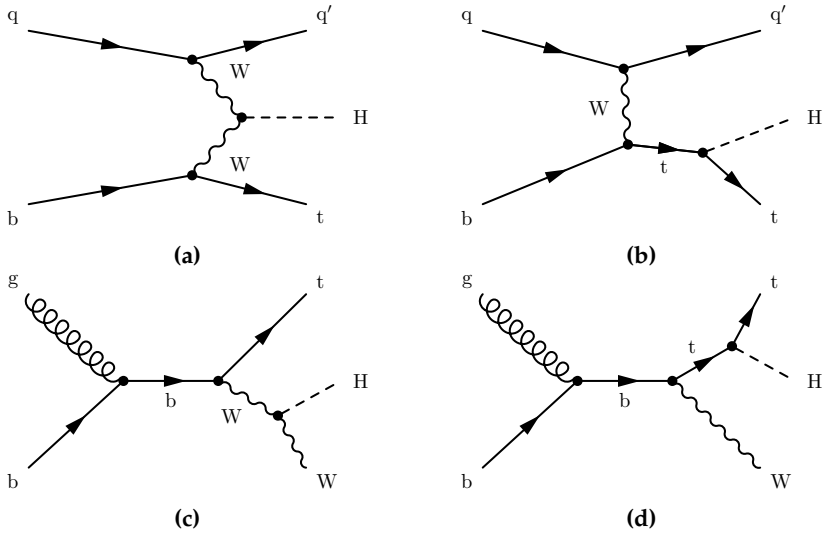

Figure 1.10.: Feynman diagrams for the tHq (a,b) and tHW (c,d) processes: the Higgs boson either couples to the W boson (a,c) or to the top quark (b,d).

1.5.3. Associated Production with a Single Top Quark

In addition to the sensitivity on the absolute value of the top quark Yukawa coupling, as present in the $t\bar{t}H$ process, single top quark production in association with a Higgs boson (tH) is also sensitive to the relative sign between the top quark Yukawa coupling and the Higgs boson to vector boson coupling. Hence, this process is able to test the top quark Yukawa coupling in a more comprehensive way.

In principle, the Higgs boson can be produced in association with any of the single top quark production modes. In this analysis, only two of them are considered: the associated production of a Higgs boson and a single top quark in the t channel (tHq) with an SM cross section of

$$\sigma_{\mathrm{tHq}}^{\mathrm{SM}} = 74.3^{+4.8}_{-11.1}\ (\mathrm{scale}) \pm 2.7\ (\mathrm{PDF} + \alpha_{\mathrm{S}})\ \mathrm{fb} \tag{1.49}$$

and the associated production with a single top quark in the tW production mode (tHW) with an SM cross section of

$$\sigma_{\mathrm{tHW}}^{\mathrm{SM}} = 15.2^{+0.7}_{-1.0}\ (\mathrm{scale}) \pm 1.0\ (\mathrm{PDF} + \alpha_{\mathrm{S}})\ \mathrm{fb} \tag{1.50}$$

at a center-of-mass energy of 13 TeV [40]. The s-channel production is negligible due to its low cross section of 2.9 fb. In tHq and tHW production, either the W boson or the top quark radiates off a Higgs boson, leading to two Feynman diagrams with the same final state for each of the production modes, see Fig. 1.10. The interference of these two diagrams leads to a sensitivity of the cross section and the kinematic distribution of the final-state particles to the strength of the coupling of the Higgs boson to the W boson, g_{HVV}, and to the top quark, y_{t}. In addition to the sensitivity to the absolute value, this process is also sensitive to the relative sign between the two couplings. The amplitudes of the Feynman diagrams are given as

$$\begin{aligned}\mathcal{A} = \frac{g}{\sqrt{2}}\Bigg[&(\kappa_{\mathrm{t}} - \kappa_{\mathrm{V}})\frac{m_{\mathrm{t}}\sqrt{s}}{m_{\mathrm{W}}v}A\left(\frac{t}{s}, \varphi; \xi_{\mathrm{t}}, \xi_{\mathrm{b}}\right) \\ &+ \left(\kappa_{\mathrm{V}}\frac{2m_{\mathrm{W}}}{v}\frac{s}{t} + (2\kappa_{\mathrm{t}} - \kappa_{\mathrm{V}})\frac{m_{\mathrm{t}}^2}{m_{\mathrm{W}}v}\right)B\left(\frac{t}{s}, \varphi; \xi_{\mathrm{t}}, \xi_{\mathrm{b}}\right)\Bigg],\end{aligned} \tag{1.51}$$

with $\kappa_{\mathrm{V}} = \frac{g_{\mathrm{HVV}}}{g_{\mathrm{HVV}}^{\mathrm{SM}}}$ and $\kappa_{\mathrm{t}} = \frac{y_{\mathrm{t}}}{y_{\mathrm{t}}^{\mathrm{SM}}}$ as the ratios of the vector boson and top quark couplings relative to the SM, respectively, the masses of the W boson m_{W} and the top quark m_{t}, the Mandelstam variables t and s, and the vacuum expectation value v. The expressions A and B are explained in Ref. [44]. The cross

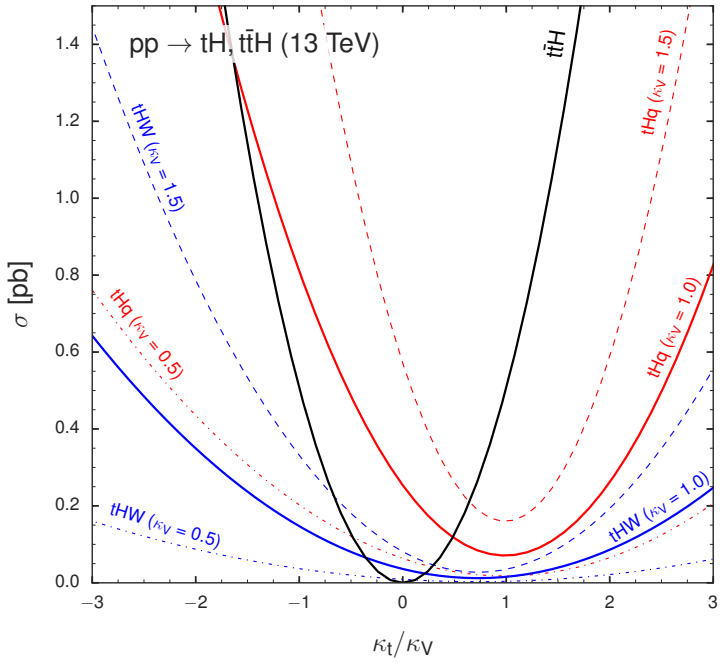

Figure 1.11.: Production cross sections for the tH and $t\bar{t}H$ processes for different coupling scenarios: the solid lines represent $\kappa_V = 1$, wheras the dashed lines represent $\kappa_V = 0.5$ and $\kappa_V = 1.5$, respectively. The SM scenario is at $\kappa_t/\kappa_V = 1$ and $\kappa_V = 1$. Taken from the additional figures to Ref. [2].

sections for the tHq and tHW processes with non-SM couplings can be approximately parametrized with:

$$\sigma_{tHq} = (2.63\,\kappa_t^2 + 3.58\,\kappa_V^2 - 5.21\,\kappa_t\kappa_V) \times \sigma_{tHq}^{SM} \tag{1.52}$$

and

$$\sigma_{tHW} = (2.91\,\kappa_t^2 + 2.31\,\kappa_V^2 - 4.22\,\kappa_t\kappa_V) \times \sigma_{tHW}^{SM}. \tag{1.53}$$

The cross sections are depicted in Fig. 1.11. Furthermore, the tH production can be used to study possible CP-violating components in the Higgs boson coupling [45]. The effective Lagrangian density for this coupling can be written as

$$\mathcal{L} = \overline{\psi}\big(\cos(\alpha)\kappa_{Htt}g_{Htt} + i\sin(\alpha)\kappa_{Att}g_{Att}\gamma_5\big)\psi X_0, \tag{1.54}$$

with the CP-mixing phase α, the coupling modifiers κ_{itt}, the coupling strengths $g_{itt} = \frac{m_i}{v}$ and the corresponding field X_0 representing the Higgs boson. The scalar component is described by Htt, where κ_{Htt} is equivalent to the above defined κ_t. The pseudoscalar component is denoted with the index Att. A mixing angle of $\alpha = 0°$ corresponds to a purely CP-even coupling, whereas $\alpha = 90°$ holds for a purely CP-odd state. The SM case can be recovered, when setting $\alpha = 0°$ and $\kappa_{Htt} = 1$. The cross sections for tH and $t\bar{t}H$ production for different mixing angles are shown in Fig. 1.12. The tH process is sensitive to the whole range of mixing angles, whereas the $t\bar{t}H$ process is degenerate in α.

The anomalous Higgs boson couplings regarding a possible CP-mixing is also studied in Higgs boson to Z boson decays [46].

1.5.4. Higgs Boson Decays

In principle, the Higgs boson can decay into any pair of massive SM particles. Decays into massless particles can also be realized via virtual particle loops. As the coupling is proportional to the mass of the decay particles, heavier daughter particles are favored. In general, the branching ratios are dependent

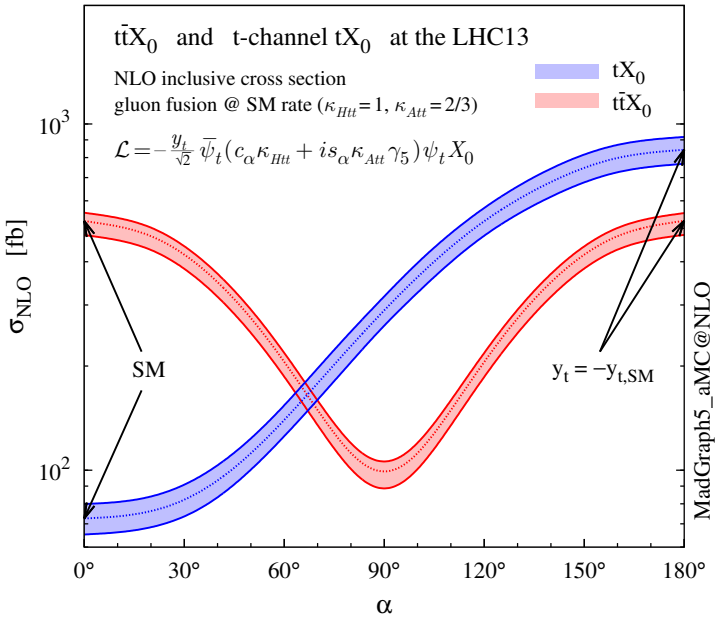

Figure 1.12.: Production cross sections for the tH and $t\bar{t}H$ processes for different CP-mixing scenarios: the tH process is shown in blue and $t\bar{t}H$ process in red. The SM case corresponds to $\alpha = 0°$. Taken from Ref. [45]

on the mass of the Higgs boson. For a Higgs boson mass of 125 GeV, the $H\to b\bar{b}$ decay is dominating. The most important decays are described in the following.

Decay into a Pair of Bottom Quarks

The decay into a pair of bottom quarks is the dominating decay channel with a predicted branching ratio of 58% at a mass of 125 GeV [18]. Although the branching ratio is large, the measurement of this decay channel is challenging as there are background processes with the same final-state particles. However, it was observed in 2018 by the CMS [13] and ATLAS [14] collaborations. The observation uses a combination of the VH, ggF, VBF, and $t\bar{t}H$ production channels.

Decay into a Pair of W bosons

With a branching ratio of about 21% [18] the decay into two W bosons is the second most common decay channel. It can be divided into three categories depending on whether the W boson decays leptonically or hadronically. The fully-leptonic channel with both W bosons decaying into a lepton and a neutrino was used for the discovery of the Higgs boson [10, 11]. This decay channel was first observed in 2014 by combining the ggF, VBF and VH production channels [47].

Decay into a Pair of Tau Leptons

The Higgs boson decay into two τ leptons has a branching ratio of about 6% [18] and was observed in 2016 for the first time [42]. This was the first direct observation of a coupling between a fermion and the Higgs boson.

Decay into a Pair of Z Bosons

Although the branching ratio into two Z bosons is only 2.6%, it was used in the search for the Higgs boson [10, 11], since it gives a clear signature when requiring a decay into charged leptons. The channel itself was observed in 2013 [48].

Decay into a Pair of Photons

The Higgs boson cannot decay directly into photons, but this decay is still possible via a fermion or boson loop. This leads to a small branching ratio of about 0.2% [18]. Nevertheless, it is a very sensitive channel since it produces a very clear signature with two isolated photons of high transverse momentum. The Higgs boson decay into a pair of photons was the most sensitive channel for the Higgs boson observation [10, 11]. Furthermore, this decay channel gives access to the top quark Yukawa coupling through the virtual quark loop.

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

To test how well theories describe nature, experimental setups are indispensable. Conclusions on the validity of the underlying theories can be drawn from the measurements of these experiments. In principle there are two goals for such experiments: either to measure already observed phenomena with higher precision in order to look for deviations from the theory prediction or to search for new phenomena either previously predicted by a theory or totally unexpected. There is a long series of experiments in particle physics, most of which belong to one of the following categories: collider, fixed-target or astroparticle experiments.

The analyses presented in this thesis use data taken with the Compact Muon Solenoid (CMS) experiment, located near Cessy, France. It is one of the four large experiments of the Large Hadron Collider (LHC), which started operation in 2008, at the European Organization of Nuclear Research (CERN), Geneva, Switzerland. In the following, the LHC and the CMS experiment are described in more detail.

This chapter is, if not stated otherwise, based on Ref. [49] regarding the LHC in Section 2.1 and on Ref. [50] for the CMS experiment in Section 2.2.

2.1. The Large Hadron Collider

The LHC is located inside a 27 km long tunnel about 100 m below ground in the Geneva area in Switzerland and France. For practical reasons, it was built in the same tunnel that was used for the Large Electron-Positron Collider (LEP) [52]. The acceleration and deflection of particles relies on the Lorentz force. When particles are deflected, they loose energy through synchrotron radiation, which is proportional to m^{-4} of the accelerated particles. The LEP used electrons, which loose a significant amount of their energy through synchrotron radiation. To search for new particles, higher center-of-mass energies are required. Hence, the collider uses protons, which are extracted from a hydrogen source and are ionized afterwards. As the protons cannot be injected directly into the LHC, there is a whole accelerator complex built to preaccelerate them, see Fig. 2.1. The preaccelerators are designed specifically to provide a large amount of protons and allow stable beams in the LHC.

The first link of the accelerator chain is the LINAC 2, which accelerates the protons to energies of 50 MeV [53]. As most modern particle accelerator, the LINAC 2 utilizes radio frequency (RF) cavities for the acceleration. In order to use the cavities several times for the acceleration, all following accelerators are build as synchrotrons. The protons are further accelerated in the Proton Synchrotron Booster to energies of 1.4 GeV [54]. The Proton Synchrotron raises the energy of the protons to 25 GeV. It has a circumference of 628 m and uses conventional electromagnets for bending and focusing. With a circumference of about seven kilometers, the Super Proton Synchrotron is the last accelerator before the LHC. It brings the protons to energies of up to 450 GeV using more than 1000 conventional electromagnets for bending and focusing [55].

Finally, the protons are injected into the LHC [49] with a circumference of 27 kilometers. The LHC was designed to operate at center-of-mass energies of up to 14 TeV. In the data taking period from 2010 to 2012, referred to as run 1, the center-of-mass energy was $\sqrt{s} = 7$ TeV for the first two years and $\sqrt{s} = 8$ TeV in 2012. In run 2, the operation period from 2015 to 2018, the center-of-mass energy was increased to $\sqrt{s} = 13$ TeV. In order to achieve this, two proton beams with energies of 6.5 TeV each are

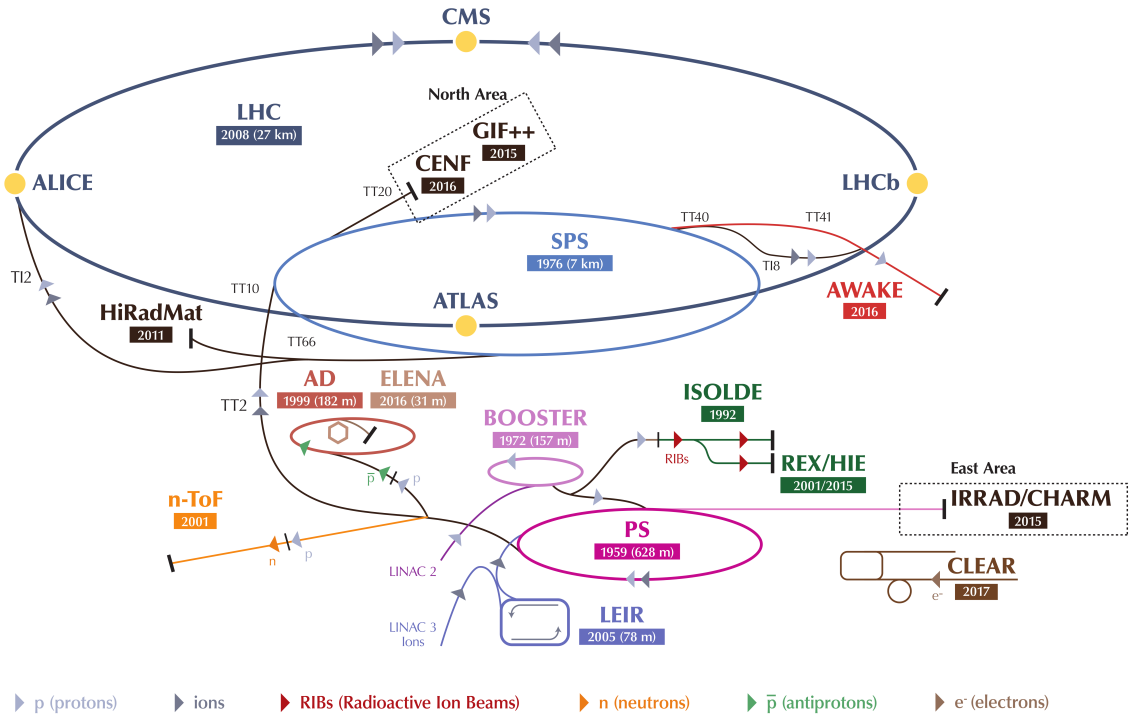

Figure 2.1.: Sketch of the CERN accelerator complex, including the LHC and its preaccelerators: starting from LINAC 2 (on the bottom in the middle) the protons are injected into the BOOSTER, then into the Proton Synchrotron (PS), Super Proton Synchrotron (SPS), and finally into the LHC. Modified, taken from Ref. [51].

collided at four points, each equipped with an experiment. Furthermore, the LHC is also able to collide heavy ions instead of protons. Nevertheless, the proton operation is described in the following.

There are two general-purpose detectors, which are mainly designed to search for the Higgs boson and new physics beyond the SM: CMS [50] and ATLAS [56] detectors, both providing data of comparable quality and therefore able to crosscheck each others results. In addition, there are two detectors targeting specific areas of particle physics: A Large Ion Collider Experiment (ALICE) [57], studying mainly heavy ion physics, and Large Hadron Collider beauty (LHCb) [58], primarily investigating physics processes containing B mesons in forward rapidity regions.

The goal of an accelerator, such as the LHC, is to provide a large number of particle interactions and hence enable the experiments to search for rare processes. A measure for the interaction rate is the instantaneous luminosity, defined as

$$L = f \cdot \frac{nN_1N_2}{4\pi\sigma_1\sigma_2}, \tag{2.1}$$

with the revolution frequency f, the number of bunches n, the number of particles per bunch in beam $i = 1, 2$, N_i, and the width of the bunches σ_i. The first accelerator with a center-of-mass energy larger than 1 TeV was the Tevatron, which collided a proton and an antiproton beam resulting in a center-of-mass energy of 1.96 TeV [59]. The LHC is designed to provide a luminosity of $L = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ at a center-of-mass energy of 14 TeVin proton-proton collisions. This is only possible by using two proton beams instead of a proton and an antiproton beam as the beam intensity would be too low otherwise, e.g., the Tevatron had an instantaneous luminosity of about $L = 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in its last year of operation [60]. In general, the experiments are interested in the total number of collisions rather than the collision rate. Therefore, the integrated luminosity

$$L_{\mathrm{int}} = \int L\mathrm{d}t, \tag{2.2}$$

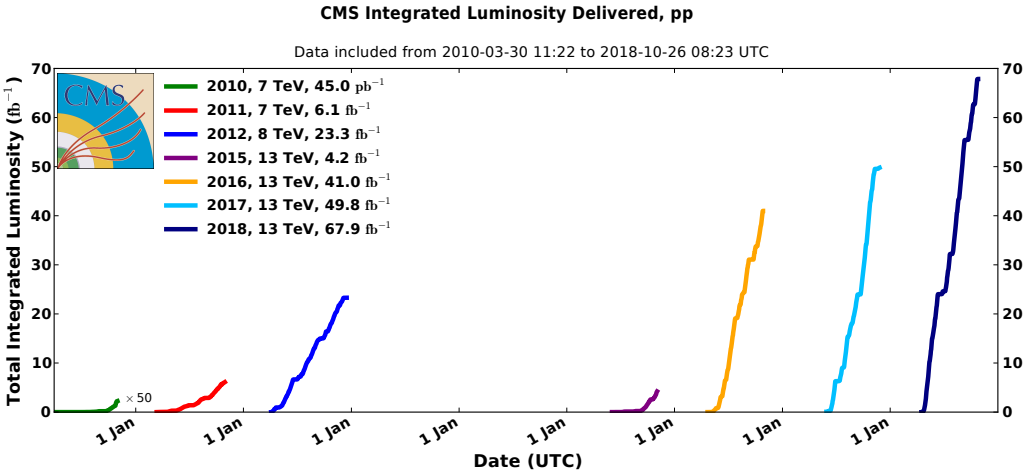

Figure 2.2.: Integrated luminosity of the LHC during the run 1 (2010-2012) and run 2 (2015-2018) data taking periods. The combined integrated luminosity for run 1 and run 2 is almost 200 fb^{-1}. Taken from Ref. [61].

of a data taking period is introduced. It is related to the number of particles

$$N = L_{\text{int}} \cdot \sigma, \tag{2.3}$$

produced in a process with a cross section σ. The integrated luminosity of the LHC over its time of operation up to 2018 is shown in Fig. 2.2.

The LHC ring is divided into eight straight sections, four of which are equipped with the experiments, and eight arcs. There are two RF cavities for the acceleration of the protons, one for each of the beams. As the protons travel in bunches with a timely distance of 25 ns, the RF cavities are operated at a frequency of 400 MHz. In total, the LHC can operate with up to 2808 bunches. Since the protons lose only 7 keV per turn through synchrotron radiation at 7 TeV beam energy, one RF cavity per beam suffices to reach the 14 TeV center-of-mass energy the LHC was designed for.

The LHC consists of 1232 dipole and 392 quadrupole magnets. The dipole magnets are superconducting, operating at a temperature of 1.9 K. The dipoles produce magnetic fields up to 8 T. Each of the dipoles has a mass of 27.5 t, a length of 16.5 m and a diameter of 57 cm; the inner diameter of the beam pipe is 5 cm. A cross section of an LHC dipole magnet is shown in Fig. 2.3. To operate the LHC with stable beams, the magnets are required to produce almost identical fields, where the relative deviations of the fields must not exceed 10^{-4}.

The beam pipes are evacuated to contain less than 10^{15} molecules per cubic meter to achieve the required beam lifetime of several hours. This corresponds to a pressure of 10^{-10} mbar. The cooling of the magnets is realized through superfluid helium at a maximum temperature of 1.9 K. The superfluidity guarantees a low temperature gradient over the whole ring, enabling uniform conditions for all dipole magnets. Therefore, only eight cooling plants, placed at distances of up to 3.3 km in the tunnel, are needed. In total, almost 100 t of helium are used in the cooling system.

About 1000 beam position monitors are installed to monitor the particle beams, which are used to adjust the magnets. The information of these monitors is used to control the magnets and RF cavities. The control system is built in three levels: the equipment level, the server level and the control room. If a beam gets out of control or if the intensity of the beam is too low to produce enough events, the beam gets dumped by kicker magnets in a dedicated beam dumping area.

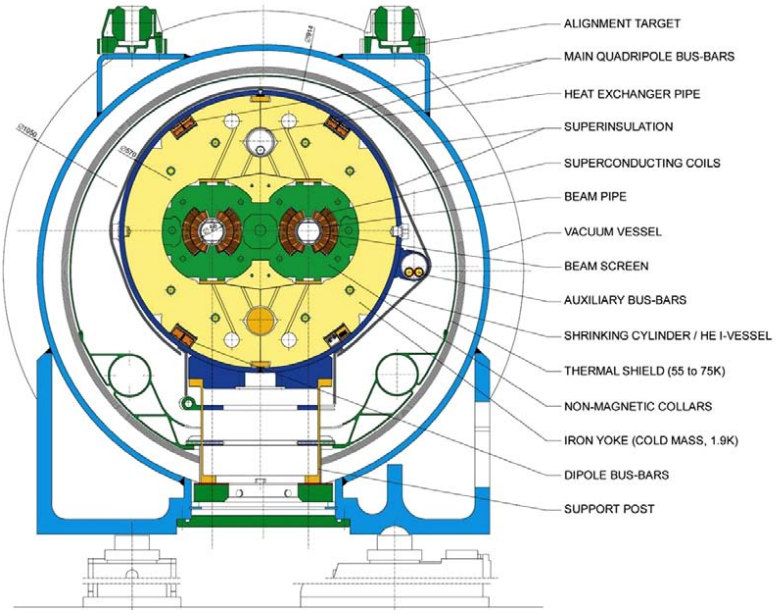

Figure 2.3.: Cross section of an LHC Dipole magnet with a diameter of 57 cm and an inner diameter of the beam pipes of 5.0 cm. The two beam pipes are placed in the center, embedded in several cooling and insulation layers. The dipoles provide magnetic fields of up to 8 T. Taken from Ref. [49].

2.2. The Compact Muon Solenoid Experiment

The CMS experiment, located in Cessy, France, is one of the two general purpose detectors at the LHC. It was mainly designed to discover the Higgs boson, which was successfully achieved in 2012, and to search for new physics at the TeV scale. Afterwards, the focus shifted to measurements of the Higgs boson properties, searches for new physics processes beyond the SM and precision measurements of SM processes. The CMS experiment, as most other particle detectors, is built in several layers around the interaction point, starting with the tracker system surrounded by the electromagnetic and hadron calorimeters, the solenoid magnet and, as the outermost part, the muon chambers. The whole CMS detector weighs about 14000 t, is 28.7 m long, and 15 m in diameter. When the LHC is operated at its design luminosity, there are about 10^9 inelastic interactions per second, requiring sophisticated techniques to select the most interesting events. A view of the whole detector can be found in Fig. 2.4 and a slice of the detector is shown in Fig. 2.5. The CMS collaboration, which operates the experiment, consists of approximately 4000 people from all over the world.

The data recorded by the CMS experiment is processed in several layers: the first layer is the trigger system, which decides whether to keep an event or to discard it. Afterwards, the events are processed by the computing infrastructure to reconstruct physics objects. Finally, the data is analyzed either on the decentralized computing infrastructure or on local computing resources. Furthermore, the computing environment is used to simulate events of relevant processes. The trigger system and the computing infrastructure are explained in this chapter, whereas the event reconstruction and simulation are explained in Chapter 4.

2.2.1. Coordinate System

The coordinate system is defined in a way that the x axis points towards the center of the LHC, the y axis points in upward direction and the z axis points in the direction of the counterclockwise traveling beam. The origin of the coordinate system is the center of the CMS detector. As the CMS detector is rotationally symmetric around the beam axis, it is beneficial to use polar instead of Cartesian coordinates.

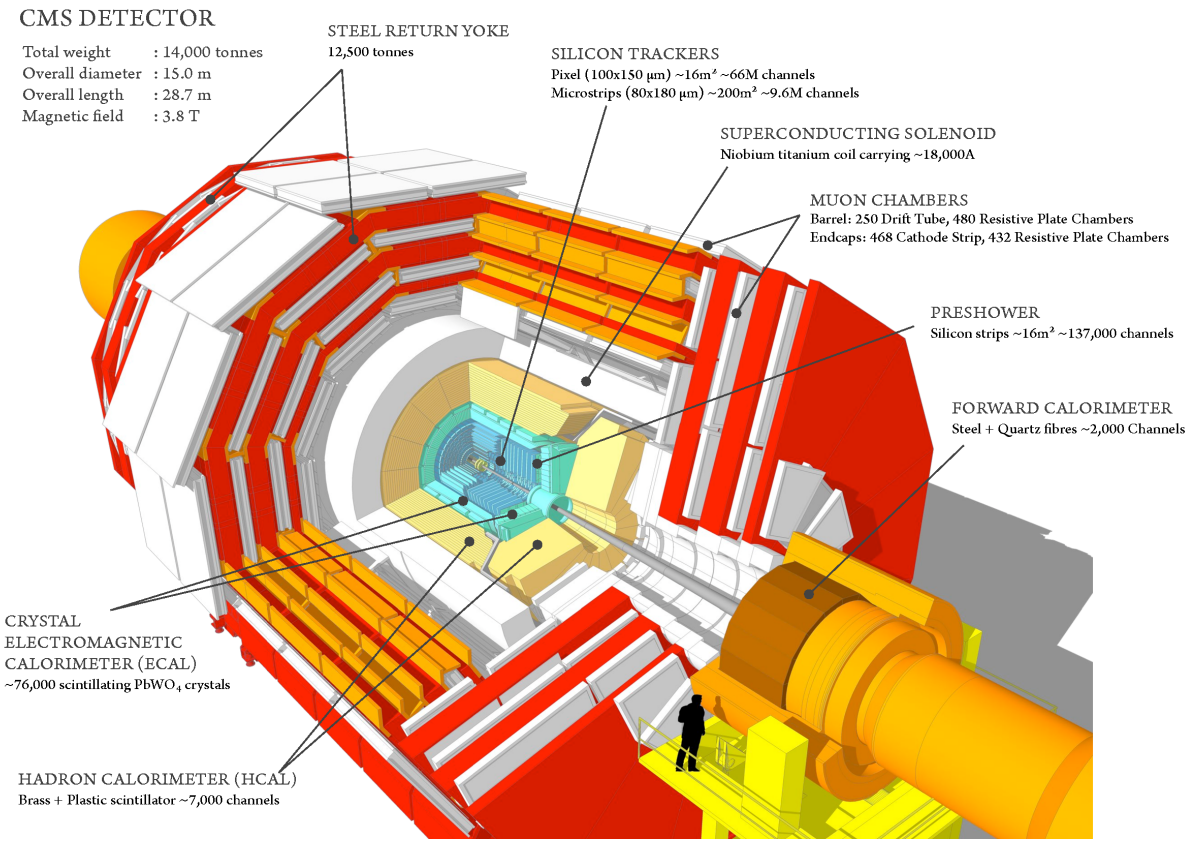

Figure 2.4.: Cutaway view of the CMS detector: the beam pipe is located in the center of the detector surrounded by the tracker system, the electromagnetic and hadron calorimeter, the superconducting solenoid and, as the outermost layer, the muon system, embedded in the steel return yoke. Taken from Ref. [62].

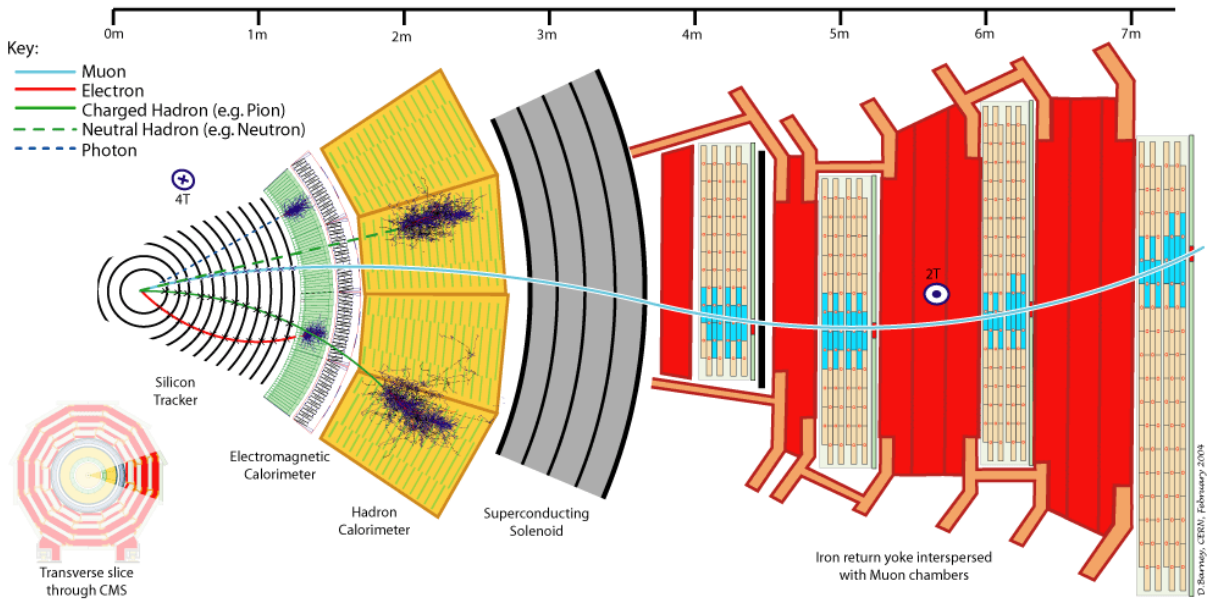

Figure 2.5.: Slice of the CMS detector: the subdetectors of the CMS experiment are shown including the expected energy depositions of muons, electrons, charged and neutral hadrons, as well as photons, in the subdetectors. Charged particles traverse the detector on curved trajectories as a result of the magnetic field produced by the superconducting solenoid. Taken from Ref. [63].

Therefore, the angle ϕ in the x-y plane is introduced, where $\phi = 0$ corresponds to the unit vector in x direction. Furthermore, the angle θ is introduced, where $\theta = 0$ corresponds to the negative vector in z direction. The momentum in z direction of the initial-state particles taking part in the interaction is unknown because the proton is not elementary. Only the transverse momentum p_{T} of the incoming particles is known to be zero. Therefore, momenta are denoted with p_{T}, ϕ and, for more convenience, pseudorapidity:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{2.4}$$

Finally, the connection of this coordinate system to the Cartesian coordinates is given by:

$$p_x = p_{\mathrm{T}} \cos\phi, \tag{2.5}$$

$$p_y = p_{\mathrm{T}} \cos\phi, \tag{2.6}$$

and

$$p_z = p_{\mathrm{T}} \sinh\eta. \tag{2.7}$$

An illustration of the coordinate system used at the CMS experiment can be found in Fig. 2.6.

2.2.2. Tracking System

The precise measurement of the trajectories of the decay products is essential for the event reconstruction and particle identification. Therefore, a high resolution tracking system is implemented. It consists of two parts surrounding the interaction point with an outer diameter of 2.5 m and a length of 5.8 m: the pixel detector and the strip detector. The pixel detector surrounds the interaction point and the strip detector encloses the pixel detector. The tracking system of the CMS detector is shown in Fig. 2.7.

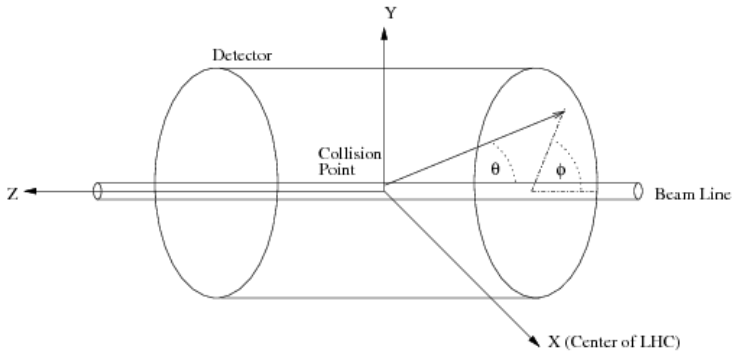

Figure 2.6.: Illustration of the coordinate system used at the CMS experiment: the x, y and z axis, as well as the θ and ϕ angles, are depicted, with the nominal collision point as the origin of the coordinate system. Taken from Ref. [64].

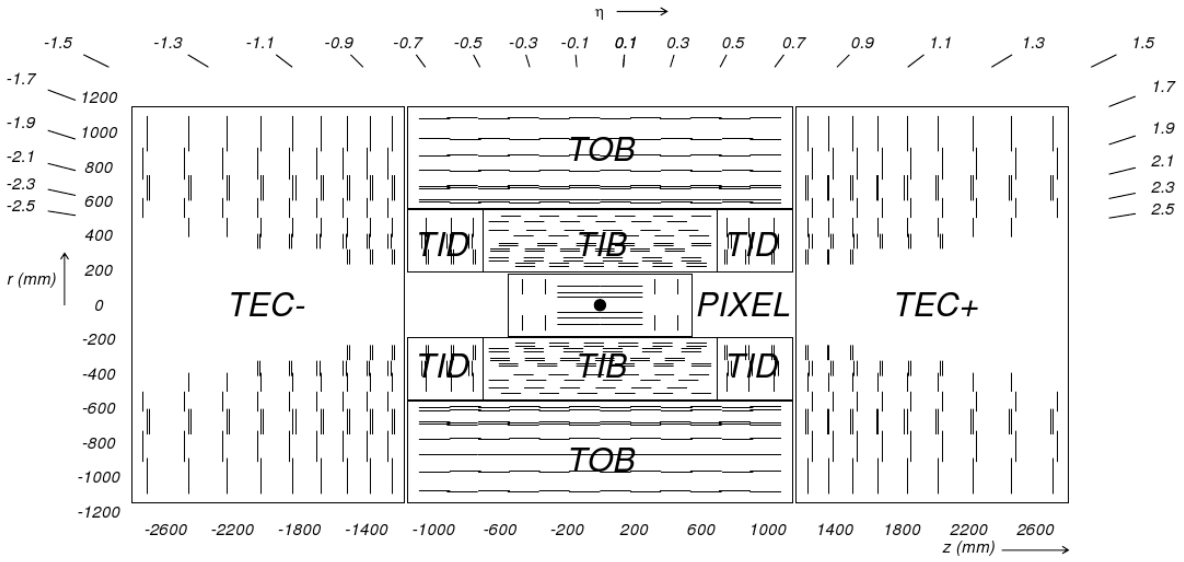

Figure 2.7.: Depiction of the CMS tracking system. The interaction point is marked with a black dot in the center. The pixel detector is located closest to the beam axis. It is surrounded by the tracker inner barrel (TIB) and the tracker inner disks (TID). The tracker outer barrel (TOB) and the tracker end caps (TEC) enclose the inner tracker parts. Taken from Ref. [65].

Silicon Pixel Detector

The pixel detector is the part of the CMS detector that is closest to the interaction point. Its purpose is the precise measurement of the tracks of charged particles in order to provide good vertex resolution, proving to be especially helpful for the identification of jets originating from bottom quarks. The pixel detector as built in the initial design consists of three layers around the beam axis with radii of 4.4 cm to 10.2 cm and endcaps consisting of two disks. In total 1440 pixel modules are used. This setup covers a pseudorapidity of $|\eta| < 2.5$. The first version of the pixel detector consisting of three layers was used until the end of 2016; from 2017 on, a new four-layer version is used, which is called Phase 1 upgrade [66].

The high particle flux close to the interaction point, as well as the high track density, require the pixel modules to be radiation tolerant. This is achieved using an n+ pixel on n- substrate design, which allows partial depletion even under the conditions close to the beam pipe. Thereby, a spatial resolution of 15 to 20 μm is reached. Each of the modules consists of 8 to 16 readout chips with 52×80 pixels each with a size of $100 \times 150\,\mu\text{m}^2$ and a thickness of 180 μm. The complete module measures $66.6 \times 26.0\,\text{mm}^2$. In order to provide the best working conditions, the modules are cooled to -10° C using liquid phase cooling. For further details see Ref. [50].

Silicon Strip Detector

The silicon strip tracker is the second part of the CMS tracking system. It consists of four parts: the tracker inner barrel (TIB) is made up of four cylinders with radii of 255 mm, 339 mm, 418.5 mm and 498 mm ranging from -700 mm to +700 mm along the beam axis. The tracker inner disks (TID) consist of disks placed between 800 mm and 900 mm along the beam axis covering a radius from 200 mm to 500 mm; three disks are placed in positive and three in negative z direction. The tracker outer barrel (TOB) encloses the TIB and the TID; it consists of a single mechanical structure with an inner radius of 555 mm and an outer radius of 1160 mm, and a length of 2180 mm. The tracker endcaps (TEC) with an inner radius of 220 mm and an outer radius of 1135 mm cover the forward regions of the detector ranging from 1240 mm to 2800 mm in positive and negative direction along the beam axis. Each of the endcaps consists of nine disks made up of modular elements, called petals. These petals allow the exchange of parts of the endcaps without the need of disassembling the whole structure.

The strip tracker is composed of about 15000 detector modules used in the TIB, TOB, TID and TEC. In total, approximately 24000 p-on-n silicon microstrip sensors are installed covering a total active area of about $200\,\text{m}^2$. The sensors have thicknesses between 320 μm to 500 μm and sizes of $6 \times 12\,\text{cm}^2$ and $10 \times 9\,\text{cm}^2$ for the inner and outer barrel, respectively. The active area per module varies between approximately $6000\,\text{mm}^2$ and $17000\,\text{mm}^2$. Depending on the position of the module within the detector, different modules and sensors are used.

The modules installed in the TIB, TID and the first four rings of the TEC have one sensor, whereas modules in the TOB and the rings 5-7 of the TEC employ two sensors. The modules are mounted either on a frame of carbon fiber or on a frame of graphite. Further details on the silicon strip detector can be found in Ref. [50].

2.2.3. Electromagnetic Calorimeter

In addition to a precise measurement of the particle tracks, determining the energy of the particles is essential for physics analyses. The electromagnetic calorimeter (ECAL) measures the energy of electrons and photons. As the energy measurement is destructive, it has to be performed after the spatial measurement. Hence, the ECAL is placed outside of the tracking system at a radius between 1.29 m and 3.15 m. It is composed of a barrel part (EB) covering a pseudorapidity of $|\eta| < 1.479$ and endcaps (EE) covering $1.479 < |\eta| < 3.0$, as well as the preshower detector (ES) at $1.653 < |\eta| < 2.6$ as shown in Fig. 2.8. It is built as a hermetic homogeneous calorimeter of lead tungstate crystals ($PbWO_4$),

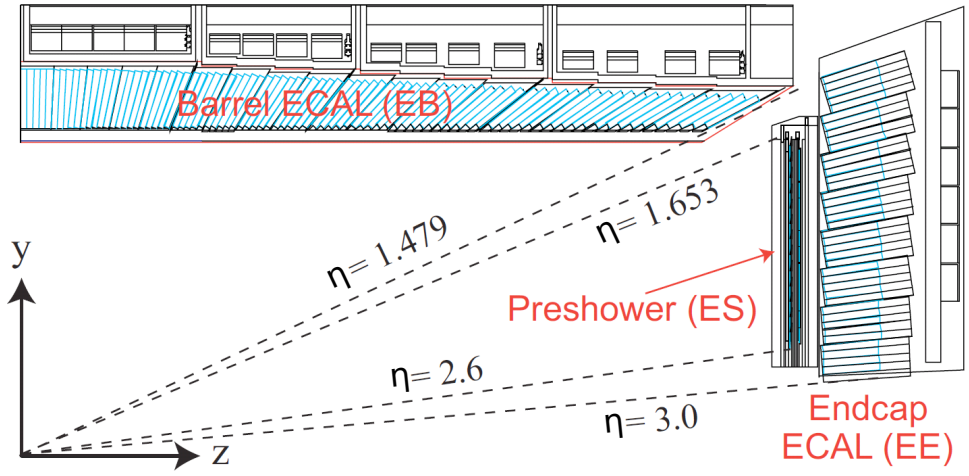

Figure 2.8.: The electromagnetic calorimeter of the CMS experiment consists of the barrel (EB), the preshower (ES) and the endcap (EE) calorimeters. Modified, based on [67].

about 60000 of the crystals are installed in the central part and about 7000 are used in each of the endcaps.

The particles produce an electromagnetic shower, which is a cascade of electrons, positrons, and photons, when passing the ECAL. The electrons and positrons either produce further electron-positron pairs or photons through bremsstrahlung. The pair production continues until the energy of the particles is lower than two times the electron rest mass. The energy of the particles is measured through scintillation, where the crystals produce photoelectrons in proportion to the energy loss of the traversing particle. Operated at a temperature of 18°C, an energy loss of one MeV results in about five photons observed by the photodetector. The dimensions of the ECAL are chosen to enclose the whole energy deposit in the active material, each of the crystals has a length of 230 mm which is about 25 times the radiation length $R_0 = 0.89\,\text{cm}$ of lead tungstate. In order to measure the energy for each collision separately, the crystals are designed to reach a scintillation decay time of about 25 ns.

The relative energy resolution of the ECAL can be described as

$$\left(\frac{\sigma_\text{E}}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \tag{2.8}$$

with the stochastic term S, the noise term N and the constant term C. The stochastic term expresses uncertainties originating from the photon statistics and fluctuations of the shower development, e.g., the track length of individual particles. The noise term considers electronics, digitization and pileup noise. Miscalibration and non-uniformities of the ECAL are characterized by the constant term. For the CMS experiment, the equation reads [50]:

$$\left(\frac{\sigma_\text{E}}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.3\%)^2. \tag{2.9}$$

The energy resolution in the central EB is 1.5%, 3-4% in the outer EB and 4% in the EE obtained from a maximum likelihood fit performed in the di-electron invariant mass range from 89 GeV to 100 GeV [68].

The EB is divided in 360 elements in the ϕ direction and 2×85 elements in the η direction which results in a granularity for each of the crystals of $(\Delta\eta, \Delta\phi) = (0.0174, 0.0174)$. The EB weighs about 70 t and has a crystal volume of about $8\,\text{m}^3$, whereas the EE weighs 24 t and has a crystal volume of about $3\,\text{m}^3$.

In order to identify neutral pions and to distinguish them from photons, the ES is installed in front of the EE. It is built as a sampling calorimeter consisting of two lead radiator planes and two silicon strip

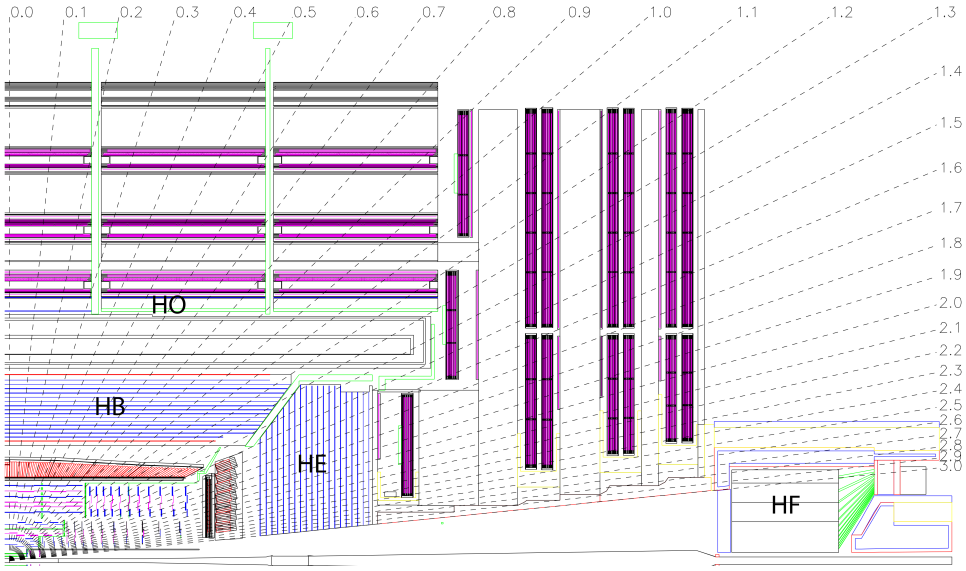

Figure 2.9.: The hadron sampling calorimeter of the CMS experiment consists of the hadron barrel (HB), the hadron endcap (HE), the hadron outer (HO) and the hadron forward (HF) calorimeters. Taken from Ref. [50].

sensor planes. The radiator plane thickness corresponds to two interaction lengths for the first plane and one radiation length for the second.

2.2.4. Hadron Calorimeter

The energy of hadrons is measured with a sampling calorimeter built of alternating layers of brass absorber material and scintillators. The hadron calorimeter (HCAL) is composed of four calorimeters: the hadron barrel (HB), the hadron endcap (HE), the hadron outer (HO) and the hadron forward (HF) calorimeters as shown in Fig. 2.9. It is placed at a radius of 1.77 m to 2.95 m and covers a pseudorapidity of $|\eta| < 5.2$.

The HB calorimeter consists of 36 azimuthal wedges made of brass absorbers composed of 70% copper and 30% zinc with a radiation length of 1.49 cm and an interaction length of 16.42 cm. Orthogonal to the beam axis, the thickness of the calorimeter corresponds to 5.82 interaction lengths. The effective thickness increases with the polar angle covering a pseudorapidity range of $|\eta| < 1.3$. In η direction, the HB calorimeter is divided into 16 segments. Hence, the plastic scintillator has a granularity of $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$.

The HE calorimeter covers a pseudorapidity range of $1.3 < |\eta| < 3$ and has a length of about 10 interaction lengths. Similar to the HB calorimeter, it is built of 36 azimuthal wedges but consists of 17 layers. The segmentation of the HE calorimeter is $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$ for $|\eta| < 1.6$ and $(\Delta\eta, \Delta\phi) = (0.17, 0.17)$ for $|\eta| \geq 1.6$.

The HB calorimeter is not able to stop all of the hadrons, therefore, another layer is needed to measure the whole energy of the hadrons, namely the HO calorimeter. It is placed outside of the solenoid magnet, which is used as another absorber material. The HB calorimeter is composed of five rings along the z axis with twelve identical sectors in ϕ direction, where the central ring consists of two scintillators at radii of 3.82 m and 4.07 m, respectively. These are interlaced with iron pieces. The more forward rings are placed at a radius of 4.07 m and consist of only one scintillator layer. This extends the depth of the HCAL to at least 11.8 interaction lengths. As for the HB calorimeter, the granularity is $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$.

In order to measure the energy of hadrons produced in the forward direction, the HF calorimeter was installed, covering a pseudorapidity up to $|\eta| = 5$. As it is close to the beam, it has to withstand very strong radiation. The average energy deposit per collision is about seven times as high as for the central

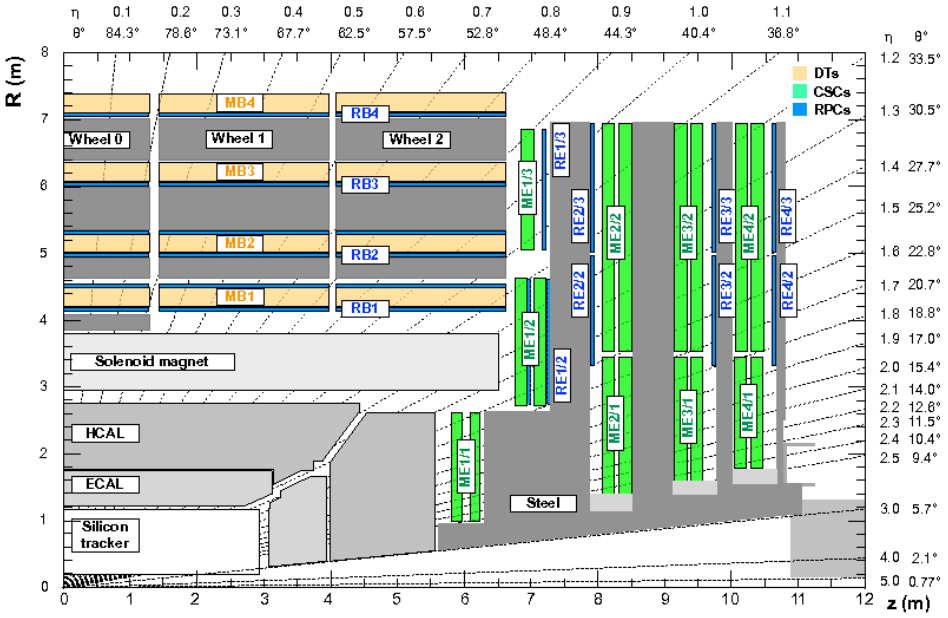

Figure 2.10.: The muon system of the CMS experiment consists of the barrel region made of drift tubes (DTs), the endcaps made of cathode strip chambers (CSCs) and the complementary trigger system built of resistive plate chambers (RPCs). Taken from Ref. [69].

region. The HF calorimeter, placed about 11 m from the interaction point, covers radii from 12.5 cm to 1.3 m and 1.65 m in z direction. In total, the HF calorimeter consists of 13 rings in radial direction divided into 18 elements in ϕ direction. The granularity is $(\Delta\eta, \Delta\phi) = (0.175, 0.175)$.

2.2.5. Solenoid Magnet

The measurement of particle momenta is possible by determining the curvature of the particle tracks in a magnetic field. As the particles travel almost at the speed of light and therefore have high momentum, a strong magnetic field is required for precise momentum measurement. The eponymous part of the CMS detector, the superconducting solenoid magnet, provides a 3.8 T field along the beam axis inside the magnet. The magnet, which is operated at a temperature of 1.8 K, stores 2.6 GJ of energy. On the outside of the magnet, where the muon system is located, the field is oriented in the opposite direction. The cold mass has a diameter of 6 m and a length of 12.5 m and weighs 220 t. The 10 kt iron return yoke, consisting of five wheels and two endcaps, returns the magnetic flux thus providing a stronger field within the detector. The magnetic flux density within the iron yoke is 2 T.

2.2.6. Muon System

The detector systems discussed so far, which are located inside the muon system, absorb almost all the particles produced in the collisions. One of the particles that passes the inner layers without too severe energy loss is the muon as it is a minimum ionizing particle. The muon system itself is embedded in the iron return yoke. Hence, tracks detected outside of the solenoid magnet can be measured and enable the identification of muons. Furthermore, the information of the muon system is used to select events of interest, called triggering, which is even possible without the use of other detector components. Additionally, the muon system is used for the momentum measurement of the muons. This is achieved for pseudorapidities of $|\eta| < 2.4$. The muon measurement is of special interest as a lot of processes, e.g., the Higgs boson decay into a pair of Z bosons, may contain one or more muons in their final state.

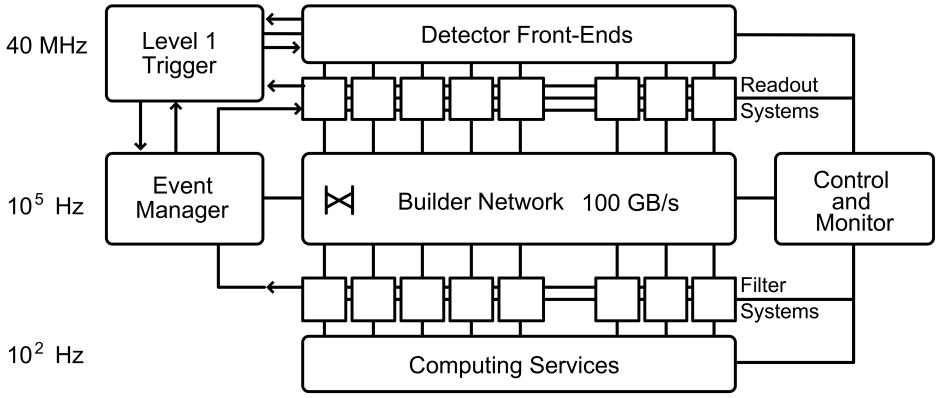

Figure 2.11.: The CMS trigger system consists of two layers: the Level 1 trigger reduces the data rate from 40 MHz to 100 kHz and the HLT reduces the rate to 400 Hz. Taken from Ref. [70].

The muon system as shown in Fig. 2.10, consists of a barrel region and two endcaps built of gaseous particle detectors. The barrel region covers a pseudorapidity range of $|\eta| < 1.2$ using rectangular drift cells and is composed of five wheels of four layers each. The endcaps cover $0.9 < |\eta| < 2.4$ by utilizing cathode strip chambers and consist of four layers. Furthermore, a complementary trigger system is installed using resistive plate chambers (RPCs) covering $|\eta| < 1.6$.

The drift tubes (DTs) used in the barrel region utilize a mixture of 85% argon and 15% carbon dioxide, which results in a maximum drift time of 380 ns. The spatial resolution is 170 µm [50]. For each of the three innermost layers of the barrel region, 60 drift chambers are installed, whereas the outermost layer uses 70.

The endcap region is equipped with 468 cathode strip chambers (CSCs) arranged in groups of 72 or 36 depending on the position in the detector. The chambers are divided in parts covering either 10° or 20° in ϕ direction. As the barrel region and the endcaps overlap for $0.9 < |\eta| < 1.2$, muons produce signals in both parts of the muon system. The CSCs are built of seven cathode panels interleaved with six anode panels and are filled with 40% argon, 50% carbon dioxide and 10% carbon tetrafluoride. This setup results in a measured combined spatial resolution of 80 µm [50].

The RPC system is built from gaseous parallel-plate detectors. Especially for the muon triggering, the RPC provides valuable input as it is able to identify an ionizing event at a time scale shorter than the bunch crossing frequency.

2.2.7. Trigger and Data Acquisition System

The collisions produced by the LHC at a rate of 40 MHz provide far too much data to be thoroughly analyzed. Therefore, a multi-level selection is applied to keep interesting events and reject the remaining. The first link in the chain is the trigger system [70, 71]. It aims to reduce the data to an amount that can be stored. These events are then processed further and different selections are applied in the analyses carried out by the collaboration.

The CMS trigger system is composed of a Level 1 (L1) hardware trigger reducing the event rate to 100 kHz and a high level software trigger (HLT) operated on a cluster of computers reducing the event rate to 400 Hz. An illustration of the CMS trigger system is shown in Fig. 2.11.

The L1 trigger uses information from the calorimeters and the muon system to decide within 4 µs whether an event is of interest or not. The global trigger combines the information of the global muon trigger and the global calorimeter trigger, each combining the information of lower stages.

For the global calorimeter trigger (GCT), the information of the regional calorimeter trigger (RCT) is combined. The RCT outputs electron and photon candidates, as well as the regional transverse energy.

The GCT then processes the electron and photon candidates further, identifies jets and computes the missing transverse energy. The GCT covers a pseudorapidity range of $|\eta| < 5$.

In the global muon trigger, the information of the DTs, CSCs and RPCs of the muon system is combined. For the DT and CSC part, the trigger identifies tracks by evaluating the hit information and provides the input for the regional track finders. The RPC signals are used in the pattern combination trigger to identify muon candidates as well. The track finders and the pattern comparator provide muon candidates to the global muon trigger.

The global trigger realizes several sets of selection requirements for different signatures, which are utilized by the high level trigger (HLT).

The HLT reconstructs actual physics objects, e.g., electrons, muons and jets, and selects events based on this information. There are several triggers implemented aiming at different physics processes. The triggers are implemented as trigger paths, which are a sequence of reconstruction and selection algorithms increasing in complexity executed on a processor farm. The per event time budget is 175 ms, meaning that the decision of keeping the event has to be made within that time period otherwise not all events passing the L1 trigger can be processed.

Events that fulfill at least one of the trigger paths requirements are transfered to the CMS Tier-0 center, described in the next section. The rate of the HLT is limited by this transfer process, hence there are also trigger paths that do not store all events fulfilling the requirements but only a fraction of it.

2.2.8. Computing

The data passing the trigger system has to be stored and physics objects have to be reconstructed. Furthermore, the data is made available to the collaboration. Additionally, simulated samples are produced, reconstructed and distributed. This is realized through the Worldwide LHC Computing Grid (WLCG) [72, 73], which is a collaboration between the LHC experiments and computing centers from all over the world.

The WLCG is organized in a hierarchal structure with four levels called Tiers: There is one Tier-0 center at CERN storing the raw data and performing a first reconstruction of it. There are several Tier-1 centers, one of which is located at the KIT, Karlsruhe, all together jointly holding a copy of the raw and reconstructed data. The Tier-1 centers perform a second-pass reconstruction of data and are also responsible for the reproduction of the reconstruction when improved algorithms and calibrations are available. Furthermore, they provide data to the Tier-2 centers. Specific analyses are carried out on the Tier-2 centers, which also store relevant data samples transferred from Tier-1 centers. Further duties are the production of simulated samples and offering resources for specialized tasks such as detector alignment or calibration. Tier-3 centers provide resources for local analysis tasks with moderate requirements.

3. Statistical Methods and Machine Learning

Particle physics, as physics in general, aims to deliver quantifiable results. For example, the observation of a new process is defined by how unlikely the observation is to be explained by the background model. Conventionally, an observation is claimed, if the deviation from the background model is larger than five standard deviations. Similarly, if the respective process has not been observed yet, an upper limit on the cross section at a certain confidence level can be set. The determination of the significance and upper limits is only possible by introducing statistical methods as described in the following. For this purpose, the implementation of the COMBINE package [74, 75] based on ROOFIT [76] is used. One or more observables have to be chosen in order to apply these methods.

In many modern analyses, machine learning techniques are applied to obtain a composite variable as a combination of observables, leading to a better discrimination between signal and background processes as compared to a classification using a single observable. Furthermore, machine learning is used for jet tagging or energy regression. Boosted Decision Trees (BDTs) and (Deep) Neural Networks (DNNs) are the most common approaches in CMS, both of which are used in this thesis.

3.1. Probability Density Function

Observables in particle physics are usually variables that follow an underlying distribution called probability density function (pdf), denoted as $f(x)$ [77]. It is related to the probability of measuring a variable x in a range of $a \leq x \leq b$

$$P(a \leq x \leq b) = \int_a^b f(x)\mathrm{d}x, \tag{3.1}$$

where

$$\forall x \in \mathbb{R} : \; f(x) \geq 0, \tag{3.2}$$

and

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1. \tag{3.3}$$

The expectation value of the variable following the pdf is defined as:

$$\langle x \rangle = \int_{-\infty}^{\infty} x f(x)\mathrm{d}x. \tag{3.4}$$

The variance $V[x]$ is given by

$$V[x] = \langle (x - \langle x \rangle)^2 \rangle = \int_{-\infty}^{\infty} (x - \langle x \rangle)^2 f(x)\mathrm{d}x, \tag{3.5}$$

and the standard deviation σ is linked to the variance by:

$$\sigma = \sqrt{V[x]}. \tag{3.6}$$

In particle physics, a lot of measurements are counting experiments, where the number of entries per bin, called events, are randomly distributed according to an underlying Poisson distribution

$$f_{\text{Poisson}}(x|\mu) = \frac{\mu^x \mathrm{e}^{-\mu}}{x!}, \tag{3.7}$$

with the mean μ and the standard deviation $\sqrt{\mu}$. In case of large μ, the Poisson distribution can be approximated as a Gaussian distribution

$$f_{\text{Gauss}}(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{3.8}$$

where μ is the mean, and σ is the standard deviation.

3.2. Maximum Likelihood Method

Hypotheses are usually described by parameters $\vec{a}$. The estimation of these parameters of a pdf $f(\vec{x}_i|\vec{a})$ in such a way that the measured data $\vec{x}_i$, where the components are the observables for each data point i, is described best, is the essential task of an analysis. Maximizing the likelihood function

$$L(\vec{x}_i|\vec{a}) = \prod_i^N f(\vec{x}_i|\vec{a}), \tag{3.9}$$

is one way to determine this point in the parameter space, which is called maximum likelihood estimate. The parameters of the pdf usually correspond to parameters in the underlying theoretical model. In this thesis, the signal strength modifier μ is the parameter of interest, which is defined as the ratio between the number of observed and expected signal events. The data, as well as the simulated events, are represented in binned histograms leading to Poisson distributed probabilities for the number of events per bin. Hence, the binned likelihood is defined as

$$L(\vec{n}|\mu,\vec{\theta}) = \prod_i^{n_{\text{bins}}} \frac{(\mu s_i(\vec{\theta}) + b_i(\vec{\theta}))^{n_i}}{n_i!} \cdot \mathrm{e}^{-(\mu s_i(\vec{\theta}) + b_i(\vec{\theta}))}, \tag{3.10}$$

with the number of observed events n_i per bin i, the number of expected signal and background events in the i-th bin s_i and b_i and the number of bins n_{bins}. Nuisance parameters $\vec{\theta}$ account for systematic uncertainties affecting the number of expected signal and background events. The product of the likelihood can become very small. Therefore, the negative logarithm of the likelihood function (NLL) is used

$$-\ln L(\vec{x}_i|\vec{a}) = -\ln \prod_i^N f(\vec{x}_i|\vec{a}) = -\sum_i^N \ln f(\vec{x}_i|\vec{a}), \tag{3.11}$$

which results in larger values. This simplifies the computation of the parameters.

3.3. Systematic Uncertainties

In an experimental setup, none of the parameters can be measured with infinite precision. Therefore, systematic experimental uncertainties have to be considered. Furthermore, theoretical uncertainties are taken into account. The experimental uncertainties can be determined from dedicated measurements, whereas the theoretical uncertainties are estimated by varying the parameters of the respective model. As they modify the expected number of events, the uncertainties are implemented as nuisance parameters into the likelihood function [78]. In principle, there are two types of uncertainties: uncertainties affecting the normalization of the prediction and uncertainties affecting the shape of the prediction.

Rate uncertainties are implemented with a log-normal prior

$$\pi(n) = \frac{1}{\sqrt{2\pi}n\sigma_n} \cdot \mathrm{e}^{-\frac{(\ln n - n_0)^2}{2\sigma_n^2}}, \tag{3.12}$$

where n is the number of observed events, n_0 is the mean number of events and σ_n is the corresponding uncertainty.

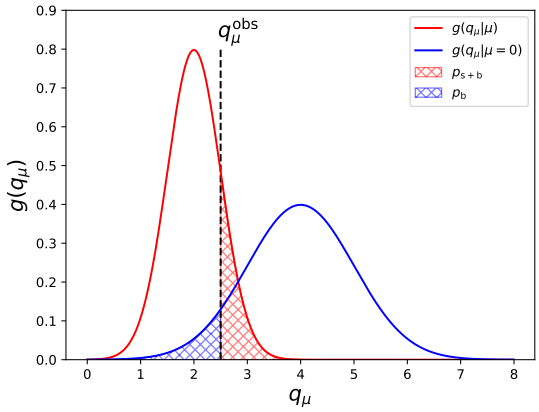

Figure 3.1.: Illustration of the pdfs for background and signal+background hypotheses: the area under the blue curve from $-\infty$ to q_μ^{obs} corresponds to p_b, whereas the area under the red curve from q_μ^{obs} to $+\infty$ corresponds to p_{s+b}.

For shape uncertainties, a template morphing method is implemented. This requires histograms not only for the nominal but also for systematically varied expectations for each of the uncertainties. Therefore, the whole analysis is rerun with systematically varied observables such as the jet energy. In the analyses presented in this thesis, the variations correspond to one standard deviation in upward and downward direction, respectively. A polynomial function is used to interpolate and extrapolate between the templates [79], leading to a change of the shape and most of the time also of the rate of the expectation.

Uncertainties introduced by the finite size of the simulated data samples are implemented using the Barlow-Beeston light method [80].

3.4. Exclusion Limit Calculation

The search for new processes or new particles aims at a result that quantifies the probability of the existence given certain signal and background models. Therefore, a hypothesis test is implemented to decide between two hypotheses: the background hypothesis containing only experimentally established processes and the signal+background hypothesis. In order to quantify the signal, the signal strength μ is introduced, which can obtain any value in general. A value of $\mu = 0$ corresponds to the background hypothesis and $\mu = 1$ corresponds to the signal+background hypothesis. A test statistic q_μ [81] is used to evaluate both hypotheses. The Neyman-Pearson lemma [82] states the ratio of the two likelihoods of the hypotheses to be the most powerful test statistic following an underlying pdf g if no nuisance parameters are present. The ratio

$$q_\mu = -2 \ln \frac{L(\mu, \hat{\theta}_\mu)}{L(\hat{\mu}, \hat{\theta}_{\hat{\mu}})} \tag{3.13}$$

of the likelihoods including nuisance parameters is still a valid and powerful test statistic. The nuisance parameters $\hat{\theta}_\mu$ maximize the likelihood for a given μ, whereas $\hat{\theta}_{\hat{\mu}}$ and $\hat{\mu}$ maximize the likelihood globally. The p-value for the background p_b is defined by

$$1 - p_b = \int_{q_{\mu^{\text{obs}}}}^{\infty} g(q_\mu|\mu = 0, \theta) \mathrm{d}q_\mu, \tag{3.14}$$

and for the signal+background hypothesis it is given by

$$p_{\text{s+b}} = \int_{q_{\mu^{\text{obs}}}}^{\infty} g(q_\mu|\mu,\theta)\mathrm{d}q_\mu, \tag{3.15}$$

which can be computed from the underlying pdf g of the test statistic q_μ. An illustration of the pdfs of the test statistic and their connection to the p-values can be found in Fig. 3.1. For a given signal strength, the p-value $p_{\text{s+b}}$ (p_{b}) represents the probability of observing a value of $q_\mu \geq q_\mu^{\text{obs}}$ under the signal+background (background) hypothesis. Whether to keep or reject a hypothesis is decided based on a significance level α chosen before the measurement is performed [83]. In particle physics, $\alpha = 5\%$ is chosen for most analyses. The signal+background hypothesis is accepted if $\alpha > p_{\text{b}}$ and rejected otherwise. From that an upper limit on the signal strength at a confidence level (C.L.) of $1-\alpha$ can be computed. When using p-values as defined above, downwards fluctuations of the background can result in an unjustified signal rejection. A more robust approach to handle signal processes that are rare compared to the background processes is the CL_S limit [84]. Here, the p-value is replaced by CL_S defined as:

$$\text{CL}_\text{S} = \frac{p_{\text{s+b}}}{1-p_{\text{b}}}. \tag{3.16}$$

The observed limit is the limit obtained by evaluating the measured data. In order to optimize analysis techniques, different approaches are tested and the corresponding exclusion limits are calculated. When evaluating the analysis methods on data, a bias can be introduced as the optimization might target statistical fluctuations instead of real physical effects. Therefore, the expected limit obtained from toy experiments is used alternately to improve the method. Once the analysis workflow is fixed, the observed limits are calculated and can be compared to the expectation.

Multiple toy experiments are usually performed to evaluate the median of the expected test statistic including the uncertainties. As this requires a lot of toy experiments and therefore a lot of computing power, asymptotic limits using an Asimov data set [81] are utilized instead, based on the theorems of Wald and Wilks [85, 86].

In case of an excess of data over the background prediction, a significance

$$Z = \phi^{-1}(1-\text{CL}_\text{S}), \tag{3.17}$$

for the signal is calculated, where ϕ returns the quantile of the Gaussian distribution. Z is the number of standard deviations of a Gaussian distribution. In particle physics, the consensus is to state a discovery if $Z \geq 5$ is observed.

3.5. Boosted Decision Trees

In particle physics, it is important to distinguish between signal and background processes. When studying processes with a comparatively small cross section, a discrimination based on a single variable, such as the invariant mass of the decay products, is not sufficient. Hence, decisions based on multiple variables, called multivariate methods, are implemented. Boosted Decision Trees (BDTs) are one of these methods. Here, the BDTs as implemented in the Toolkit for Multivariate Analysis (TMVA) [87] as part of the ROOT framework [88] are described. BDTs are used to combine the discrimination power of several variables into one single output variable giving information on whether the event is signal- or background-like.

A simple decision tree (DT) consists of a series of binary decisions based on a set of different input variables $\vec{x}$ dividing the data set into many regions, called leaves. The starting point of each tree is called root node. It contains all events, which are split into the regions as follows. Starting from the root node, a binary decision is made based on one variable, leading to two daughter nodes containing the events with a value of the considered variable larger or smaller than the cut value, respectively. Further, discrimination within a daughter node is achieved by applying another cut on a variable using the

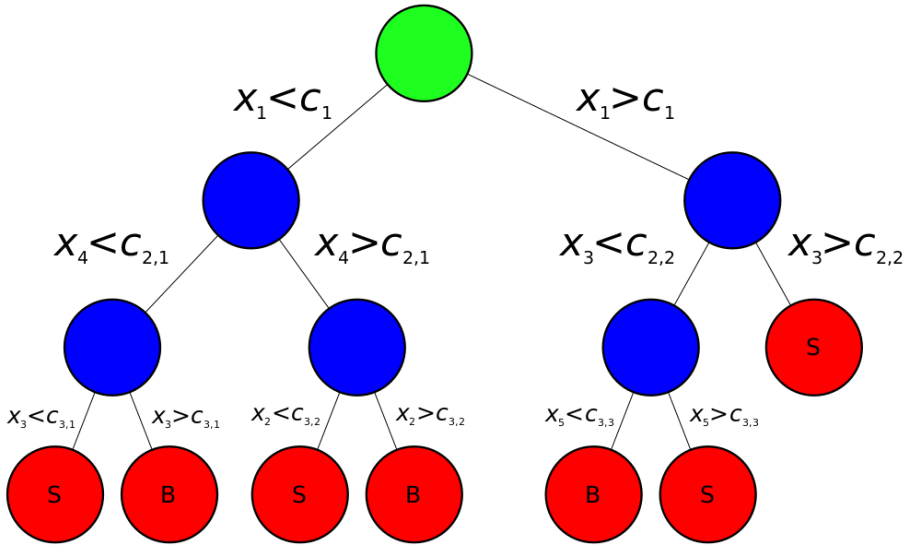

Figure 3.2.: Illustration of a decision tree: starting from the root node at the top (green), the events are classified based on cuts c on different variables x and end up in a leaf (red). Each of the leaves is either labeled with signal (S) or background (B) depending on the label of majority of events in the respective leaf.

former daughter node as the starting point, called mother node, for the next decision. A depiction of a DT can be found in Fig. 3.2. The depth of the tree is given by a stop criterion, e.g., the number of events in the node, a predefined maximum depth or if no further discrimination between signal and background can be achieved. The nodes at the lowest level are called leaves, where each of the leaves is classified as signal or background depending on which of the categories is dominating. In particle physics, truth information obtained from simulation is used to provide labeled events in the training process. All cuts and the weights of the BDTs are calculated in the training as described in the following. The purity of a node is defined as

$$P = \frac{\sum_i^{N_\mathrm{s}} w_{i,\mathrm{s}}}{\sum_i^{N_\mathrm{b}} w_{i,\mathrm{b}} + \sum_i^{N_\mathrm{s}} w_{i,\mathrm{s}}}, \tag{3.18}$$

where N_s (N_b) is the number of signal (background) events and $w_{i,\mathrm{s}}$ ($w_{i,\mathrm{b}}$) are the corresponding event weights. The cut value is chosen to maximize the separation

$$S = G_\mathrm{mother} - G_\mathrm{daughter1} - G_\mathrm{daugther2}, \tag{3.19}$$

between signal and background, where G is the Gini coefficient of the respective nodes [89]

$$G = P(1-P) \sum_{i=1}^{N_\mathrm{events}} w_i, \tag{3.20}$$

where P is the purity of the node, w_i is the weight of the events and N_events is the total number of entries. For unweighted events, the maximal Gini index is $G = 0.25$ for $P = 0.5$, which is equal to no separation. A perfect separation results in $G = 0$.

When utilizing only one decision tree, the result turns out to be very sensitive to statistical fluctuations of the data. In order to obtain a more robust and more accurate classifier, a number of N_trees trees are trained iteratively using the AdaBoost algorithm [90, 91]. The final output is then composed of the output of the individual trees, where each tree is only a weak classifier.

The boosting algorithm aims at events that have been classified wrongly in the previously trained tree and gives them a higher weight in the next iteration of the training, called boost weight. The weights

of the events for the training of the first tree are set to $1/N_{\text{events}}$. The weights $w_{i,n}$ of the n-th tree are connected to its misclassification rate

$$r_n^{\text{mis}} = \frac{\sum_i w_{i,n}\delta_{i,n}^{\text{mis}}}{\sum_i w_{i,n}}, \tag{3.21}$$

where $\delta_{i,n}^{\text{mis}}$ is 0 for correctly classified events and 1 for wrongly classified events. The boost weight of tree n is given by:

$$\alpha_n = \ln \frac{1 - r_n^{\text{mis}}}{r_n^{\text{mis}}}. \tag{3.22}$$

The weights for the following tree $n+1$ of a given event i are:

$$w_{i,n+1} = w_{i,n} \cdot \mathrm{e}^{\alpha_n \delta_{i,n}^{\text{mis}}}. \tag{3.23}$$

The sum of the weights is normalized to 1 after training all trees and the weights, as well as the cuts, are stored. The BDT output $\hat{y}$ for event i is given by

$$\hat{y}(\vec{x}_i) = \sum_n^{N_{\text{trees}}} \alpha_n \mathrm{T}_n(\vec{x}_i), \tag{3.24}$$

where $\mathrm{T}_n(\vec{x}_i)$ is the classification output of tree n for event i. The output of each tree T_n is either 1 in case of classifying the event as signal or -1 in case of classifying it as background. The final BDT output lies between -1 and 1.

The variables used for the BDT can be ranked according to their importance in the training. The importance is derived by counting the number of occurrences in the split decisions and weighting for the separation gain and the number of events in the node [92].

3.6. Artificial Neural Networks

Artificial neural networks (ANNs) map an input vector $\vec{x}$ containing the input variables for an event to an output vector $\vec{y}(\vec{x})$, containing one or more variables $\hat{y}_i$ that should be derived from the input. Neural networks are used for a huge variety of tasks, e.g., image, video, and audio processing, regression, classification, or translation of texts [93]. The structure connecting the input and output layers is inspired by the human brain: an ANN consists of one or more layers made up of neurons. In a fully connected feed forward neural network as shown in Fig. 3.3, the neurons of each layer are connected to all neurons of the preceding layer.

The inputs x_i are weighted with weights w_i and processed by a non-linear activation function a in order to obtain the output $\hat{y}_j$ of the neuron. Additionally, a bias w_0 can be introduced. The output of the neuron then yields

$$\hat{y}_j = a\left(\sum_i w_i \cdot x_i + w_0\right), \tag{3.25}$$

which can be rewritten as

$$x_i^{(k)} = a_i^{(k)}\left(\sum_{j=1}^{N(k-1)} w_{ij}^{(k-1)} x_j^{(k-1)} + w_{i0}^{(k-1)}\right), \tag{3.26}$$

for the output of neuron i in layer k with the number of neurons $N(k)$ in layer k, the weights $w_{ij}^{(k-1)}$ of the neurons in layer k multiplied by the output of neuron j in layer $(k-1)$, and its bias w_{i0}^{k}. The weights

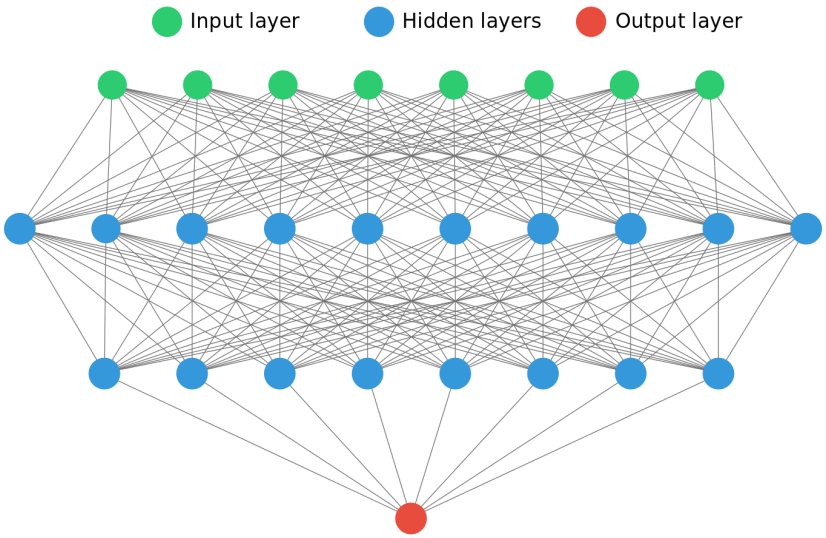

Figure 3.3.: Illustration of a fully connected ANN: starting from the input layer at the top (green) all nodes are connected to the nodes of the first hidden layer (blue). All nodes of the first hidden layer are connected to the second hidden layer and all nodes of this layer are connected to the output node (red).

and biases are summarized in a matrix **w**. The output of an ANN $\hat{y}$ with one hidden layer is then given by

$$\hat{y} = a\left(\sum_{j=1}^{N_{\text{hidden}}} w_j \cdot a\left(\sum_{i=1}^{N_{\text{input}}} w_{ij}x_i + w_{i0}\right)\right), \tag{3.27}$$

where N_{input} (N_{hidden}) is the number of input (hidden) nodes. The activation function is chosen depending on the task of the network. The most commonly used functions are shown in Fig. 3.4. One of these activation functions is the rectified linear unit (ReLU):

$$a_{\text{ReLU}}(x) = \begin{cases} x & \text{for } x \geq 0 \\ 0 & \text{for } x < 0. \end{cases} \tag{3.28}$$

It has a disadvantage as it looses information about values smaller than zero. This can be fixed by introducing the leaky ReLU

$$a_{\text{leakyReLU}}(x) = \begin{cases} x & \text{for } x \geq 0 \\ \beta x & \text{otherwise,} \end{cases} \tag{3.29}$$

which allows for information with $x < 0$ to be leaked into the output, suppressed by a parameter β. Hence, usually the leaky ReLU is preferred over the ReLU activation function. Another variation of the ReLU is the exponential linear unit (ELU)

$$a_{\text{ELU}}(x) = \begin{cases} x & \text{for } x \geq 0 \\ \alpha(e^x - 1) & \text{otherwise,} \end{cases} \tag{3.30}$$

with an additional parameter $\alpha > 0$. This can result in faster converging trainings for specific cases. Furthermore, the sigmoid function

$$a_{\text{sigmoid}}(x) = \frac{1}{1 + \mathrm{e}^{-x}}, \tag{3.31}$$

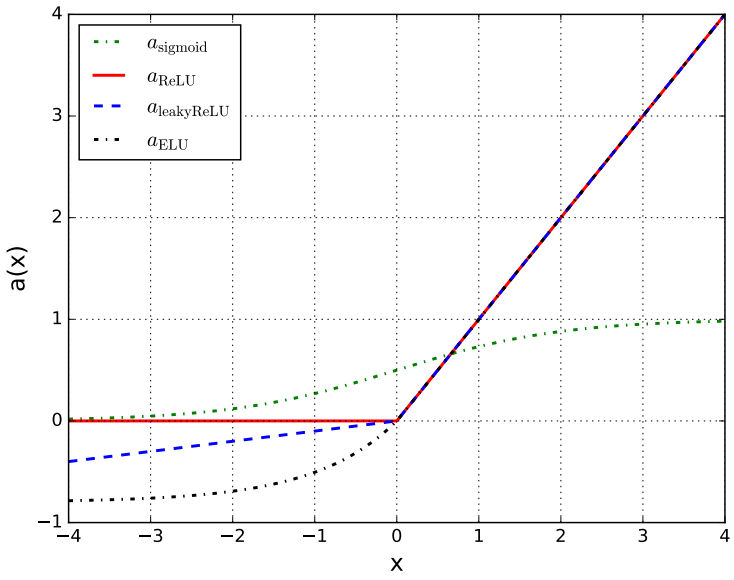

Figure 3.4.: Illustration of activation functions: the ReLU activation function is shown in red, the leaky ReLU in blue, the ELU in black and the sigmoid in green.

is another activation function commonly used. The output of this function is constrained to values between -1 and 1, which makes it an ideal candidate for the activation of output neurons in a classification ANN resulting in the output $\vec{y}(\vec{x})$.

The output $\hat{y}(\vec{x})$ of the ANN is meant to approximate the target $\vec{t}(\vec{x})$. In case of a classification task, the target is the label of the dataset. For regression tasks, the target is the respective value of the variable. The weights, as well as the biases, are determined in the training of the ANN. The training is performed on a dataset, where the target vector is known for each of the events. The discrepancy of the target and the ANN output is described by a loss function $\lambda(\mathbf{w})$, which is minimized in the training.

As for the activation functions, there is a large variety of loss functions optimized for the specific problem faced. Often, the binary cross entropy

$$\lambda_{\text{cross entropy}} = -\frac{1}{n}\sum_{i=0}^{N} y_i \log(\hat{y}_i) + (1-t_i)\log(1-\hat{y}_i), \tag{3.32}$$

is used for classification problems, where N is the number of output nodes. For regression tasks, the χ^2 loss

$$\lambda_{\chi^2} = \frac{1}{2}\sum_{i=0}^{N}(t_i - \hat{y}_i)^2 \tag{3.33}$$

is commonly used. A more sophisticated function is the Huber loss

$$\lambda_{\text{Huber}} = \sum_{i=0}^{N} \begin{cases} \frac{1}{2}(t_i-\hat{y}_i)^2 & \text{for } |t_i-\hat{y}_i| \leq \delta \\ \delta|t_i-\hat{y}_i| - \frac{1}{2}\delta^2 & \text{otherwise} \end{cases}, \tag{3.34}$$

with the tuneable hyperparameter δ. It is less sensitive to outliers than the squared error loss and hence can be used to obtain stable results without special outlier treatment.

After defining the layout of the network by choosing the architecture, the activation functions and the loss function, the network is trained. Firstly, all weights are set randomly. The loss function is then

minimized by a dedicated algorithm such as the iterative stochastic gradient descent. It calculates the gradient of the loss function with respect to the weights and adjusts the weights by applying

$$\mathbf{w} \leftarrow \mathbf{w} + \Delta\mathbf{w} = \mathbf{w} - \eta \nabla_{\mathbf{w}} \lambda(\mathbf{w}), \tag{3.35}$$

where η is the learning rate defining the magnitude of the change from one iteration to the next. In general, a low learning rate requires the algorithm to perform more iterations but allows for a more precise computation of the minimum. Nevertheless, it might cause the algorithm to find a local minimum instead of the global minimum. A vanishing gradient $\nabla_{\mathbf{w}} \lambda(\mathbf{w})$ either occurs at a minimum of the loss function or at a saddle point. Ending at a saddle point or a local minimum is not intended as this leads to suboptimal results. As the global minimum is very hard to find, at least finding a local minimum close to the global minimum is desirable. The computation of the gradient is performed by backpropagation [94].

In order to avoid the minimization to converge at a saddle point, the training dataset can be split into subsets, called batches. The updating of the weights is then performed for each of the batches subsequently. The handling of the whole dataset, i.e., all batches, is called one epoch. After each epoch, the batches are newly drawn from the whole dataset. A more sophisticated approach based on stochastic gradient descent is the Adam optimizer [95], which also takes gradients of previous iterations into account. Hence, the updating of the weights is changed to

$$\mathbf{w} \leftarrow \mathbf{w} - \frac{\eta}{\sqrt{\hat{v}} + \epsilon} \cdot \hat{\mathbf{m}}, \tag{3.36}$$

with the dampening parameter ϵ and momentum vectors:

$$\hat{v} = \frac{v}{1 - \beta_1} \tag{3.37}$$

and

$$\hat{\mathbf{m}} = \frac{\mathbf{m}}{1 - \beta_2}. \tag{3.38}$$

β_1 and β_2 are the decay rates that regulate the influence of the previous gradients. The initial values of the momentum vectors are zero for both. In each iteration they are updated by:

$$v \leftarrow \beta_1 v + (1 - \beta_1)(\nabla_{\mathbf{w}} \lambda(\mathbf{w}))^2 \tag{3.39}$$

and

$$\mathbf{m} \leftarrow \beta_2 \mathbf{m} + (1 - \beta_2) \nabla_{\mathbf{w}} \lambda(\mathbf{w}). \tag{3.40}$$

The recommended parameter set is $\beta_1 = 0.999$, $\beta_2 = 0.9$, $\epsilon = 10^{-8}$, and $\eta = 0.001$ [95].

In order to obtain a more robust ANN, a dropout regularization [96] can be introduced, which deactivates a percentage of the neurons of a certain layer for the updating of the weights. The deactivated neurons are chosen randomly in each iteration. Furthermore, the batch normalization method [97] is often utilized. It introduces another layer, the batch normalization layer, that rescales the output y_i of each neuron for a batch B

$$\hat{y}_i = \sigma_i \cdot \frac{y_i - \mathrm{E_B}(y_i)}{\sqrt{\mathrm{V_B}(y_i) + \epsilon}} + \mu_i, \tag{3.41}$$

with the mean value $\mathrm{E_B}$ and the Variance $\mathrm{V_B}$ of y_i for batch B. Additionally, a constant value ϵ to avoid divisions by zero, and the adjustable parameters of the model μ_i and σ_i to rescale the distribution to a certain mean and standard deviation, respectively, are introduced. The rescaling is introduced to avoid divergences and hence to ensure a convergence of the training. Consequently, only two values have to be changed during the training instead of shifting a large number of weights. This enables a faster training with a network output with lower dependence on single neurons.

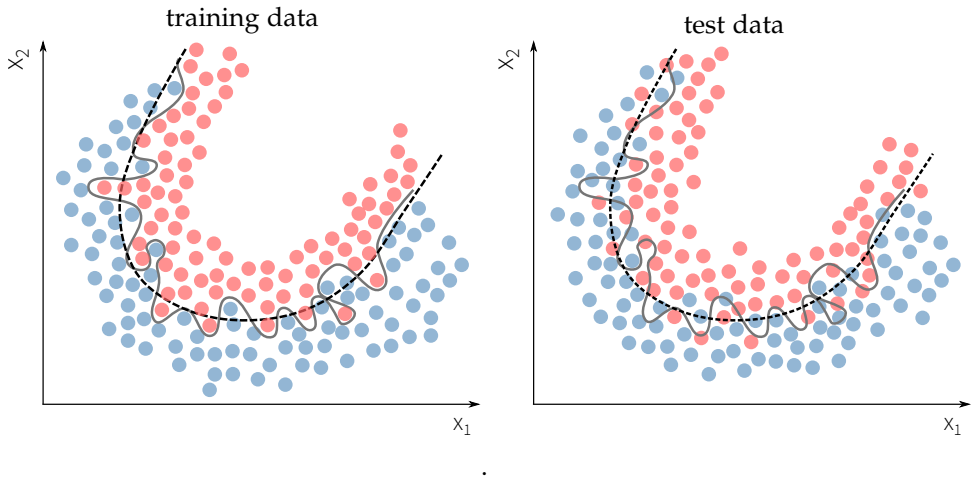

Figure 3.5.: The left (right) figure shows the distribution of the training (test) datasets for signal in red and background in blue in the two variables x_1 and x_2. The dashed line illustrates the general underlying distribution of the datasets and results in a good separation on both datasets. The solid line allows for perfect discrimination on the training dataset by learning statistical fluctuations but results in a less powerful discrimination on the test dataset. This is called overtraining. Taken from Ref. [98]

3.7. Validation of Machine Learning Methods

The methods discussed above need to be validated in order to be used in a physics analysis. Firstly, the modeling of the input variables has to be checked. Therefore, the distribution of the variable for simulated events and data has to be compared. Only variables with reasonable agreement are used as input.

In addition, this possible physics specific issue, an evaluation of the method itself has to be performed. As the datasets used for the training are of finite size, they suffer from statistical fluctuations. The aim of a machine learning method is to learn the underlying connections between the input and the target but not the fluctuations. In case this is achieved, the performance of the classifier is equal on the training dataset and on an independent test dataset. If the performance on the test dataset is significantly worse, the classifier is overfitting, also called overtraining. An illustration of overtraining can be found in Fig. 3.5.

To have a more systematic approach to investigate overfitting the dataset with known labels is split into a dataset for training and testing. The training is performed on the train dataset. Subsequently, the classifier is applied to the test dataset and the output for the test and train sets is compared for the different labels. A Kolmogorov-Smirnov test [99, 100] can be implemented to obtain a measure for the agreement between the two distributions.

Overtraining can be avoided by reducing the ability of the classifier to learn too many features. This can be achieved by reducing the number of free parameters, e.g., in a neural net, the number of nodes or layers or by introducing a dropout regularization. For BDTs, a reduction of the number of trees or the depth of the trees is able to mitigate overtraining effects.

4. Simulation and Reconstruction of Events

As described in the previous chapter, the recorded data is compared to theory predictions. The detectors are not able to measure the particles of interest directly. Therefore, simulations of the predicted processes, consisting of a simulation of the hard scattering process and the parton shower, are performed. Subsequently, a simulation of the detector response to the scattering process is carried out. The simulated samples now have the same format as the recorded data. For simulation and data, the detector signals are translated into particles including information about their kinematic properties. Depending on the particle, different reconstruction methods are implemented.

4.1. Event Simulation

The simulation of events consists of several steps as illustrated in Fig. 4.1. With input from the PDFs as described in Section 1.3, the hard scattering subprocess is simulated. The quarks and gluons produced by the hard subprocess are not stable. Hence, they undergo a subsequent parton shower followed by the hadronization. Finally, the decay of unstable hadrons is simulated.

4.1.1. Proton-Proton Scattering Process

The study of the hard scattering process is the actual target of most analyses. The simulation of the hard scattering process depends on the four-vectors of the initial partons taking part in the interaction. Several authors provide parton distribution functions used to describe the kinematic distributions of these partons. The NNPDF set [30, 102, 103] is one of the most commonly used in the CMS collaboration.

There are two possibilities how to treat bottom quarks in the initial state as their mass is larger than the mass of the proton. The proton can either be modeled in the four flavor scheme (4FS) or in the five flavor scheme (5FS). The 4FS assumes the proton to contain quarks of the four lightest flavors, up, down, strange and charm, whereas the 5FS also includes bottom quarks. For processes with bottom quarks in the initial state, a description in the 4FS requires a gluon to split into a pair of bottom quarks, one of which is then the initial quark of the process. This gluon splitting is described as part of the hard scattering process. Therefore, the kinematic properties of the bottom quark are described more accurately. Nevertheless, the 5FS results in easier calculations as there is one less final-state quark. Sometimes choosing one of the schemes removes interferences between two processes, e.g., between the tHW and ttH processes by choosing the 5FS for the tHW process as this results in different final states. This is described in more detail in Section 5.1.

In order to compute the matrix element and therefore the cross section of a certain process, a perturbation series corresponding to all possible Feynman diagrams, which represent the process, has to be evaluated. As the number of Feynman diagrams is infinite, an exact calculation is not possible. Nevertheless, very good approximations can be made. Therefore, the Feynman diagrams are ordered by the number of additional vertices involving couplings of the strong interaction. Calculations considering diagrams without additional vertices are called leading order (LO) with the corresponding leading order cross section σ_{LO}. Each subsequent order then adds one additional vertex, resulting in real emissions or virtual corrections. Since the coupling constant α_{S} is smaller than one for high momentum transfer, the cross section can be expressed as a perturbation series in α_{S}:

$$\sigma = \sigma_{\mathrm{LO}} \left(1 + \sum_i \sigma_i \left(\frac{\alpha_{\mathrm{S}}}{2\pi} \right)^i \right). \tag{4.1}$$

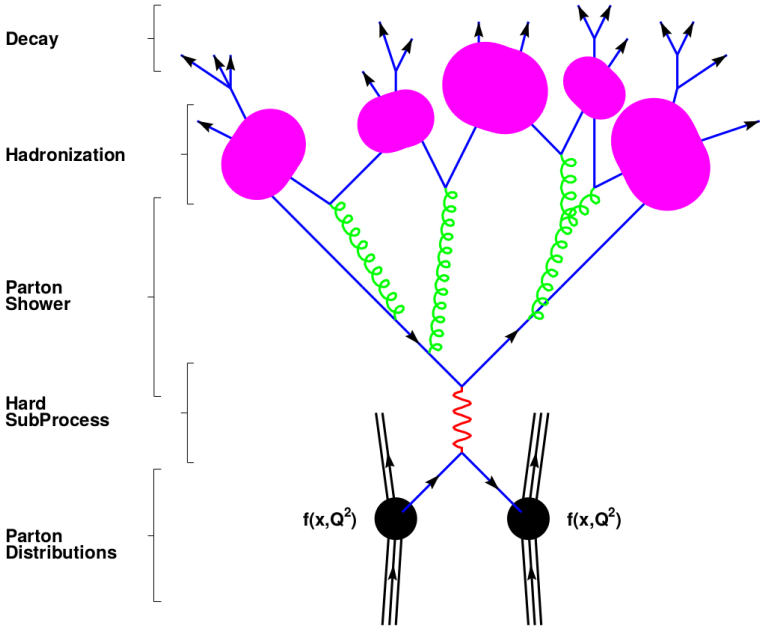

Figure 4.1.: Illustration of the event simulation workflow: starting with partons sampled from a pdf, the hard subprocess is simulated. The final decay products are simulated based on parton shower and hadronization models. Modified, based on: Ref. [101].

Hence, the higher the order the more precise is the calculated cross section. For most of the processes, calculations at next-to-leading order (NLO) or next-to-next-to-leading order (NNLO) are available. Since the strong coupling constant is significantly larger than the coupling constant of the electromagnetic interaction, electromagnetic corrections are of smaller impact at hadron colliders for most processes.

4.1.2. Parton Shower and Hadronization

Most of the final-state particles produced in the hard subprocess are instable and the quarks are not color-neutral. Therefore, they cannot propagate freely but form bound states. In a first step, the parton shower is performed, adding further gluon radiation to the hard scattering process. In general, the gluons can couple to any color-charged particle. The gluon radiation is grouped into initial-state radiation (ISR), where the gluon is radiated off by a parton in the initial state, and final-state radiation (FSR), where the gluon is radiated off by a final-state particle. As subsequent radiations are of lower energy compared to the couplings considered in the matrix element, the coupling constant is larger. Therefore, a perturbative approach is insufficient to model the hadronization. Instead, the parton shower is described by the Altarelli-Parisi splitting functions [32] and Sudakov form factors [104]. Improved accuracy can be reached by simulating several matrix elements with different additional emissions and merging them. These are matched to the parton shower, e.g., using FxFx merging [105].

Confinement forbids the color-charged particles produced in the hard scattering process and in the parton shower to propagate freely. Hence, the particles form color-neutral bound states of two or three quarks, i.e., mesons and baryons. This takes place at relatively low energies, leading to a large strong coupling constant. Therefore, models from first principles are not feasible, instead, phenomenological models are used. The Lund string model [106] is one of the most common models, describing gluons as field lines. The self-interaction of the gluons is described by tubes of field lines, which are formed between color-charged particles. New quark-antiquark pairs are created if the energy stored in the field lines is large enough, i.e., they are separated from each other. This process continues until color-neutral

states are reached. Nonetheless, these are not stable particles in general. Consequently, the decay of these particles is simulated as well.

4.1.3. Underlying Event and Pileup

One parton of each proton participates in the hard scattering process. The remaining partons are usually not of direct interest for an analysis. Nevertheless, they leave traces in the detector and have to be considered in the simulation. This effect is called underlying event. As the remnants are color-charged as well, they undergo hadronization.

Furthermore, several collisions per bunch crossing occur as the LHC does not collide single protons but bunches of them. The additional collisions, called pileup, lead to additional signatures in the detector, which have to be taken into account as well. There are two types of pileup: in-time pileup arises from multiple collisions taking place in the same bunch crossing. Out-of-time pileup is caused by the finite response time of the detector being larger than the bunch crossing frequency of one crossing every 25 ns. Therefore, signals from the previous bunch crossing can still be seen in the detector.

4.1.4. Monte Carlo Event Generators

Dedicated software packages, called Monte Carlo event generators [107], exist to simulate some of the steps mentioned above. In general, there are different programs for matrix element generation at NLO and programs for the subsequent parton shower and hadronization. Some of them incorporate the whole simulation chain and some need to be combined with others in order to perform the whole workflow. In this thesis, simulated event samples generated with MadGraph5_aMC@NLO [108–110], POWHEG [111, 112] and PYTHIA [113, 114] are used. The Les Houches event (LHE) data format [115] is commonly used for interfacing the elements of the simulation chain.

MadGraph5_aMC@NLO is an NLO matrix element generator used within this thesis. It is an advancement of MadGraph [109], which only provides LO calculations, and MC@NLO [110]. Once the process is defined, events can be simulated at LO or NLO precision. For the NLO simulation, double counting of a LO diagram plus parton shower and an NLO diagram with an additional gluon can occur, i.e., adding a gluon via the NLO calculation or via the parton shower results in identical diagrams. Hence, the parton shower and the matrix element calculation have to be matched. The overlapping parton shower contributions are removed by the MC@NLO method. This method introduces negative event weights to deliver correct NLO predictions. The number of effective events is reduced by the negative weights and therefore more events need to be simulated for the same number of effective events. The negative weights can result in unphysical negative entries in distributions of observables. This can be avoided by generating more events, which then compensates for the negative weights.

The issue of negative event weights can be avoided by considering the hardest emission of the parton shower already in the NLO matrix element calculation. Thus, a parton shower generator with p_{T}-ordered partons is required, one of which is PYTHIA as described below. This approach is utilized by the positive weight hardest emission generator (POWHEG), which only yields positive event weights. Therefore, all simulated events can be used without a caveat. However, POWHEG is only able to simulate predefined processes. Defining new processes is rather complex.

The two aforementioned generators are only able to simulate the hard scattering process. The simulation of the parton shower has to be provided by another software package, one of which is PYTHIA. It is an event generator, which provides all links of the event simulation chain, starting from the matrix element and ending with the hadronization. The parton shower provided by PYTHIA is ordered in p_{T} of the partons and therefore is suited for the combination with POWHEG. The Lund String model is utilized for the hadronization. The phenomenological model uses several parameters to adjust the parton shower, underlying event and hadronization. A set of these parameters is called tune. The samples used in this thesis either employ CUETP8M1, CUETP8M2T4 [116, 117] or CP5 [118]. The first two are used in the analysis presented in Chapter 6, whereas the latter is used in Chapter 7. This is a result of

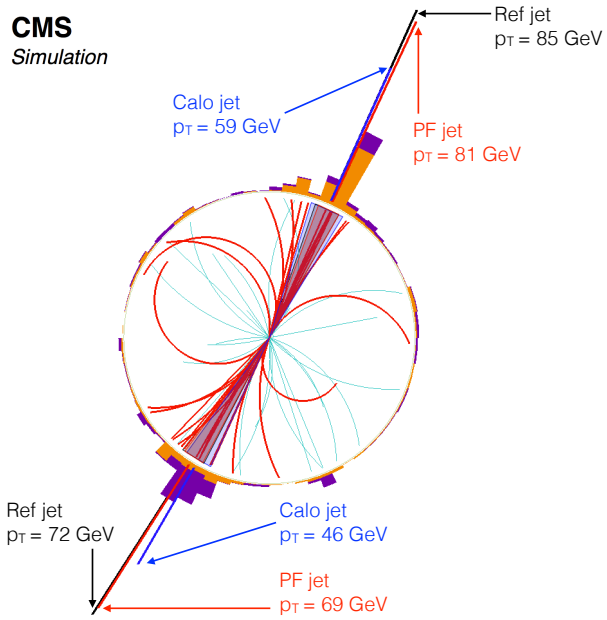

Figure 4.2.: Comparison of a PF jet and a calorimeter jet: the jet based on simulation truth information (Ref) is compared to the calorimeter jet (Calo) and the PF jet. The deviation w.r.t the Ref jet p_T is reduced significantly by the PF algorithm. Taken from Ref. [121].

samples using the latter tune not being available at the time when the analysis based on data taken in 2016 was carried out.

4.1.5. Detector Simulation

Finally, the interactions of the particles with the detector are simulated to obtain the signatures, which the events leave in the detector, allowing for a comparison with data. Therefore, the whole CMS detector is simulated including all its subdetectors, the magnet, the cabling, and the supporting structure. The detector simulation is performed using the GEANT4 package [119, 120].

GEANT4 simulates interactions via electromagnetic or hadronic processes in an energy range from 250 eV up to several TeV, covering the whole range of energies relevant at the LHC. The output of the detector simulation is similar to the real detector response. The simulation and data are processed further through the event reconstruction to obtain particle candidates.

4.2. Event Reconstruction

The simulated and recorded events are processed further as not the raw detector signals but the physical particles produced in the hard scattering are of interest. The inference of physical objects from the detector signals is called reconstruction. In order to provide the best possible identification and kinematic reconstruction of particles, all signals provided by the detector are taken into account. In contrast to, e.g., only using calorimeter information to reconstruct jets.

At the CMS experiment, the particle-flow (PF) algorithm [121–125] unites the subdetector information to obtain particle candidates. In the first step, energy clusters in the electromagnetic and hadron calorimeter, as well as trajectories in the tracking system, are reconstructed. These are called PF elements. Matching combinations of tracks and calorimeter entries are combined into PF blocks, e.g., matching a track and an ECAL cluster to obtain an electron signature.

Based on the PF blocks, the reconstruction and identification is performed. As the identification of muons is straightforward, muons are reconstructed first. Entries in the muon chambers are matched to signatures of the inner detector parts. The PF blocks used for the muon reconstruction are then removed. The remaining blocks are used for the electron and photon reconstruction. Again, the matched PF blocks are removed and the remaining blocks are used to reconstruct hadrons, charged, as well as neutral. Finally, jets are built from the hadrons. The improvement by using the PF jets instead of solely calorimeter-based jets is shown in Fig. 4.2.

4.2.1. Particle Tracks and Primary Vertex Reconstruction

The hits in the tracking subdetectors are combined into particle trajectories using the Combinatorial Track Finder (CTF) [126]. The CTF algorithm utilizes Kalman filters [65, 127–130] in an iterative approach. In each iteration, a Kalman filter is applied to add matching hits to track candidates starting from a seed. Additionally, the tracks are required to fulfill quality criteria. At first, tracks close to the interaction point that have transverse momenta of $p_T > 0.8\,\text{GeV}$ are reconstructed and the corresponding hits are removed from the set.

The track information is used to reconstruct their origins, called vertices. The vertex reconstruction is performed using high quality tracks only [131]. There are several vertices in each event, one of which is the primary vertex. It is characterized by the highest sum of squared transverse momenta of the assigned tracks that are associated to physical objects. Other proton-proton scatterings in the same bunch crossing are called in-time pileup, causing further vertices. Particles produced in the primary vertex and decaying after traveling a short distance produce secondary vertices. Furthermore, the underlying event can yield additional vertices. The selected tracks are grouped with respect to their distance to the beam spot in z direction by the deterministic annealing algorithm [132]. The position of the vertex is determined with the adaptive vertex fitting algorithm [133] for vertices with at least two corresponding tracks. Each track receives a weight, which states the probability of the track belonging to the vertex. In this thesis, the primary vertex is required to lie within a cylinder of radius 2 cm and length 24 cm around the center of the detector. Furthermore, it is required to have at least four degrees of freedom. The charge hadron subtraction method [134] is utilized to remove PF objects originating from in-time pileup effects.

4.2.2. Muon Reconstruction

Muons leave a unique signature in the detector as they pass the whole detector without being absorbed. In contrast to all other particles, they produce signals in the muon system. Therefore, the reconstruction of muons is easier than for the rest of the particles. Furthermore, hits in the tracking system can be used to reconstruct muons. Hence, the PF algorithm reconstructs muons first and then subtracts the corresponding hits from the event.

Firstly, the tracks in the muon system and in the tracker are determined independently. The corresponding particle candidates are called standalone and tracker muons, respectively. There are two ways of combining the two: inside-out (tracker muon) or outside-in (global muon). Tracker muons are reconstructed starting from tracks with $p_T > 0.5\,\text{GeV}$, which are matched to muon segments built from hits in the muon system. This proves especially efficient for muons with $p_T < 5\,\text{GeV}$. The global muon reconstruction starts with a standalone muon, which is matched to a track. Both together are fitted using a Kalman filter. This increases the resolution for muons with $p_T > 200\,\text{GeV}$ significantly. In principle, standalone muons can be used without the matching. Nevertheless, the resolution is insufficient and therefore they are not used in this thesis.

In addition to the PF requirements, further criteria as recommended by the CMS Muon Physics Objects Group have to be fulfilled by muon candidates. These criteria are combined into identification (ID) flags, which are binary variables that indicate whether the criteria are fulfilled. For muons, there is a loose and a tight ID. The loose ID accepts global and tracker muons, whereas the tight ID only accepts

global muons and furthermore requires additional criteria to be fulfilled to enhance the reconstruction quality. Nonprompt muons, e.g., muons produced in a jet, can be suppressed by introducing the muon isolation

$$I_\mu = \frac{1}{p_{T,\mu}} \Big(p_{T,CH} + \max\left(p_{T,NH} + p_{T\gamma} - \Delta\beta \cdot p_{T,CH(PU)}, 0 \right) \Big), \tag{4.2}$$

where $p_{T,\mu}$ is the transverse momentum of the muon, $p_{T,CH}$, $p_{T,NH}$, $p_{T\gamma}$ is the transverse momentum of the charged and neutral hadrons, and the photons, respectively. The transverse momentum of charged hadrons within a cone of radius 0.4 around the muon stemming from pileup is given by $p_{T,CH(PU)}$. The pileup contribution to the charged particles can be estimated. This leads to a correction factor $\Delta\beta = 0.5$, representing the fraction of neutral to charged hadrons originating from pileup [135].

4.2.3. Electron Reconstruction

After the reconstruction of the muons and the subtraction of the corresponding PF elements from the event, electrons are reconstructed. Since their mass is comparatively low, they deposit a significant share of their energy in the tracker and the rest of it in the ECAL. Consequently, electrons typically produce broad showers in the ECAL, which distinguishes them from the remaining particles. This also raises difficulties in the reconstruction as narrow showers are easier to reconstruct.

The reconstruction of electrons is especially challenging as they loose energy in the tracker and therefore produce tracks with non-constant curvature. These tracks can be classified as bad tracks by the CTF algorithm and are difficult to be matched to calorimeter entries. Consequently, Gaussian sum filters (GSF) are utilized for bad quality tracks. Here, the energy loss of an electron in each layer is assumed to follow a Gaussian distribution resulting in a better description of the electron trajectory.

The reconstruction of the energy clusters in the ECAL starts from the crystal with the highest energy deposit as the seed. The seed is clustered with neighboring energy deposits to form a supercluster (SC). There are two possibilities of matching the energy deposits in the ECAL to the tracks: for electrons with low transverse momenta, CTF and GSF tracks are matched to the ECAL clusters utilizing multivariate methods. For high-p_T electrons, SC seeds are matched to CTF or GSF tracks. Again, the matched PF elements are removed from the event.

As for muons, there are additional criteria defined for electrons to be used in an analysis. The CMS E/Gamma Physics Objects Group defines three ID flags used in this thesis with electron tagging efficiencies of 75%, 90% and 95% for tight, loose and veto ID, respectively. The criteria for the ID flags vary with LHC beam conditions in order to keep the efficiencies constant over the data taking period. The ID is based on, e.g., the supercluster shower shape, the $\Delta\eta$ and $\Delta\phi$ between the track and the supercluster and the electron isolation defined as

$$I_e = \frac{1}{p_{T,e}} \Big(p_{T,CH} + \max(p_{T,NH} + p_{T,\gamma} - \rho \cdot A_{eff}, 0) \Big), \tag{4.3}$$

where $p_{T,e}$ is the transverse momentum of the electron. The average transverse momentum density is ρ and A_{eff} is the effective area for the estimation of the pileup contribution. The electron isolation cone size is set to a radius of 0.3.

4.2.4. Photon and Hadron Reconstruction

The remaining detector entries are assumed to stem from hadrons or photons in the event. For these particles, there are two categories: charged particles leaving traces in the tracker and producing calorimeter entries and neutral particles causing calorimeter entries only. The charged particles are mostly charged pions, kaons, and protons while the neutral particles can be prompt photons and neutral hadrons such as neutral kaons, neutrons and photons from decays of neutral pions. The reconstruction of isolated prompt photons is similar to the electron reconstruction, despite the matching to the tracker entries [136]. After removing the corresponding hits, the ECAL and HCAL clusters are associated to photons and

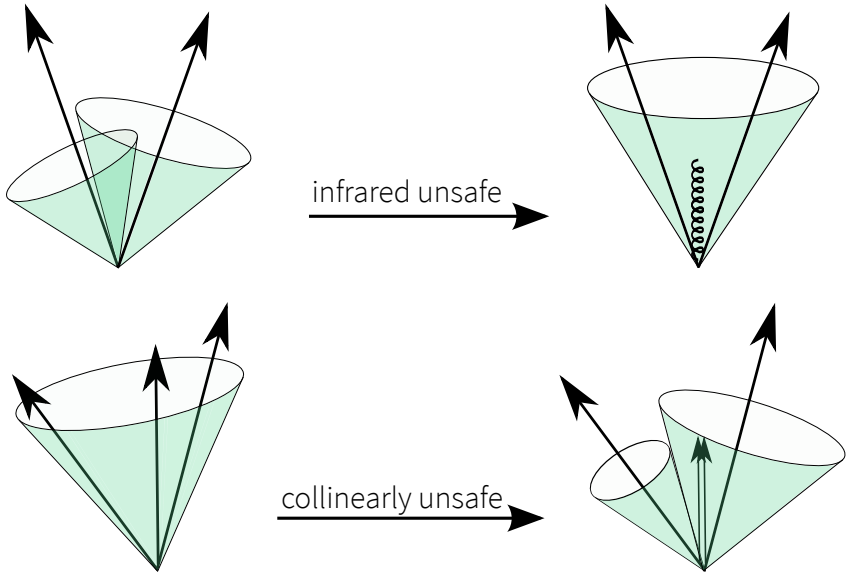

Figure 4.3.: Infrared and collinear safe algorithms: The jet clustering algorithm is required to be insensitive to gluon radiation (infrared safe, top) and insensitive to collinear parton splitting (collinear safe, bottom). Taken from Ref. [98]

neutral hadrons for $|\eta| < 2.5$, the coverage of the tracker. In the forward region, no tracker information is available and therefore the distinction between charged and neutral hadrons is more difficult. If HCAL clusters can be linked to charged hadrons, they are reconstructed as charged hadrons as well.

The remaining ECAL and HCAL clusters, as well as tracks, are then linked to charged hadrons and muons that fulfill not all quality criteria or, in case no tracks can be matched, to photons and neutral hadrons. The criteria used for the matching can be found in Ref. [136].

4.2.5. Jet Reconstruction

The single reconstructed hadrons are not of interest for most analyses but rather groups of hadrons originating from the same final-state particle. These groups of hadrons are called jets. Although the jets are not physical objects themselves, the sum of their components gives access to the features of the final-state particles from which they originate. Furthermore, jets provide a basis that enables easier comparisons between data and simulation.

The reconstruction of jets is based on well defined algorithms [137]. In general, there are two types of jet clustering algorithms: sequential and cone-based algorithms. Both need to fulfill two criteria, i.e., collinear safety and infrared safety as shown in Fig. 4.3. An algorithm is collinear safe if its result is independent of collinear parton splitting. Insensitivity against soft emissions of gluons is called infrared safety. If both criteria are fulfilled, the algorithm provides robust results, which enables a comparison with theory predictions. The objects are clustered using a distance metric

$$d_{ij} = \min(p_{\mathrm{T},i}^{2n}, p_{\mathrm{T},j}^{2n}) \frac{\Delta_{ij}^2}{D^2}, \tag{4.4}$$

with the transverse momentum p_{T} of the objects, the algorithm specific parameter n, the difference between the two objects in the y-ϕ plane $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and a fixed parameter D defining the size of the clustered objects. In this thesis, $D = 0.4$ is chosen. The value of n is chosen differently in certain algorithms: the anti-k_{T} algorithm [138] uses $n = -1$, the Cambridge-Aachen [139] algorithm sets n to zero and the k_{T} algorithm [140] relies on $n = 1$. An example of the jet clustering for all

three algorithms can be found in Fig. 4.4. The Cambridge-Aachen algorithm is mostly used for boosted topologies. Setting $n = -1$ as in the anti-k_T algorithm favors the clustering of high-p_T jets by integrating nearby objects into the jet instead of forming a new jet from them. In this thesis, the anti-k_T jet clustering algorithm is utilized.

It joins objects with the smallest distance d_{ij} until a certain distance to the beam is reached. The distance between an object i and the beam is defined as:

$$d_{i,\text{beam}} = p_{T,i}^{-2}. \tag{4.5}$$

Furthermore, the PF Jet ID [141], which is a quality flag similar to those for electrons and muons, is used. It is implemented to distinguish between jets clustered from detector noise and physical particles. The PF Jet ID combines different criteria, e.g., the energy fraction of neutral particles in ECAL and HCAL or the number of constituents of the jet. Similar to muons, there are two IDs, loose and tight.

Jet Energy Corrections

The jet energy is corrected for several effects. In general, the calorimeter response is calibrated using particles with precisely known energies. Furthermore, the energy of the jets is biased. On the one hand, the CMS detector is not fully homogeneous and hence the reconstructed energy depends on where in the detector the jet is located. On the other hand, particles originating from pileup events influence the jet energy. Therefore, certain corrections are applied [142–145]. Here, a sequential multilayer approach is applied, where each layer targets the mitigation of a dedicated effect:

- **L1 pileup:** At first, the jet energies are corrected for pileup contributions. For this, QCD simulations are compared to data events with high and low pileup. Furthermore, residual differences between data and simulation are obtained with the random cone (RC) method [142]. This leads to a parametrization of the energy density, which is used to obtain η-dependent residual corrections [146] for data and simulated samples.

- **L2L3 MC truth:** The reconstructed p_T is compared to the simulated p_T based on a QCD dijet sample to obtain jet response corrections as a function of jet p_T and η.

- **L2L3 residual corrections:** Remaining small differences at the order of a few percent are corrected by the L2L3 residual corrections. These are obtained from dijet events to correct the jet energy scale (JES) in dependence of p_T and η. Jets in the central region ($|\eta| < 2.4$) are corrected based on Z+jet, photon+jet and multijet background (MJB) events.

The whole chain of the jet energy corrections is shown in Fig. 4.5. The flavor MC correction is not applied in this thesis and therefore not described further.

b Tagging

All final-state quarks, as well as gluons, can only be measured as jets. In most cases it is very hard to derive the flavor of the final-state quark. Nonetheless, jets originating from bottom quarks have distinguishing properties, which can be utilized to tag those jets. The bottom quarks form B hadrons with a finite lifetime. The lifetime is larger than that of the other hadrons as the off-diagonal CKM matrix elements are very small and the mass of the decay products is close to the mass of the B hadrons leading to a small phase space. Therefore, the B hadrons travel a certain distance in the beam pipe before decaying. In the event, this is visible as a displaced secondary vertex as shown in Fig. 4.6. Dedicated b tagging algorithms are implemented to obtain the best discrimination between b jets and jets stemming from gluons or other quarks. These algorithms utilize different machine learning techniques in order to combine information on secondary vertices, tracks, and jet properties.

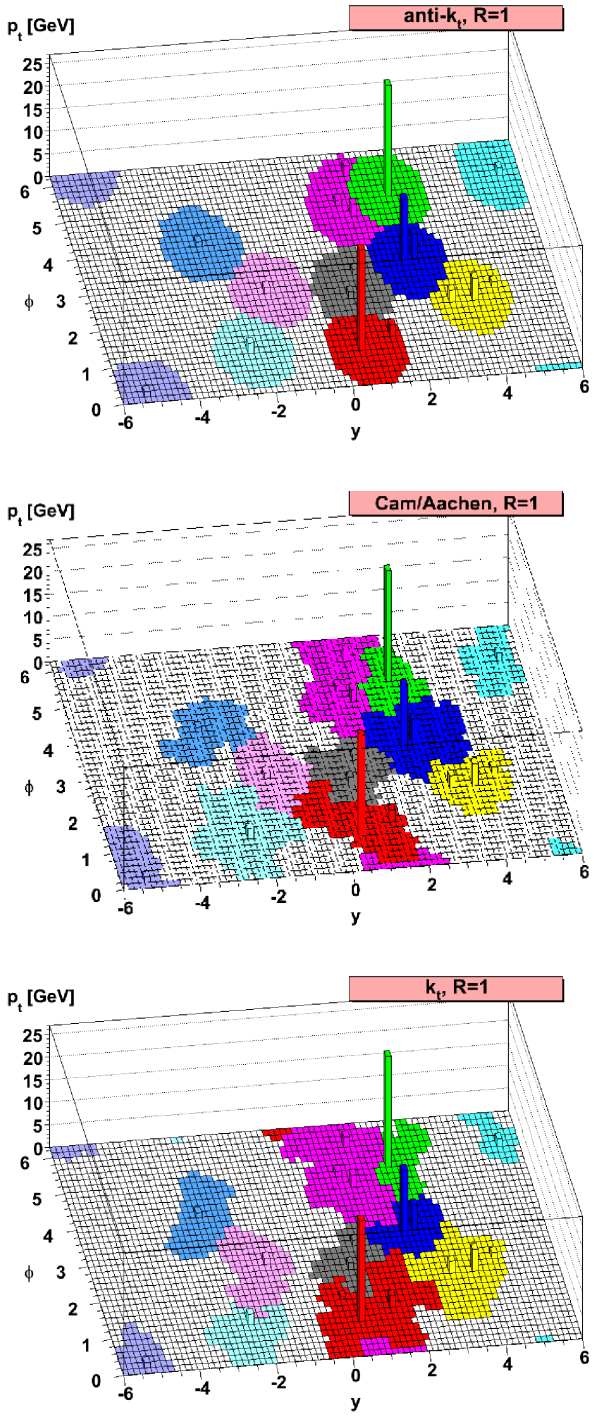

Figure 4.4.: Clustering of jets with different algorithms: the clustering with the anti-k_{T} algorithm is shown at the top, the Cambridge-Aachen algorithm in the middle, and the k_{T} algorithm on the bottom. Each color corresponds to a different jet reconstructed by the respective algorithm. The ϕ-y grid corresponds to the calorimeter cells. Taken from Ref. [140].

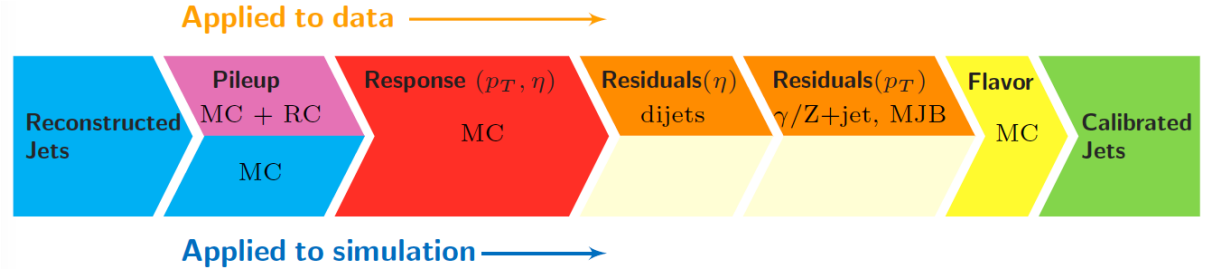

Figure 4.5.: Workflow for the jet energy corrections: starting with the reconstructed jets, corrections for pileup and MC simulation information are applied. Furthermore, residual corrections are implemented. All corrections labeled with MC are derived from simulation. The flavor corrections are not used within this thesis. Taken from Ref. [142]

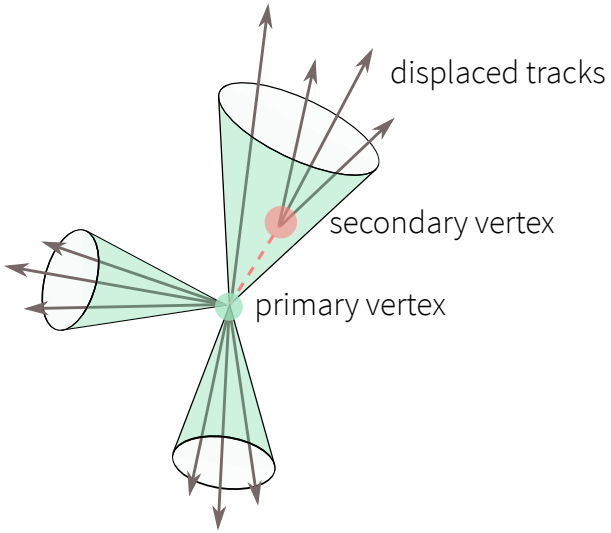

Figure 4.6.: b jet properties: as the B hadrons posses a larger lifetime than other hadrons, they travel a certain distance (red dashed line) before decaying. This results in displaced tracks, from which a secondary vertex is reconstructed. Taken from Ref. [98]

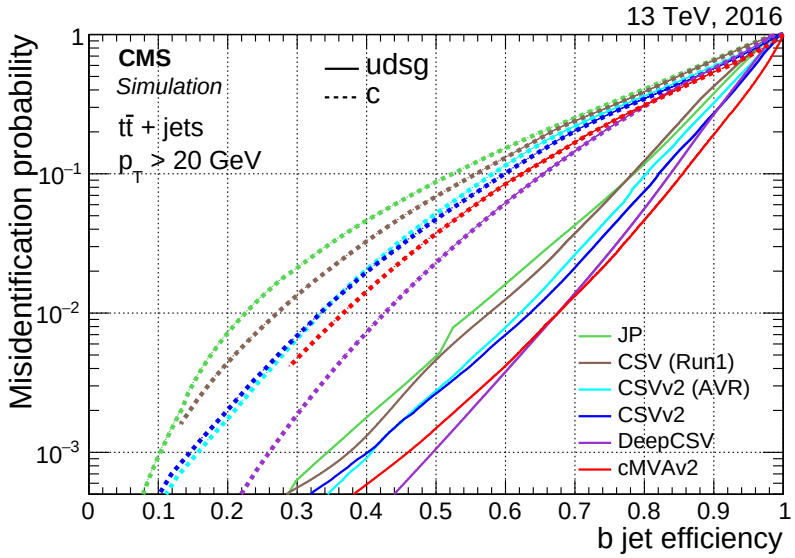

Figure 4.7.: c and light-flavor jet misidentification probability over b jet identification efficiency for different b tagging algorithms obtained from simulated $t\bar{t}$ events: The best discrimination is given by the DeepCSV algorithm. Taken from Ref. [149]

In this thesis, two different algorithms are used: the Combined Secondary Vertex Tagger v2 (CSV) [147, 148] and the DeepCSV [149] tagger. The CSV tagger is based on three neural networks, which are combined using a likelihood ratio. Each of the neural networks targets a different kind of events. One is used for events with a reconstructed secondary vertex, the second targets events without a secondary vertex but at least two tracks with a combined mass above a certain threshold and the third targets events that do not fit the first two categories. The resulting discriminator yields outputs between 0 and 1, where higher values indicate a higher probability for the jet to originate from a bottom quark. The training is performed using two independent sets, one with charm quarks as background and the other one with up, down, and strange quarks as background. The final discriminator is the weighted sum of the two trainings, where the latter is weighted with a factor of three.

The DeepCSV algorithm is an improvement of the CSV algorithm using the newest machine learning techniques. Here, a deep neural network instead of a shallow network is used and a larger number of input variables is utilized. Furthermore, the training is performed in all vertex categories at once. Additionally, the output is not binary but multidimensional. In total, there are five output nodes: P(b) targeting jets with exactly one B hadron in the jet, P(bb) targeting jets with at least two B hadrons in the jet, P(c) targeting jets without a B hadron and one C hadron, P(cc) targeting jets without a B hadron and at least two C hadrons, and P(udsg) targeting all other jets. This allows for a more detailed distinction of the jets. A comparable measure to the CSV tagger can be obtained by adding P(b) and P(bb).

For each of the taggers, three working points are defined: loose, medium and tight. These working points correspond to mistagging rates for light-flavor or gluon jets of 10%, 1% and 0.1%, respectively for jets with a b-tagger output greater or equal to the respective threshold. Studies on simulated top quark pair events yield a b tagging efficiency of 81% (loose), 63% (medium) and 41% (tight), for the CSVv2 algorithm and 84% (loose), 68% (medium) and 50% (tight), for the DeepCSV algorithm [149]. As can be seen, the efficiencies are higher for the DeepCSV algorithm for all working points compared to the CSVv2 algorithm. A comparison of the misidentification probabilities as a function of the b jet efficiency is shown in Fig. 4.7.

The CSVv2 tagger is used in the analysis with data recorded in 2016 by the CMS experiment presented in Chapter 6 as this was the default tagging algorithm at the time when the analysis was carried out.

For the analysis of the 2016–2018 data set presented in Chapter 7, the DeepCSV tagger is used since it yields higher tagging efficiencies.

4.2.6. Missing Transverse Momentum

All the methods mentioned above focus on particles that are either directly or through their decay products observable by the detector. Nevertheless, there are particles that are not observable by the detector. In the SM, the only particles that do not leave a trace in the detector are the neutrinos as they are only interacting weakly. Additionally, in BSM models there are further particles, which are not directly detectable. There is a possibility to obtain the missing transverse momentum associated with undetectable particles [150, 151]. In the initial state of the hard scattering process, the total transverse momentum is approximately zero. As a result of momentum conservation, the total transverse momentum of all particles produced in a collision has to be zero as well. Therefore, a deviation in the total transverse momentum from zero can be associated to these particles. Further, smaller contributions to $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ come from areas that are not covered by the detector and the uncertainties on the transverse momentum introduced by the jet reconstruction. Summing over all visible transverse momenta leads to a raw momentum vector pointing in the opposite direction of the missing transverse momentum:

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss,raw}} = -\sum_{i}^{N} p_{\mathrm{T,i}}. \tag{4.6}$$

Taking the pileup correction $\vec{C}_{\mathrm{T}}^{0}$ and the propagated jet energy correction $\vec{C}_{\mathrm{T}}^{1}$ [152] into account, the missing transverse momentum is defined as:

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} = \vec{p}_{\mathrm{T}}^{\,\mathrm{miss,raw}} + \vec{C}_{\mathrm{T}}^{0} + \vec{C}_{\mathrm{T}}^{1}. \tag{4.7}$$

4.2.7. W Boson Reconstruction

This thesis studies processes involving top quarks, which dominantly decay into a W boson and a bottom quark. As the W boson is not stable and therefore decays further, it cannot be observed directly in the detector. In general, there are two classes of W decays: leptonic and hadronic. Here, the focus lies on leptonic decays as this allows for better background suppression. This results in two decay products, the lepton and the corresponding neutrino, where leptons are electrons or muons within this thesis. Tauons are only considered indirectly if they decay into electrons or muons.

In order to reconstruct the four-vector of the W boson, the lepton and neutrino have to be added. Unfortunately, the neutrino is only indirectly detectable through the missing transverse momentum. Hence, there is no information about the longitudinal momentum of the neutrino $p_{z,\nu}$. This can be solved by constraining the W boson mass to the literature value of $m_{\mathrm{W}} = 80.395\,\mathrm{GeV}$ [18]. Here, the method as described in Ref. [153] is utilized. The W mass is connected to the decay products via

$$m_{\mathrm{W}}^2 = \left(E_\ell + \sqrt{(p_{\mathrm{T}}^{\mathrm{miss}})^2 + p_{z,\nu}^2}\right)^2 - (\vec{p}_{\mathrm{T},\ell} + \vec{p}_{T}^{\,\mathrm{miss}})^2 - (p_{z,\ell} + p_{z,\nu})^2, \tag{4.8}$$

with the energy of the lepton

$$E_\ell = \sqrt{p_{\mathrm{T},\ell}^2 + p_{z,\ell}^2}, \tag{4.9}$$

the transverse momenta of the electron $p_{\mathrm{T},\ell}$ and the neutrino $p_{\mathrm{T},\nu}$, and the longitudinal momenta $p_{z,\ell}$ and $p_{z,\nu}$, respectively. Solving for the longitudinal neutrino momentum leads to

$$p_{z,\nu}^{\pm} = \frac{\Delta \cdot p_{z,\ell}}{p_{\mathrm{T},\ell}^2} \pm \sqrt{\frac{\Delta^2 \cdot p_{z,\ell}^2}{p_{\mathrm{T},\ell}^4} - \frac{E_\ell^2 \cdot (p_{\mathrm{T}}^{\mathrm{miss}})^2 - \Delta^2}{p_{\mathrm{T},\ell}^2}}, \tag{4.10}$$

with

$$\Delta = \frac{m_{\mathrm{W}}^2}{2} + \vec{p}_{\mathrm{T},\ell} \cdot \vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} \tag{4.11}$$

$$= \frac{m_{\mathrm{W}}^2}{2} + p_{\mathrm{T},\ell} \cdot p_{\mathrm{T}}^{\mathrm{miss}} \cdot \cos(\phi), \tag{4.12}$$

with the angle ϕ between lepton and $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. There are two cases: for a positive discriminant, two solutions occur, the one with the smaller absolute value is chosen in this case as the longitudinal momentum of the neutrino. In case of a negative discriminant, two complex solutions are obtained, which can be caused by systematic uncertainties in the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ or energy measurement. In this case, the momenta of the neutrino in x and y direction are varied to obtain one real solution. The variation is determined by the minimization procedure as described in Ref. [153].

In events with high jet multiplicities, hadronically decaying W bosons are more complex to reconstruct, as there are typically several jets, which are candidates for the reconstruction. Therefore, sophisticated machine learning methods can be implemented to find the jets in the event, which belong to the W boson. In the associated tHW production, e.g., a leptonically and a hadronically decaying W boson have to be reconstructed. The jet assignment under dedicated hypotheses is described in the next chapter.

5. Event Topology and Jet Assignment

This chapter introduces the processes relevant for the analyses described in the following two chapters. For both analyses, Higgs boson production in association with single top quarks or in association with top quark pairs are considered as signal. The main background process in the analyses is the top quark pair production. Additionally, the single top quark production and further minor backgrounds are discussed.

As described in the previous chapter, the events are reconstructed, which results in objects such as jets, photons, and leptons. For the latter, the signatures in the detector differ between the flavors. Hence, the lepton flavor can be determined. However, it is not trivial to find out which jet originates from which final-state particle. Therefore, a BDT-based jet assignment method is implemented. This method is tailored to the signal and background processes relevant for the analyses presented in the following chapters.

5.1. Signal Processes

As described in Chapter 1, there are two processes that are directly sensitive to the top quark Yukawa coupling: single top quark production in association with a Higgs boson (tH) and top quark pair production in association with a Higgs boson ($t\bar{t}H$); both are considered as signal processes within this thesis. For the tH production, two signal processes are considered: single top t-channel production in association with a Higgs boson (tHq) and the associated production of a single top quark, a Higgs boson, and a W boson (tHW). Modifications in the Yukawa coupling directly affect the cross section of these processes, see Appendix A.1, and their kinematic properties, see Appendix A.2.

Here, the Higgs boson decay into a pair of bottom quarks is studied. Furthermore, a leptonically decaying W boson is required, leading to exactly one lepton in each event. The following descriptions are made under these restrictions.

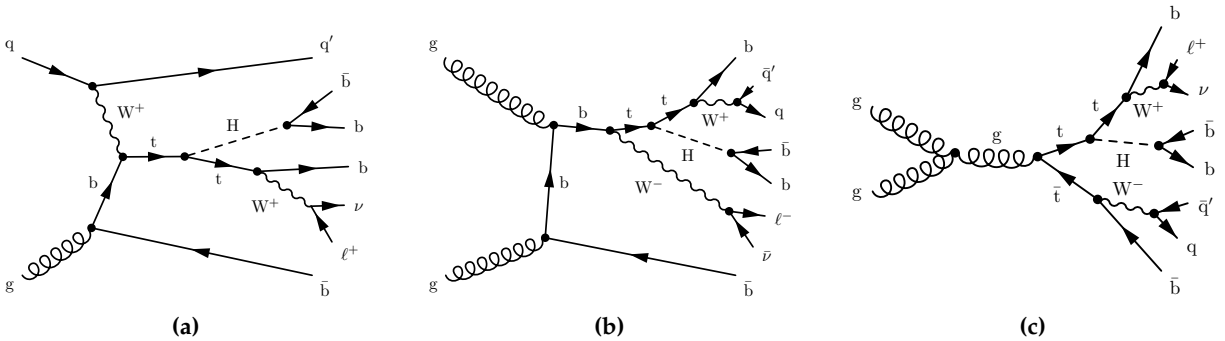

Figure 5.1.: Feynman diagrams including final-state particle decays for the signal processes: the tHq process (a), the tHW process (b), and the $t\bar{t}H$ process (c) are shown. Only semileptonic decays are shown here since this thesis focuses on single lepton events.

5.1.1. t-channel Single Top Quark Production in Association with a Higgs Boson

In the tHq process, a single top quark, a Higgs boson, and a light quark are produced. For the simulated samples, the 4FS is utilized. Hence, an additional bottom quark is produced via gluon splitting. The Higgs boson either couples to the top quark or to the W boson, leading to an interference, which allows the measurement of the top quark Yukawa coupling, the coupling of the Higgs boson to the W boson, and the relative sign between the two couplings.

The top quark, as well as the Higgs boson, decay further. The top quark decays into a bottom quark and a W boson, which further decays into a lepton and the corresponding neutrino as only leptonically decaying W bosons are considered. The Higgs boson is required to decay into a pair of bottom quarks. In total, the final state of the tHq process consists of four bottom quarks, a light quark, a lepton, and the corresponding neutrino. The Feynman diagram of the tHq process including the decay of the final-state particles is shown in Fig. 5.1(a).

The light quark produced in this process has a characteristic feature: it is mostly produced in forward direction, as shown in Fig. 5.2 for the inverted top quark coupling (ITC) case with $-\kappa_t = \kappa_V = 1$. This feature is rather unique and can be used to distinguish the light quark from the remaining quarks, as well as to discriminate the tHq process from background processes. Depending on the Higgs boson coupling, the kinematic properties of the final-state particles change as shown in Appendix A.2.

5.1.2. tW Single Top Quark Production in Association with a Higgs Boson

The tHW process produces a single top quark, a Higgs boson, a W boson, and an additional bottom quark from gluon splitting. In order to avoid interference effects with the $t\bar{t}H$ process, the tHW process is simulated in the 5FS. This setup is chosen as there is currently no simulated sample available that contains tHW and $t\bar{t}H$ events with interference effects included. Similar to the tHq process, the Higgs boson either couples to the top quark or to the W boson. Therefore, this process is sensitive to the top quark Yukawa coupling, the Higgs boson to W boson coupling and the relative sign between the two.

The Higgs boson is required to decay into a pair of bottom quarks and the top quark to decay into a W boson and a bottom quark. In general, the W boson from the top quark decay and the additional W boson radiated off by the bottom quark can decay leptonically or hadronically. In this thesis, only cases where one of them decays hadronically and the other one decays leptonically are considered, but there is no constraint on which W boson decays how. Altogether, the final state of the tHW process results in four bottom quarks, two quarks from the hadronic W boson decay, a lepton, and the corresponding neutrino.

The Feynman diagram including the decay of the final-state particles for the tHW process is shown in Fig. 5.1(b). The simulated distributions for the pseudorapidity and the transverse momentum of the final-state particles are shown in Fig. 5.2 for the ITC scenario. Depending on the Higgs boson coupling, the kinematic properties of the final-state particles change as shown in Appendix A.2.

5.1.3. Top Quark Pair Production in Association with a Higgs Boson

In the $t\bar{t}H$ process, the Higgs boson couples directly to the top quark. Consequently, this process is an excellent candidate to study the top quark Yukawa coupling. Since the cross section is proportional to κ_t^2, only constraints on the amplitude but not on the sign of the top quark Yukawa coupling can be set. Nonetheless, this process enhances the sensitivity of the analyses presented in the following chapters significantly. Since there is no interference with other relevant processes, the $t\bar{t}H$ process is simulated in the 4FS.

In the $t\bar{t}H$ process, two top quarks and a Higgs boson are produced, where each of the top quarks decays into a bottom quark and a W boson and the Higgs boson is required to decay into a pair of bottom quarks. As this thesis studies single lepton events, one of the top quarks has to decay hadronically and the other one leptonically. This results in four bottom quarks, two from the top quark decays and two

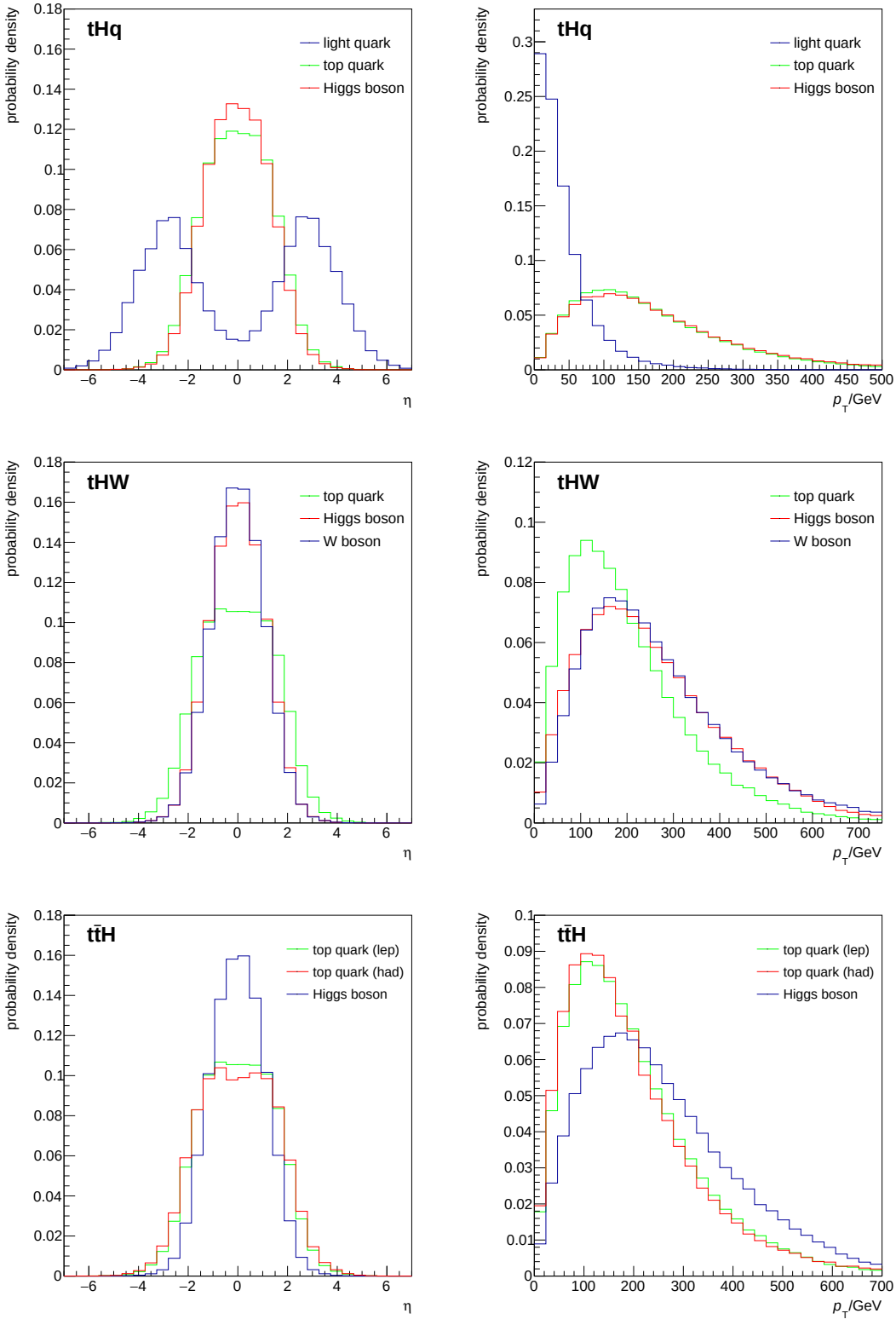

Figure 5.2.: Parton level distributions for the final-state particles: the η distribution (left) and p_T distribution (right) are shown for the ITC scenario. For the tHq process (top), the light quark is preferably produced in forward direction and has lower transverse momentum. All other particles are favorably produced in the central region. For the tHW process (middle), the transverse momentum of the top quark is smaller than that of the W boson and that of the Higgs boson. For the $t\bar{t}H$ process (bottom), the transverse momentum of the Higgs boson is typically lower than that of the leptonically (lep) and hadronically (had) decaying top quarks.

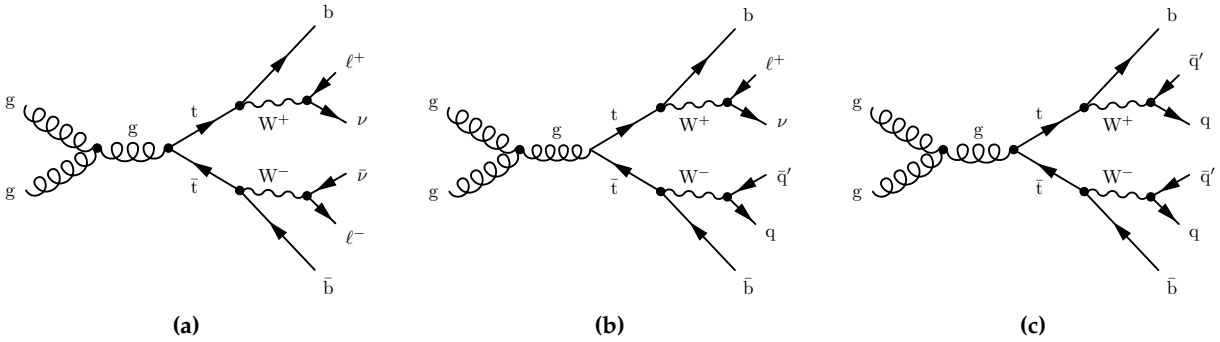

Figure 5.3.: Feynman diagrams including the decay of the final-state particles for the $t\bar{t}$ background process: the fully-leptonic decay mode is shown on the left, the semileptonic mode is shown in the middle, and the fully-hadronic mode can be found on the right.

from the Higgs boson decay, two light quarks from the hadronic W boson decay, and a lepton and the corresponding neutrino from the leptonically decaying W boson.

The Feynman diagram including the decay of the final-state particles for the $t\bar{t}H$ process is shown in Fig. 5.1(c). The η and p_T distributions on parton level for the decay particles are shown in Fig. 5.2. All final-state particles are produced in the central region; the Higgs boson typically has a lower transverse momentum than the top quarks. Different κ_t scenarios only modify the cross section of the $t\bar{t}H$ process leaving the kinematic properties unchanged.

5.2. Background Processes

There are several SM processes that produce very similar signatures to that of the signal processes. All these processes have to be considered in the analysis. The most important background process is the top quark pair production ($t\bar{t}$). Furthermore, the single top quark production can produce signal-like event topologies as well. For both processes, there is only the Higgs boson missing in comparison to the signal events. In case of initial-state (ISR) or final-state radiation (FSR), where a gluon is radiated off by the respective particles, the number of final-state particles for the signal and background processes is identical. If the gluon splits into two bottom quarks, all of the final-state particles are exactly the same. This circumstance makes the distinction between signal and background processes very challenging.

5.2.1. Top Quark Pair Production

The $t\bar{t}$ production does not only have a similar signature as the signal, but does also happen far more often since the cross section is three orders of magnitude larger than the $t\bar{t}H$ cross section and even four orders larger than the SM tH cross section. For this thesis, the most important background process is the semileptonic $t\bar{t}$ production, where one of the W bosons decays leptonically and the other one decays hadronically. This results in two bottom quarks, two light quarks, one lepton, and the corresponding neutrino produced in the semileptonic $t\bar{t}$ process. The fully-hadronic decay mode is negligible, when requiring a charged lepton as done in this thesis. Further background contributions stem from the fully-leptonic decay since some of these events yield only one reconstructed lepton. The Feynman diagrams including the decay of the final-state particles for the three different $t\bar{t}$ modes are shown in Fig. 5.3.

The discrimination between the signal processes and the $t\bar{t}$ background process becomes especially challenging in case of ISR or FSR, where the radiated gluons split into pairs of quarks. The modeling of the flavor composition for the additional quarks is important to describe this background correctly. Unfortunately, it is hard to model by theory, which results in large systematic uncertainties on the

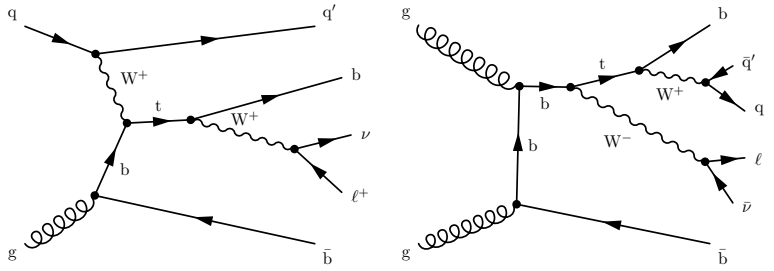

Figure 5.4.: Feynman diagrams including the decay of the final-state particles for the single top quark background process: the *t*-channel production is shown on the left and the tW associated production is shown on the right.

rate of contributions from different flavors. Therefore, the simulated $t\bar{t}$ samples are split into different categories based on a jet-quark matching:

- **$t\bar{t}$+lf**: events with additional jets stemming from up, down, or strange quarks, and events without any additional jets
- **$t\bar{t}$+$c\bar{c}$**: events with two additional jets stemming from charm quarks, events with one additional jet originating from two charm quarks, and events with one additional jet matched to a charm quark and the other jet produced by the second charm quark not matched
- **$t\bar{t}$+$b\bar{b}$**: events with two additional jets stemming from bottom quarks
- **$t\bar{t}$+2b**: events with one additional jet originating from two bottom quarks
- **$t\bar{t}$+b**: events with one additional jet matched to a bottom quark and the other jet produced by the second bottom quark not matched

The splitting is used to provide different background shapes, which are used to constrain the systematic uncertainties on the $t\bar{t}$+jets production in the final fitting procedure.

5.2.2. Single Top Quark Production

The single top quark production is another important background because it can produce the same signature as the signal, when radiating off an additional gluon. In this thesis, only the two single top quark production processes with the highest cross section are considered, i.e., the *t*-channel and the associated tW production. The single top quark production in the *s*-channel is negligible since its cross section is a factor of 20 smaller than that of the *t*-channel production. Similar to the signal processes, the *t*-channel process is modeled in the 4FS, whereas the tW associated production is modeled in the 5FS.

In the *t* channel, the process yields two bottom quarks, a light quark, a lepton, and the corresponding neutrino. The tW associated production features two W bosons, which leads to three possible decay channels: fully-leptonic, semileptonic and fully-hadronic. In the analyses presented in this thesis, only the semileptonic channel is relevant. The decay products are two bottom quarks, two light quarks, a lepton, and the corresponding neutrino. The Feynman diagrams including the decay of the final-state particles for both processes are shown in Fig. 5.4.

5.2.3. Minor Background Processes

In addition to the single top quark and top quark pair production, other processes of minor impact on the analysis are considered. These minor background processes are summarized as miscellaneous backgrounds (Misc).

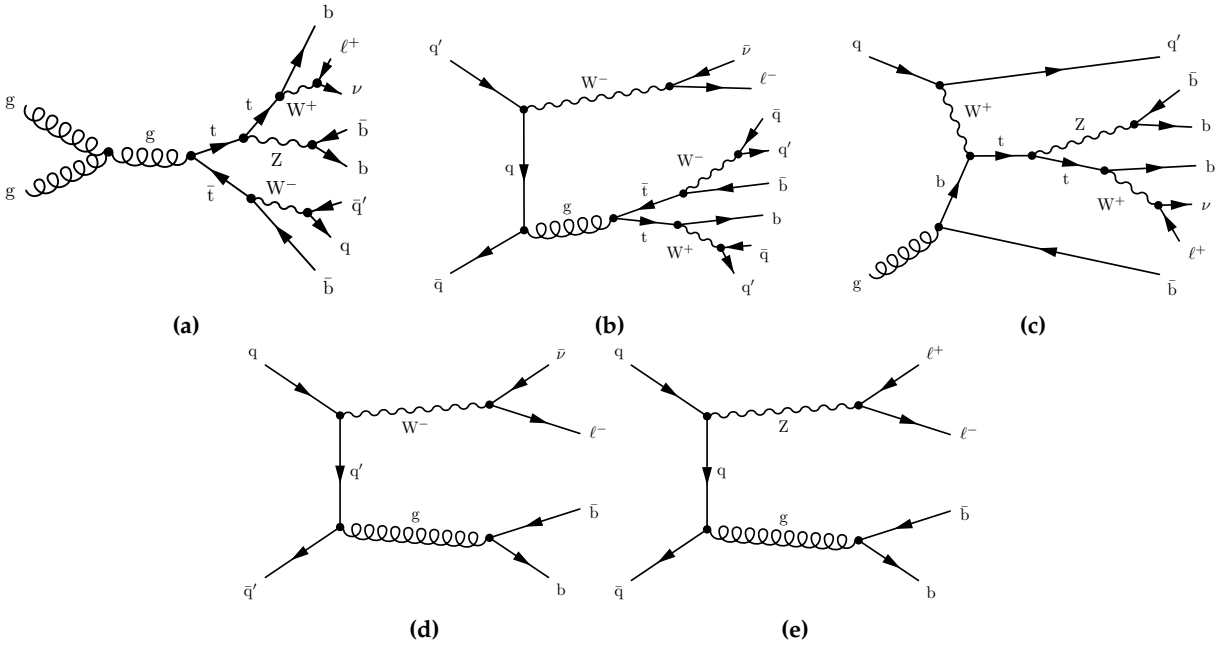

Figure 5.5.: Feynman diagrams including the decay of the final-state particles for minor background processes: the top quark pair production in association with a Z boson (a) or a W boson (b), the single top quark production in association with a Z boson (c), the W boson production with jets (d) and the Z boson production with jets (e) are shown.

The top quark pair production in association with a Z boson ($t\bar{t}Z$) or a W boson ($t\bar{t}W$) are considered as minor background processes. Especially the $t\bar{t}Z$ process can cause a signature similar to that of the $t\bar{t}H$ signal if the Z boson decays into a pair of bottom quarks. Nevertheless, the contribution is only minor because the branching fraction for a Z boson decaying to $b\bar{b}$ is only a fourth of the Higgs boson into $b\bar{b}$ branching fraction. Furthermore, the mass of the Z boson is lower than the Higgs boson mass, which results in lower transverse momenta of the bottom quarks. Hence, some of the events are not fulfilling the event selections as presented in Chapter 6 or Chapter 7. The Feynman diagrams for the $t\bar{t}V$ production can be found in Fig. 5.5(a,b).

Single top quark production in association with a Z boson (tZq) is another minor background process. It can result in the same final-state particles as the tH signal processes if the Z boson decays into a pair of bottom quarks. The Feynman diagram for the tZq production can be found in Fig. 5.5(c).

The production of W bosons or Z bosons with jets also contributes to the background events. These processes can be suppressed by requiring a missing transverse momentum above a certain threshold in combination with three or more b-tagged jets. The exact event selections can be found in the respective analysis chapter. The corresponding Feynman diagrams can be found in Fig. 5.5(d,e).

The production of multiple jets via the strong force, i.e., QCD multijet events, is negligible when requiring a certain missing transverse momentum. The production of vector boson pairs has a comparatively low cross section, i.e., about two orders of magnitude smaller than the $t\bar{t}$ production cross section and low branching ratios for signal-like final states. Therefore, it is negligible as well.

5.3. Jet Assignment

The kinematic properties of the final-state particles provide valuable information for the discrimination between signal and background events. Unfortunately, this information cannot be obtained directly. Events as recorded by the detector and reconstructed with the methods described in Section 4.2 feature

jets, leptons, and missing transverse momentum. For data, it is not easily possible to tell which jet should be assigned to which final-state quark.

The most basic approach is to utilize a χ^2-based method, which uses the masses of the reconstructed particles to find the best assignment. As there are further properties of the jets, which can be used to find the best assignment, more sophisticated methods are implemented. A machine learning based method is introduced to assign jets to the final-state quarks. In this thesis, the jet assignment utilizes BDTs. This is implemented for different hypotheses, i.e., interpreting the event as stemming from a certain process. The jet assignment is performed under three signal hypotheses, namely tHq, tHW, and $t\bar{t}H$, and the dominating background hypothesis, i.e., $t\bar{t}$. A jet assignment under other background hypotheses is not implemented as their combined contribution to the background composition is smaller than 10%. Furthermore, the number of simulated events fulfilling the event selections is too small to perform a reasonable training. The four assignments result in four interpretations for each event yielding kinematic variables for each of the reconstructed particles, e.g., the top quark or the Higgs boson. The four-vectors of these particles are obtained by adding the four vectors of the assigned jets. The jet-assignment variables are utilized to enhance the separation of signal and background processes.

The BDTs combine several variables, which include the mass, the transverse momentum and the angular distributions of the assigned jets. A study shows that including the mass of the final-state particles as an input of the BDT yields a higher sensitivity, even though this results in a more signal-like distribution of the Higgs boson mass in background events [154].

In order to train the BDTs, a correctly assigned and a wrongly assigned sample is needed. For each of the hypotheses, the jet assignment BDT (JA BDT) is trained on simulated samples, which only contain events corresponding to the hypothesis in question, e.g., for the training under the tHq hypothesis, the tHq simulation sample is used. Here, the correctly assigned sample contains the correct jet-quark assignment and the wrongly assigned sample contains all wrong jet-quark permutations. The wrongly assigned sample is weighted in a way to give each event the same weight in the training [154]. If this were not the case, events with more jets and therefore a larger number of wrong permutations would contribute more to the wrongly assigned composition than events with fewer jets. This is especially problematic as the number of wrong permutations neglecting additional constraints scales with $n_{\text{jets}}!$, where n_{jets} is the number of jets per event.

The correct and wrong assignments for the training samples can only be determined by using simulation truth information. An angular distance metric

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \tag{5.1}$$

is introduced, where η and ϕ are the coordinates as defined in Chapter 2, and Δ represents the distance between the jet and the quark. The correct jet-quark assignment is defined as the assignment with the lowest sum of ΔR for all jet-quark pairs that are to be assigned. These correct assignments are stored in the correctly assigned sample. Additionally, for each of the quarks, the distance has to be smaller than $\Delta R < 0.4$. This cutoff was chosen as this results in the best results when applying the assignment method to an independent test sample.

Only jets with $\Delta R > 0.4$ between a jet and the lepton are considered in the assignment as these are reconstructed with higher quality. Furthermore, jets with $|\eta| > 2.4$ are required to have $p_{\text{T}} > 40\,\text{GeV}$ and central jets are required to have $p_{\text{T}} > 30\,\text{GeV}$. Jets fulfilling these criteria are referred to as assignable jets.

Events with at least one correct assignment are referred to as matchable events. Only these matchable events are considered in the training process. Assignments, which do not change the properties of the final-state particles are only considered once, e.g., the order of two jets assigned to the Higgs boson does not affect the Higgs boson kinematics. All other assignments are considered wrong. Implementing a more sophisticated approach to define the correct assignment by considering not only the angular distance but also the transverse momentum did not provide better results [155].

The training is performed using the JA BDT configuration as shown in Table 5.1. The JA BDT uses the kinematic properties of the respective final-state particles, as well as their angular distributions,

Table 5.1.: Settings used in the training of the JA BDTs: these settings are used for the jet assignment under the tHq, tHW, $t\bar{t}H$ and $t\bar{t}$ hypotheses. The parameters are defined in Ref. [87].

Parameter	Value
`NTrees`	300
`MinNodeSize`	5%
`MaxDepth`	4
`nCuts`	20
`BoostType`	AdaBoost
`AdaBoostBeta`	0.3
`SeparationType`	GiniIndex

to separate the correctly assigned and wrongly assigned samples. The samples are divided into two categories A and B based on their event ID. Each of the samples contains about 50% of the events. The training is executed in a two-fold approach. Two BDTs are trained: the first on category A events and the second on category B events. For the training on category A events, the category B events are used as a test dataset and vice versa. As only samples that have not been utilized in the respective training, are used in the analysis itself, this approach allows to take advantage of the whole samples in the further steps of the analysis.

The JA BDT is applied on data and simulated events. The JA BDT trained on set A is applied on set B and vice versa. For each event, all possible jet-quark assignments are evaluated, the input variables are calculated and the assignment with the highest JA BDT output is saved. Therefore, for each of the events the best possible assignment under each of the hypotheses is stored. Events with too few jets for a certain assignment receive default values for the respective variables. These default values lie outside of the distributions for the assignable events.

The individual jet assignments under the respective hypotheses are discussed in the following. For all hypotheses, certain restrictions on the assigned jets are introduced to reduce the computing effort. These requirements do not reduce the performance of the assignment as has been studied.

5.3.1. tHq Hypothesis

In the assignment under the tHq hypothesis, three particles are reconstructed: the leptonically decaying top quark, the Higgs boson decaying into a pair of bottom quarks and the light quark. This requires at least four assignable jets: one jet for the bottom quark from the top quark decay, two jets for the Higgs boson and one jet for the light quark.

In order to reduce the computing effort, the number of permutations is decreased: at least one of the jets assigned to the bottom quarks has to be b-tagged. Furthermore, all jets assigned to bottom quarks have to be in the central region with $|\eta| < 2.4$ since only in this region b tagging is possible. Additionally, the jet assigned to the light quark has to be untagged. The variables used in the tHq jet assignment are listed in Table 5.2. The distributions of these variables can be found in the analysis chapters.

5.3.2. tHW Hypothesis

The jet assignment under the tHW hypothesis matches jets to the top quark, the Higgs boson and the hadronically decaying W boson. This results in five jets: one jet from the bottom quark of the top quark decay, two jets for the Higgs boson decay and two quarks for the hadronically decaying W boson. As there are two W bosons present in this process, it is unclear which of them stems from the top quark

Table 5.2.: Input variables used in the tHq JA BDTs: these variables describe the properties of the final-state particles, i.e., the Higgs boson, the top quark, and the light quark. For variables with long tails, e.g., the reconstructed mass of the Higgs boson, the logarithm of the variable is used as narrower distributions are better handled by the BDT. These variables are newly computed for each assignment, i.e., each event has several sets of these variables. In the training step, the set corresponding to the correct assignment is stored in the signal sample and all other sets are stored in the background sample. In the application step, the set resulting in the highest BDT output is chosen and stored in the samples.

Variable	Description
	Higgs boson related variables
log m(H)/GeV	invariant mass of the reconstructed Higgs boson
log p_{T}(H)/GeV	transverse momentum of the reconstructed Higgs boson
ΔR(H)	ΔR between the two jets assigned to the Higgs boson decay
CSV($\mathrm{b}_{\mathrm{H},1}$)	output of the b-tagging algorithm for the first jet assigned to the Higgs boson decay
CSV($\mathrm{b}_{\mathrm{H},2}$)	output of the b-tagging algorithm for the second jet assigned to the Higgs boson decay
	top quark related variables
log m(t)/GeV	invariant mass of the reconstructed top quark
log p_{T}(t)/GeV	transverse momentum of the reconstructed top quark
$\Delta R(\mathrm{b_t}$,W)	ΔR between the jet assigned to the bottom quark of the top quark decay and the leptonically decaying W boson
CSV($\mathrm{b_t}$)	output of the b-tagging algorithm for the jet assigned to the bottom quark of the top quark decay
$\lvert\eta(\mathrm{b_t})\rvert$	absolute pseudorapidity of the jet assigned to the bottom quark of the top quark decay
	light quark related variables
log p_{T}(light jet)/GeV	transverse momentum of the jet assigned to the light quark
CSV(light jet)	output of the b-tagging algorithm for the jet assigned to the light quark
$\lvert\eta(\textbf{light jet})\rvert$	absolute pseudorapidity of the jet assigned to the light quark
	further variables
relative $\mathrm{H_T}$	scalar sum of the transverse momentum for all jets assigned under the tHq hypothesis divided by the scalar sum of the total transverse momentum (jets, lepton, $p_{\mathrm{T}}^{\mathrm{miss}}$)
$\Delta R(\mathrm{t},\mathrm{H})$	ΔR between the reconstructed top quark and the reconstructed Higgs boson
$\lvert\eta(\textbf{light jet}) - \eta(\mathrm{b_t})\rvert$	absolute difference in pseudorapidity of the jet assigned to the light quark and the reconstructed top quark

decay and which one is the additional W boson. This decision is also made by the JA BDT. Similar to the jet assignment itself, the BDT output is calculated for both matchings of the W boson and the assignment with the higher JA BDT output is chosen.

As a consequence of the additional quark compared to the tHq hypothesis, as well as the ambiguity of the W boson assignment, the computing effort is even higher than for the tHq assignment. In order to reduce this effort, requirements on the permutations are set: at least one of the jets assigned to bottom quarks has to be b tagged and all jets assigned to bottom quarks are required to be in the central region. The variables of the tHW JA BDT are listed in Table 5.3. The distributions of these variables can be found in the analysis chapters. Thorough studies on the impact of the tHW jet assignment on the tH analysis can be found in Ref. [156].

5.3.3. $t\bar{t}H$ Hypothesis

In the $t\bar{t}H$ jet assignment, jets are matched to the decay products of the two top quarks and the Higgs boson. In total, six quarks are matched: one bottom quark from the leptonic top quark decay, a bottom quark and two light quarks from the hadronic top quark decay and two bottom quarks from the Higgs boson decay. This results in a larger wrongly assigned sample compared to the tH samples, hence the computing effort is even higher than in the tHW jet assignment. Here, two of the jets assigned to the bottom quarks are required to be tagged. Furthermore, all jets assigned to bottom quarks have to lie in the central region. The $t\bar{t}H$ JA BDT is trained with the variables listed in Table 5.4. The $t\bar{t}H$ jet assignment is only performed in the analysis presented in Chapter 7 since this analysis is tailored to $t\bar{t}H$ and tH production. The distributions of the input variables can be found there as well.

5.3.4. $t\bar{t}$ Hypothesis

In addition to the jet assignment for the signal processes as described above, an assignment under the hypothesis of the dominating background, i.e., the $t\bar{t}$ process is performed. Therefore, jets are matched to the hadronic decay products of the two top quarks produced in this process. This leads to four quarks in total: one bottom quark from the leptonic top quark decay, and a bottom and two light quarks from the hadronic top quark decay.

The computing effort is reduced by requiring at least one b-tagged jet to be assigned to a bottom quark. Again, jets assigned to bottom quarks are required to lie in the central region of the detector. The input variables for the $t\bar{t}$ JA BDT are listed in Table 5.5. The distributions of these variables can be found in the analysis chapters.

5.3.5. Evaluation

The JA BDT can be evaluated by calculating in what fraction of the events it is able to find the correct assignment. For this, a statistically independent simulated sample representing the same hypothesis is used, i.e., a $t\bar{t}$ sample is used to evaluate the $t\bar{t}$ JA BDT. In a first step, the JA BDT is evaluated on this sample, resulting in a jet-quark assignment for each of the events. In a second step, the correct assignment is derived by using the truth information of the simulated sample. These two assignments are compared for all events, for which a correct assignment is found. If both assignments agree, all jets of the event are correctly assigned. The ratio between correctly assignment events and total number of events, for which a correct assignment would have been possible is called assignment efficiency. These efficiencies can then be compared to other jet assignment methods, e.g., a χ^2 method. The efficiencies for the respective trainings, as well as the distributions of the outputs can be found in the analyses presented in Chapters 6 and 7.

Table 5.3.: Input variables used in the tHW JA BDTs: these variables describe the properties of the final-state particles, i.e., the Higgs boson, the top quark, and the additional W boson. In order to figure out which of the W bosons decays leptonically, also properties of the hadronically and leptonically decaying W bosons are utilized. For variables with long tails, e.g., the reconstructed mass of the Higgs boson, the logarithm of the variable is used as narrower distributions are better handled by the BDT. These variables are newly computed for each assignment, i.e., each event has several sets of these variables. In the training step, the set corresponding to the correct assignment is stored in the signal sample and all other sets are stored in the background sample. In the application step, the set resulting in the highest BDT output is chosen and stored in the samples.

Variable	Description
	Higgs boson related variables
log m(H)/GeV	invariant mass of the reconstructed Higgs boson
log p_T(H)/GeV	transverse momentum of the reconstructed Higgs boson
CSV($b_{H,1}$)	output of the b-tagging algorithm for the first jet assigned to the Higgs boson decay
CSV($b_{H,2}$)	output of the b-tagging algorithm for the second jet assigned to the Higgs boson decay
ΔR(H)	ΔR between the two jets assigned to the Higgs boson decay
	top quark related variables
log m(t)/GeV	invariant mass of the reconstructed top quark
log p_T(t)/GeV	transverse momentum of the reconstructed top quark
CSV(b_t)	output of the b-tagging algorithm for the jet assigned to the bottom quark of the top quark decay
$\lvert\eta(t)\rvert$	absolute pseudorapidity of the reconstructed top quark
$\lvert\eta(b_t)\rvert$	absolute pseudorapidity of the jet assigned to the bottom quark of the top quark decay
ΔR(b_t,W_{lep})	ΔR between the jet assigned to the bottom quark of the top quark decay and the leptonically decaying W boson
	W boson related variables
log m(W_{had})/GeV	invariant mass of the reconstructed hadronically decaying W boson
log p_T(W_{had})/GeV	transverse momentum of the reconstructed hadronically decaying W boson
CSV(W_{had} jet 1)	output of the b-tagging algorithm for the first jet assigned to the hadronically decaying W boson
CSV(W_{had} jet 2)	output of the b-tagging algorithm for the second jet assigned to the hadronically decaying W boson
log m(W_b)/GeV	invariant mass of the reconstructed additional W boson
log p_T(W_b)/GeV	transverse momentum of the reconstructed additional W boson
p_T(W_{lep})-p_T(W_{had})	difference in transverse momentum between reconstructed leptonically decaying and hadronically decaying W boson
$\lvert\eta(W_{lep}) - \eta(W_{had})\rvert$	absolute difference in pseudorapidity of the reconstructed leptonically decaying and hadronically decaying W boson
$\lvert\eta(W_b)\rvert$	absolute pseudorapidity of the reconstructed additional W boson
	further variables
relative H_T	scalar sum of the transverse momentum for all jets assigned under the tHW hypothesis divided by the scalar sum of the total transverse momentum (jets, lepton, p_T^{miss})
cos $\theta(b_t, \ell)$	cosine of the angle between the lepton and the jet assigned to the bottom quark from the top quark decay

Table 5.4.: Input variables used in the $t\bar{t}H$ JA BDTs: these variables describe the properties of the final-state particles, i.e., the Higgs boson, the leptonically decaying top quark, and the hadronically decaying top quark. For variables with long tails, e.g., the reconstructed mass of the Higgs boson, the logarithm of the variable is used as narrower distributions are better handled by the BDT. These variables are newly computed for each assignment, i.e., each event has several sets of these variables. In the training step, the set corresponding to the correct assignment is stored in the signal sample and all other sets are stored in the background sample. In the application step, the set resulting in the highest BDT output is chosen and stored in the samples.

Variable	Description
	Higgs boson related variables
log m(H)/GeV	invariant mass of the reconstructed Higgs boson
log p_T(H)/GeV	transverse momentum of the reconstructed Higgs boson
ΔR(H)	ΔR between the two jets assigned to the Higgs boson decay
CSV($b_{H,1}$)	output of the b-tagging algorithm for the first jet assigned to the Higgs boson decay
CSV($b_{H,2}$)	output of the b-tagging algorithm for the second jet assigned to the Higgs boson decay
	leptonically decaying top quark related variables
log m(t_{lep})/GeV	invariant mass of the reconstructed leptonically decaying top quark
log p_T(t_{lep}/GeV)	transverse momentum of the reconstructed leptonically decaying top quark
	hadronically decaying top quark related variables
log p_T(t_{had})/GeV	transverse momentum of the reconstructed hadronically decaying top quark
log Δm(t_{had},W_{had})/GeV	difference between the invariant masses of the reconstructed hadronically decaying top quark and the reconstructed hadronically decaying W boson
log m(W_{had})/GeV	invariant mass of the two jets assigned to the hadronically decaying W boson of the top quark decay
CSV(W_{had} jet 1)	output of the b-tagging algorithm for the first jet assigned to the hadronically decaying W boson
CSV(W_{had} jet 2)	output of the b-tagging algorithm for the second jet assigned to the hadronically decaying W boson
ΔR (W_{had})	ΔR between the two jets assigned to the W boson of the hadronically decaying top quark
	further variables
relative H_T	scalar sum of the transverse momentum for all jets assigned under the $t\bar{t}H$ hypothesis divided by the scalar sum of the total transverse momentum (jets, lepton, p_T^{miss})

Table 5.5.: Input variables used in the $t\bar{t}$ JA BDTs: these variables describe the properties of the final-state particles, i.e., the leptonically and hadronically decaying top quark. For variables with long tails, e.g., the reconstructed mass of the top quark, the logarithm of the variable is used as narrower distributions are better handled by the BDT. These variables are newly computed for each assignment, i.e., each event has several sets of these variables. In the training step, the set corresponding to the correct assginment is stored in the signal sample and all other sets are stored in the background sample. In the application step, the set resulting in the highest BDT output is chosen and stored in the samples.

Variable	Description
	leptonically decaying top quark related variables
log m(t_{lep})/GeV	invariant mass of the reconstructed leptonically decaying top quark
log p_T(t_{lep})/GeV	transverse momentum of the reconstructed leptonically decaying top quark
	hadronically decaying top quark related variables
log p_T(t_{had})/GeV	transverse momentum of the reconstructed hadronically decaying top quark
log Δm(t_{had},W_{had})/GeV	difference between the invariant masses of the reconstructed hadronically decaying top quark and the reconstructed hadronically decaying W boson
log m(W_{had})/GeV	invariant mass of the two jets assigned to the hadronically decaying W boson of the top quark decay
CSV(W_{had} jet 1)	output of the b-tagging algorithm for the first jet assigned to the hadronically decaying W boson
CSV(W_{had} jet 2)	output of the b-tagging algorithm for the second jet assigned to the hadronically decaying W boson
ΔR (W_{had})	ΔR between the two jets assigned to the W boson of the hadronically decaying top quark
	further variables
relative H_T	scalar sum of the transverse momentum for the jets assigned under the $t\bar{t}$ hypothesis divided by the scalar sum of the total transverse momentum (jets, lepton, $\vec{p}_T^{\,miss}$)

6. Search for tH with H $\rightarrow b\bar{b}$ with the 2016 Data Set

The analysis presented in this chapter is performed with the data set recorded by the CMS experiment in 2016 with an integrated luminosity of 35.9 fb^{-1}. It is tailored to the search for single top quark production in association with a Higgs boson, where the Higgs boson decays into a pair of bottom quarks and the top quark decays leptonically. The signal and background processes considered in this analysis are described in the previous chapter. The tH process gives direct access to the top quark Yukawa coupling. Although the SM cross section of the tH process is very small, deviations in the Higgs boson couplings can lead to a strong increase of the cross section. A scan of the ratio between the top quark Yukawa coupling and the Higgs boson to W boson coupling is performed. Furthermore, limits on possible CP-violating contributions in the Higgs boson coupling are set. Multivariate methods, as described previously, are used to assign the jets of the high multiplicity final states to the quarks under different hypotheses. Based on these assignments, a classifier is trained to discriminate signal from background events.

A special focus lies on two specific scenarios: the SM scenario with $\kappa_t = \kappa_V = 1$ and the inverted top coupling (ITC) case with $-\kappa_t = \kappa_V = 1$. These two scenarios were already studied in the first run period of the LHC in several Higgs bosons decay channels [157] and on the data taken in 2015 by the CMS experiment [158]. In addition, the run 2 analysis features a scan of the κ_t/κ_V-ratio for values of $-4 \leq \kappa_t/\kappa_V \leq 4$.

The analysis presented in this chapter has been published by the CMS collaboration [1]. In addition, it is also part of Ref. [159]. However, several improvements compared to the published analysis are introduced in this chapter: the b-jet energy regression as described in Section 6.2 is added. Moreover, additional signal events are simulated, which expand the number of simulated events by a factor of 2.5 for the tHq and 3.5 for the tHW process. This leads to better results in the training of the machine learning methods and to smaller statistical uncertainties associated with the simulated samples. Additionally, a two-fold training approach is introduced, which increases the fraction of events used for the training from 20% in the published analysis to 100% in this thesis. This also increases the events used in the final fit to 100% of the simulated samples compared to 60% in the published analysis. Therefore, the uncertainties are reduced further and the training results are even better.

The search for single top quark production in association with a Higgs boson is also performed in the multilepton decay channel [160]. On top of that, the results of the diphoton [161] channel have been reinterpreted to search for tH production. Consequently, the results of these channels are combined with the $b\bar{b}$ decay channel [2], which is also part of the work presented in this thesis. The results of the combination are shown in Section 6.9.

6.1. Search Strategy

This analysis is optimized for the search for tH production at a center-of-mass energy of 13 TeV in the Higgs boson to $b\bar{b}$ decay channel with a leptonically decaying top quark. The SM cross section at NLO is

$$\sigma_{tHq}^{SM} = 74.3^{+4.8}_{-11.1}\ (\text{scale}) \pm 2.7\ (\text{PDF} + \alpha_S)\ \text{fb} \tag{6.1}$$

for the tHq process and

$$\sigma_{tHW}^{SM} = 15.2^{+0.7}_{-1.0}\ (\text{scale}) \pm 1.0\ (\text{PDF} + \alpha_S)\ \text{fb} \tag{6.2}$$

for the tHW process [40]. Since the signal samples are produced at LO, the cross section is scaled to the NLO cross section by applying a scale factor. The tH production is sensitive to the coupling modifiers κ_V and κ_t, see Equations (1.52) and (1.53). Here, the vector boson coupling is assumed to be equal to the W boson coupling, i.e., $\kappa_V = \kappa_W$. Furthermore, $t\bar{t}H$ production is considered as signal, too. The $t\bar{t}H$ process with an SM cross section of

$$\sigma^{SM}_{t\bar{t}H} = 507.1^{+11.1}_{-46.6}\ (\text{scale}) \pm 18.3\ (\text{PDF} + \alpha_S)\ \text{fb} \tag{6.3}$$

provides additional sensitivity to the absolute value of the top quark Yukawa coupling [40]. The cross sections as a function of the coupling modifiers are shown in Fig. 1.11. Modifying the Higgs boson couplings does not only affect the cross section of the signal processes but also the kinematic properties of the final-state products for the tHq and tHW processes as shown in Appendix A.2.

For the ITC scenario, the cross section of the tHq process is enhanced by roughly a factor of eleven to

$$\sigma^{ITC}_{tHq} = 792.7^{+21.4}_{-30.9}\ (\text{scale}) \pm 7.1\ (\text{PDF} + \alpha_S)\ \text{fb} \tag{6.4}$$

and by a factor of ten for the tHW process:

$$\sigma^{SM}_{tHW} = 147.2^{+2.9}_{-2.6}\ (\text{scale}) \pm 5.7\ (\text{PDF} + \alpha_S)\ \text{fb}. \tag{6.5}$$

As the $t\bar{t}H$ process is insensitive to the sign of the top quark Yukawa coupling, the cross section of the SM and the ITC point are equal for this process.

This analysis does not only study these two points but a range of coupling scenarios. Since the kinematic properties of the processes are the same for a certain coupling ratio κ_t/κ_V and only the cross section is modified, it is beneficial to set limits on $\sigma \times \text{BR}$ for different ratios. Here, a range of $-4 \leq \kappa_t/\kappa_V \leq 4$ is covered. Technically, a grid of different κ_t-κ_V scenarios is evaluated, from which the limits on the ratio are derived. The cross section for these specific points in the coupling plane can be found in Appendix A.1.

In addition to the scan of the coupling ratio, limits on possible CP violation in the Higgs boson couplings are determined. The SM scenario corresponds to the CP-even state with a mixing angle of $\alpha = 0^\circ$ and the ITC point corresponds to $\alpha = 180^\circ$. In total, 21 points for $-1 \leq \cos\alpha \leq 1$ in steps of $\Delta\cos\alpha = 0.1$ are evaluated.

The processes studied in this thesis all lead to several bottom quarks in the final state. The jets produced by these quarks typically loose some of their energy to neutrinos produced in leptonic decays. Furthermore, jets originating from bottom quarks have other properties than jets from light quarks. An energy regression is implemented to mitigate these effects.

In order to suppress the background events, selections are applied to obtain signal-like signatures, e.g., by requiring a certain number of jets, b tags or leptons. Based on these selected events, the jet assignment as described in Section 5.3 is applied to obtain hypothesis-specific variables. These variables are used, together with global variables, to train a classification BDT, which discriminates between signal and background processes. The BDT output distributions are then fitted to obtain exclusion limits on the SM tH production, the coupling ratio and the CP-mixing angle α.

The tHq, tHW and $t\bar{t}H$ signal processes are described in Section 5.1. The event topology of the $t\bar{t}$, the single top quark and the minor background processes can be found in Section 5.2.

6.2. b-jet Energy Regression

In this analysis, top quarks and Higgs bosons are the particles of interest. The top quark decays almost exclusively into a bottom quark and a W boson and the Higgs boson is required to decay into bottom quarks. Hence, an optimal measurement of the properties of the jets produced by bottom quarks, called b jets, is favorable. The energy of b jets determined by the jet clustering algorithm considers the energy of the particles produced by the jet, which leave a trace in the detector. For decays that produce

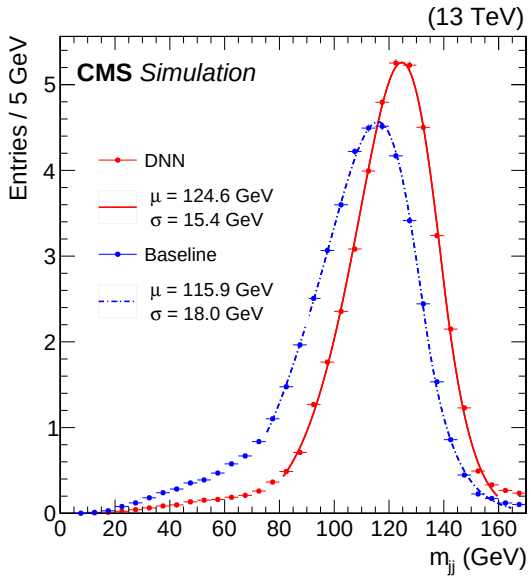

Figure 6.1.: Improvement for the dijet mass resolution when applying the b-jet energy regression on a simulated Z($\ell\ell$)H($b\bar{b}$) sample: the baseline mass distribution is shown in blue and the improved mass distribution by applying the b-jet energy regression is shown in red. The mean value is shifted toward the Higgs boson mass that has been used in the simulation (125 GeV) and the width of the distribution is reduced by 2.6 GeV. Taken from Ref. [162].

neutrinos, some of the energy of the b jet escapes the detector without being measured. The b-jet energy regression [162] tries to compensate this loss by implementing a DNN to obtain a correction factor to the transverse momentum of the jet. The impact of the b-jet energy regression on the tH analysis, as described in the following, has been studied in detail in Ref. [163].

Jets originating from bottom quarks have several features that distinguish them from jets stemming from light quarks. As already discussed in Section 4.2.5 on b tagging, these jets often produce secondary vertices. Furthermore, the mass of B hadrons is larger than for the hadrons produced in light-flavored jets. Therefore, the decay products within the b jets have higher momenta with respect to the jet axis, which leads to broader energy distributions within the jet. Additionally, the energy carried by neutrinos affects the energy distribution within the jet as well. Consequently, variables describing the energy distribution within a jet are used as an input to the b-jet energy regression DNN.

The input variables of the DNN are derived from the raw jets without any corrections since the DNN learns the corrections, obtains a bias and produces different results for simulated events and data otherwise. Nonetheless, the output nodes use the corrected jet momenta. The DNN has three output nodes: the 25% and the 75% quantile, and the correction factor

$$y = \frac{p_{\mathrm{T}}^{\mathrm{gen}}}{p_{\mathrm{T}}^{\mathrm{reco}}(\mathrm{JEC})}, \tag{6.6}$$

with the generated transverse momentum $p_{\mathrm{T}}^{\mathrm{gen}}$ and the reconstructed transverse momentum after the JEC $p_{\mathrm{T}}^{\mathrm{reco}}(\mathrm{JEC})$. The quantiles can be used to estimate the resolution of the jet. The correction factor is then applied to the transverse momentum of the jet after the JEC. All variables that are dependent on the jet momentum are recomputed with the regressed jets, e.g., the missing transverse energy. The DNN uses over 40 input variables, which describe the jets, pileup, the leptons, the secondary vertex, the spatial energy distribution and the energy distribution among the jet constituents. The full list can be found in Ref. [163].

The DNN comprises six layers containing between 1024 and 128 nodes each. Each of the layers consists of a dense unit, a batch normalization layer, a dropout unit and an activation function. All of these constituents are described in Section 3.6. The training is performed on a simulated $t\bar{t}$ sample. Only jets with generated and reconstructed transverse momentum above 15 GeV and $|\eta| < 2.5$ are used in the training and application of the regression. Jets outside of the η region do not have any b tag information and transverse momenta below the p_T threshold are typically not considered in an analysis.

The improvement for the dijet mass resolution of the reconstructed Higgs boson mass is shown in Fig. 6.1. The mean value is regressed toward 125 GeV and the width of the distribution is decreased. This is exactly the desired behavior of the energy regression.

6.3. Event Selection

The signal processes studied in this analysis contain certain signatures. Based on these signatures, events are selected to obtain a signal enriched phase space, especially targeting the tHq process. For all of the signal processes, exactly one leptonically decaying top quark is required. This leads to one lepton and the corresponding neutrino, which manifests as missing transverse momentum in the detector. For the tHq process, a bottom quark is produced in the top quark decay, two bottom quarks stem from the Higgs boson decay, and one bottom quark originates from the gluon splitting. Additionally, the tHq process features the characteristic light quark. Within this analysis, the event selection is tailored to this process.

An event selection requiring exactly four b-tagged jets, at least one additional untagged jet, exactly one lepton (electron or muon), and a certain missing transverse momentum defines one signal region, called 4m region. The CSVv2 algorithm [147, 148] is used within this analysis for b tagging.

A jet is only considered in the selection if it has a $p_T > 30\,\text{GeV}$ in the central region ($|\eta| < 2.4$) or $p_T > 40\,\text{GeV}$ in the forward region ($|\eta| \geq 2.4$). Jets are considered as b tagged if they fulfill the medium working point of the b tagger. This corresponds to a mistag rate of 1%. As a result of the detector geometry, these b-tagged jets are required to be in the central region. The electrons (muons) are required to fulfill the tight ID flag and $p_T > 35\,\text{GeV}$ (27 GeV). Moreover, the missing transverse momentum is required to be $p_T^{\text{miss}} > 45\,\text{GeV}$ (35 GeV). The different requirements for the missing transverse momentum are chosen based on the study presented in Ref. [164] to suppress QCD processes.

As the bottom quark of the gluon splitting often has a low transverse momentum, it does not always fulfill the p_T requirements of the selection. Therefore, another signal region is defined, which has only one difference compared to the 4m region: it requires only three b-tagged jets. This region is called 3m region.

Furthermore, a 2m control region is introduced to validate the data-to-simulation agreement as shown in Fig. 6.2. This region requires the same objects as the signal regions, except for requiring exactly two b-tagged jets. The selection requirements for the single-lepton regions are summarized in Table 6.1.

The systematic uncertainties on the flavor composition of the $t\bar{t}$+jets background are large. Especially, the $t\bar{t}+b\bar{b}$ contribution is of interest as it has a signature similar to that of the signal processes. Consequently, another region is introduced to target this background process, which is then used in the final fit to constrain the uncertainties. In order to avoid any overlap with the signal regions, dileptonic $t\bar{t}$ events are targeted in this region. Therefore, this region is called dileptonic background region (BR). The flavor composition of the additional jets is independent of the $t\bar{t}$ decay mode. Hence, the dileptonic channel of the $t\bar{t}$ process can be used without a caveat.

The dileptonic BR contains events with two muons ($\mu\mu$), an electron and a muon ($e\mu/\mu e$) or two electrons (ee). In case of one muon and one electron, the particle first mentioned has the higher transverse momentum. As a consequence of the multiple combinations, different trigger paths are utilized. The selection criteria of the dileptonic BR are summarized in Table 6.2. Events passing at least one of the triggers are further required to have exactly one tight and one additional loose lepton. As the tight leptons also fulfill the loose criteria, events with two tight leptons are accepted as well. The tight and loose

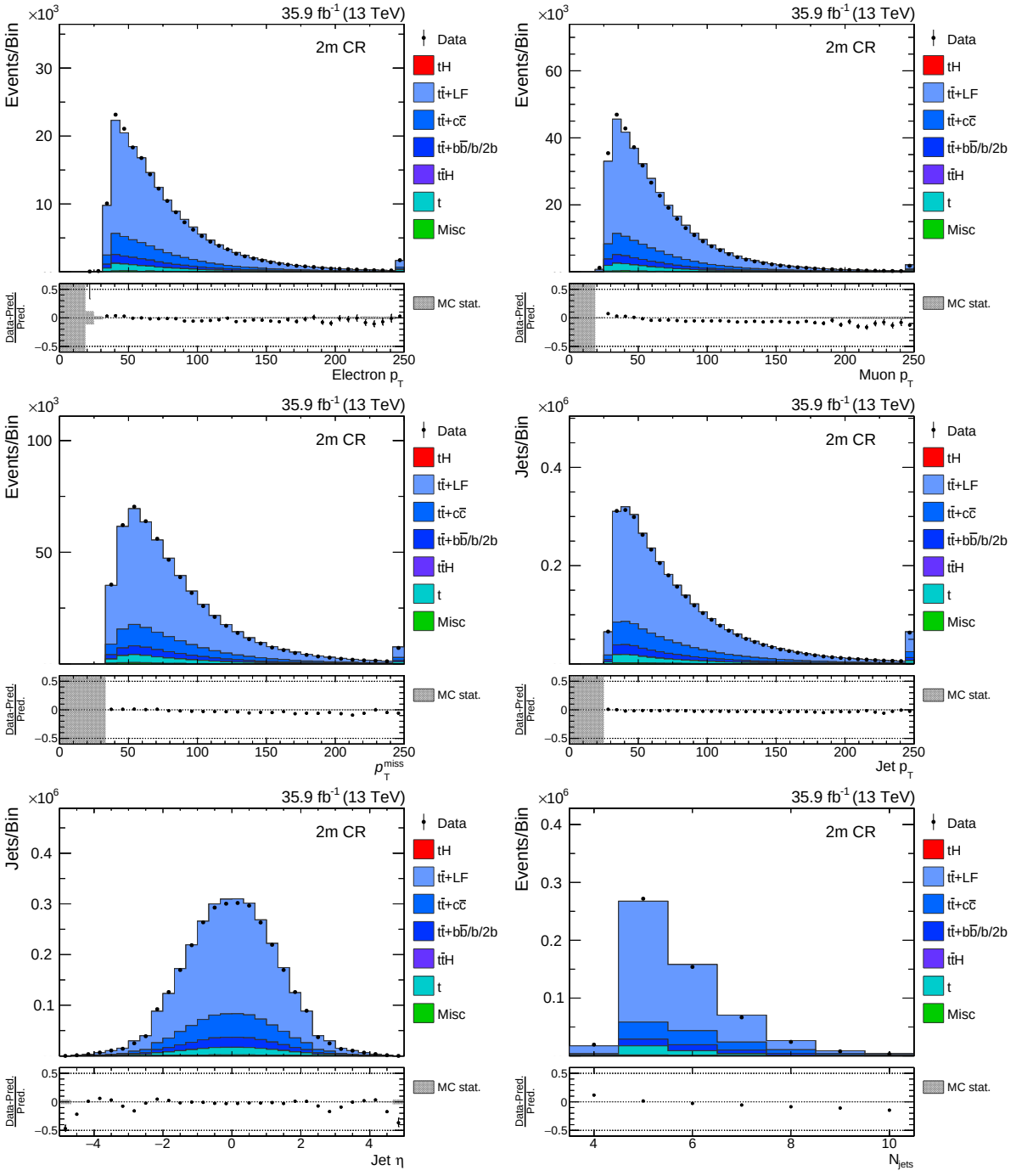

Figure 6.2.: Modeling of event variables: the distributions are shown for the 2m CR. The data is well modeled for all of the distributions.

Table 6.1.: Selection criteria for the single-lepton regions: each event has to pass at least one of the triggers, all of them require exactly one tight muon or electron. Furthermore, the lepton, p_T^{miss}, and jet selections have to be fulfilled. The objects are defined in Section 4.2.

	2m CR	**3m SR**	**4m SR**
muon triggers	HLT_IsoMu24_v* HLT_IsoTkMu24_v*		
electron trigger	HLT_Ele32_eta2p1_WPTight_Gsf_v*		
# tight leptons	1	1	1
# additional loose leptons	0	0	0
p_T^{miss}	$> 45/35\,\text{GeV}(e/\mu)$	$> 45/35\,\text{GeV}(e/\mu)$	$> 45/35\,\text{GeV}(e/\mu)$
# jets	≥ 4	≥ 4	≥ 5
# b-tagged jets	2	3	4

muons are required to have a transverse momentum of $p_T \geq 20\,\text{GeV}$ and the loose (tight) electrons have to fulfill $p_T \geq 15\,\text{GeV}$ (20 GeV). The purity of the region is enhanced by requiring a missing transverse momentum of $p_T^{miss} \geq 40\,\text{GeV}$, which suppresses Z+jets events to a fraction of 4%. In addition, at least two jets have to fulfill the medium b-tag criterion. Additionally, one more jet is required to be loosely tagged, where the loose working point corresponds to a mistag rate of 10%. Similar to the single lepton regions, the b-tagged jets have to lie in the central region with a transverse momentum of $p_T \geq 30\,\text{GeV}$. This region is introduced based on studies of the tH production on the data taken in 2015 by the CMS experiment [165].

The yields of the individual processes for the 3m and 4m SR, as well as the dileptonic BR are listed in Table 6.3. The yield of the Z+jets production is very low in the signal regions, therefore it is not considered in these regions. This is also the case for the tZq production in the dileptonic BR.

Table 6.2.: Selection criteria for the dileptonic background region: each event has to pass one of the triggers and fulfill the lepton, $p_{\mathrm{T}}^{\mathrm{miss}}$, and jet selections to be assigned to the dileptonic category. The objects are defined in Section 4.2.

	dileptonic BR
ee trigger	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_vX
eμ trigger	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 trigger	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_vX HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_vX
μμ trigger	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_vX HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_vX
# tight leptons	1
# additional loose leptons	1
$p_{\mathrm{T}}^{\mathrm{miss}}$	$> 40\,\mathrm{GeV}$
# b-tagged jets (medium)	≥ 2
# b-tagged jets (loose)	≥ 3

Table 6.3.: Prefit expected event yields: the prefit event yields of all considered processes are listed. The Z+jets production is negligible in the signal regions, therefore it is not considered. The same holds for the tZq process in the dileptonic BR. The yields for the $t\bar{t}H$ process are identical for the SM and the ITC case since it is only sensitive to the magnitude of κ_t.

Process	**3m SR**	**4m SR**	**dileptonic BR**
$t\bar{t}$+lf	29475±8159	477±259	7469±1462
$t\bar{t}$+$c\bar{c}$	10195±4878	455±301	3303±1687
$t\bar{t}$+b	4733±1958	253±137	1269±560
$t\bar{t}$+$b\bar{b}$	4720±1828	992±414	1226±512
$t\bar{t}$+2b	2570±1057	183±86	640±299
Single top	2448±659	105±27	435±53
$t\bar{t}Z$	222±48	37.3±7.9	107±49
$t\bar{t}W$	101±25	5.7±14.5	47.6±15.1
tZq	31.3±11.5	4.2±22.5	-
Z+jets	-	-	642±333
Sum of backgrounds	54495±9955	2512±598	15139±2401
$t\bar{t}H$ (SM/ITC)	288±55	73.0±7.5	76.8±10.6
tHq (SM)	12.9±3.7	1.6±0.5	0.5±0.2
tHW (SM)	8.4±1.8	1.3±0.2	2.3±0.5
tHq (ITC)	147±39	18.5±5.6	5.2±1.7
tHW (ITC)	79.3±13.6	12.2±2.2	22.0±3.9
Observed	53555	2627	15140

6.4. Corrections of Simulated Events

The simulation does not always deliver the same distributions as the data. Furthermore, the configuration of the MC event generators is set at a time when the exact conditions for the data taking are unknown. Therefore, some differences occur between data and simulated samples. These differences are mitigated by applying dedicated corrections. Here, corrections for the b-tagging and lepton efficiencies, as well as for the number of primary vertices, are applied as explained in the following.

b-Tagging Scale Factors

This analysis studies events with several b-tagged jets. Therefore, a good description of these jets in simulated events is desirable. The difference in the tagging efficiencies for data and simulation is mitigated by applying event weights obtained with a tag-and-probe method [149]. These weights change the whole output distribution of the b tagger in a way that the simulation to data agreement is improved. The scale factors as a function of p_T and η are derived for heavy-flavored and light-flavored jets separately. The light flavor correction is obtained from Z+jets events and the heavy-flavor scale factor is derived from dileptonic $t\bar{t}$ events. For light-flavored jets, the correction factor is given by

$$\mathrm{SF_{LF}}(p_T, \eta, \mathrm{CSV}) = \frac{\mathrm{Data} - \mathrm{MC_{HF}}}{\mathrm{MC_{LF}}}, \tag{6.7}$$

and for heavy-flavored jets, the scale factor is:

$$\mathrm{SF_{HF}}(p_T, \eta, \mathrm{CSV}) = \frac{\mathrm{Data} - \mathrm{MC_{LF}}}{\mathrm{MC_{HF}}}. \tag{6.8}$$

These scale factors are applied to the simulated events. The scale factor $\mathrm{SF_{LF}}$ is applied to light-flavored jets and $\mathrm{SF_{HF}}$ is applied to b jets, whereas c jets do not receive any correction factor. The event weight is the product of all selected jets of an event:

$$\mathrm{weight_{CSV}} = \prod_{i}^{N_{\mathrm{jets}}} \mathrm{SF}_i. \tag{6.9}$$

The effect of the b-tagging scale factors on the distributions of the CSV output is shown in Fig. 6.3.

Lepton Efficiency Scale Factors

The reconstruction of leptons as described in Section 4.2 introduces several selection efficiencies stemming from the tracking, trigger, isolation and the ID flags. These efficiencies differ between data and simulation. Hence, lepton efficiency scale factors are introduced to mitigate these differences. Dedicated corrections are provided by the CMS Muon POG for muons and the E/γ POG for electrons [166–168]. This results in event-specific corrections for the simulated samples. The distributions of the lepton kinematics with and without the corrections are shown in Fig. 6.4.

Pileup Reweighting

The number of primary vertices is adjusted in the simulated samples with an event-based pileup weight [169] to fit the distribution obtained from the measurement. The weights are determined from minimum-bias events. The distribution of the primary vertices per event with and without the correction is shown in Fig. 6.5. The pileup weights mitigate the difference between data and simulation.

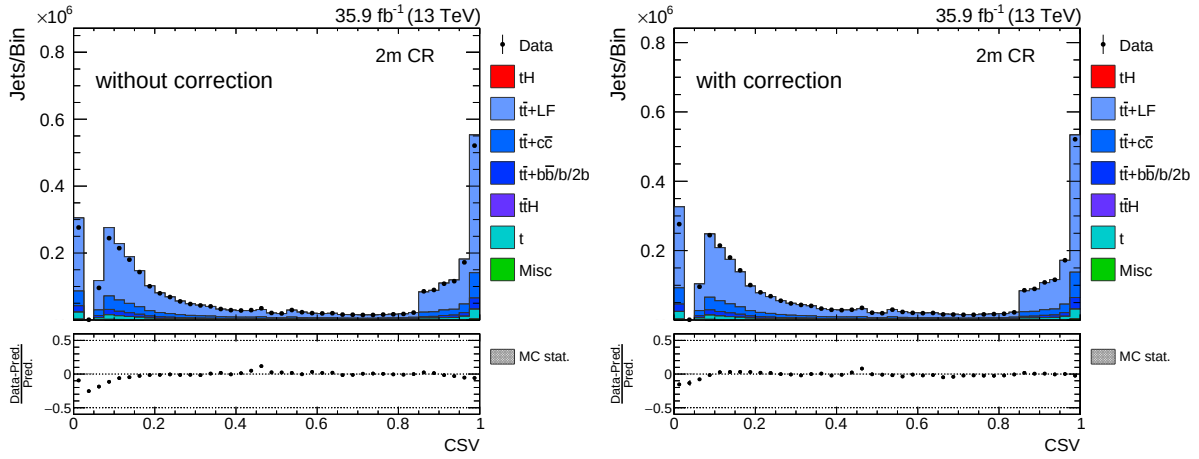

Figure 6.3.: Influence of the b-tagging scale factors: the distribution on the left shows the CSV output of all jets without applying the scale factors and the right figure shows the distribution with the correction. The modeling of the b-tagger output is slightly improved by the correction.

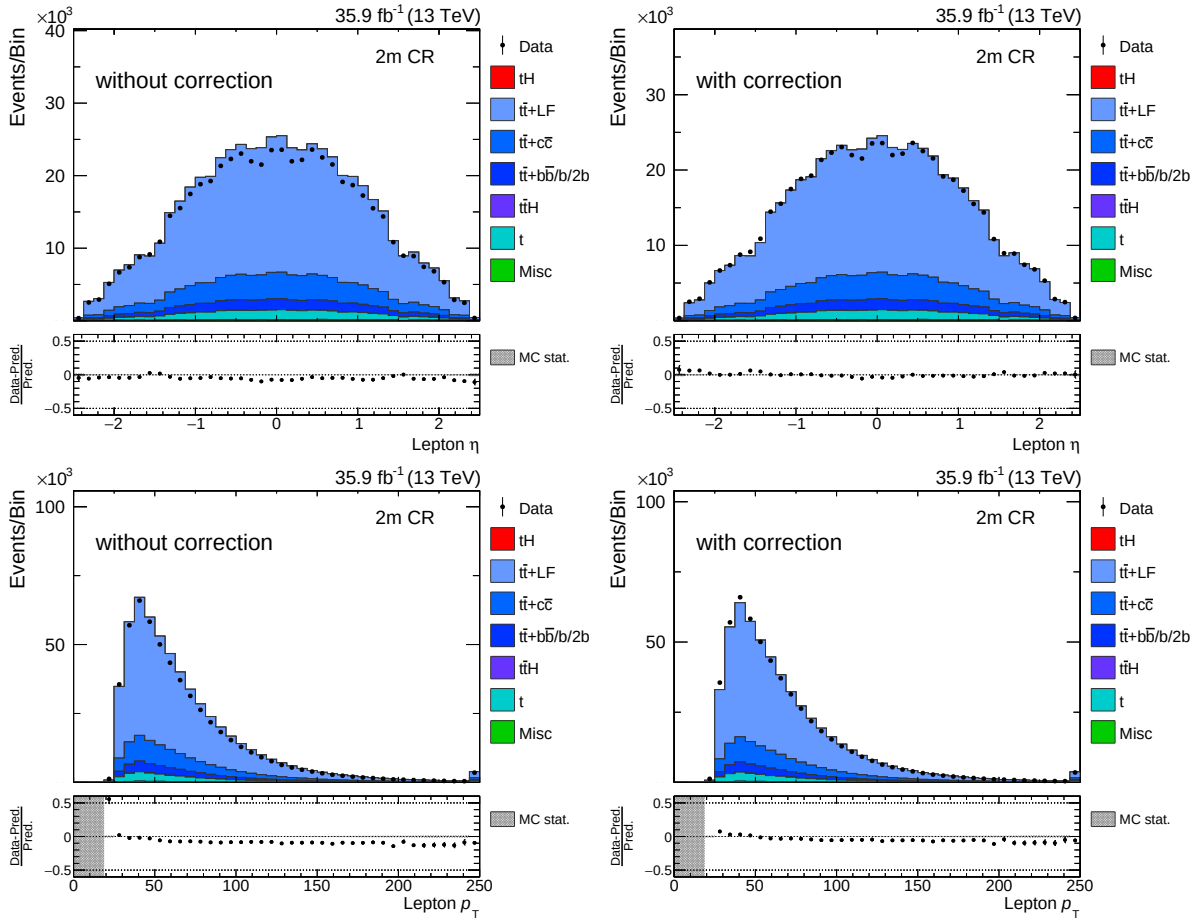

Figure 6.4.: Influence of the lepton scale factors: the distributions on the left show the η (top) and p_T (bottom) of the leptons without applying the scale factors and the figures on the right show the distributions with the correction applied. The distributions with the lepton scale factors applied show good agreement between simulation and data.

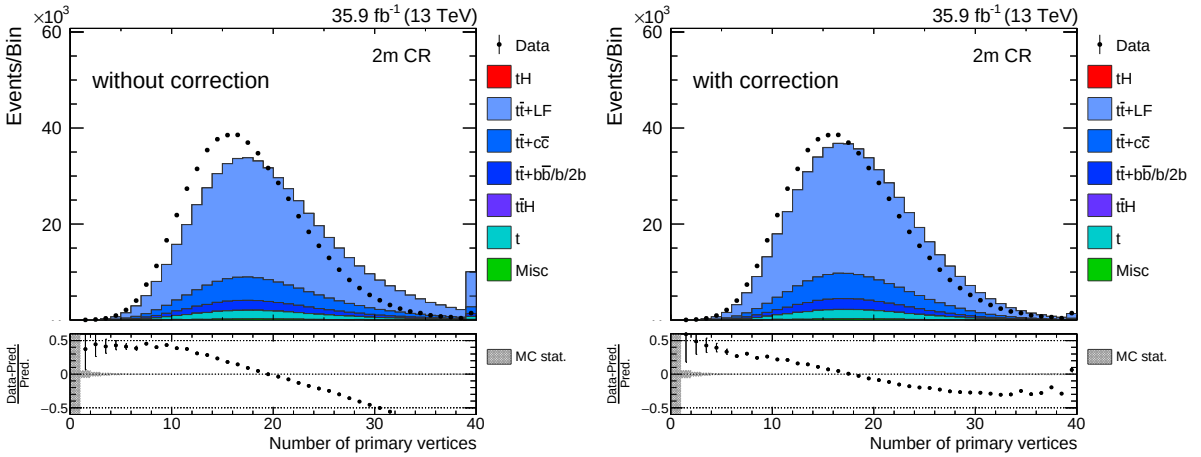

Figure 6.5.: Influence of the pileup correction: the distribution on the left shows the number of primary vertices without applying the scale factors and the right figure shows the distribution with the correction. Although the agreement is still not perfect, the correction leads to an improvement.

6.5. Uncertainty Treatment

All measurements are subject to systematic uncertainties. These have to be considered in an analysis to yield a reliable result. Both, experimental and theoretical uncertainties are considered. The first group contains all uncertainties related to the experimental setup and the event reconstruction and the uncertainties on the theory predictions are covered by the theoretical uncertainties. Some of these uncertainties only affect the overall event rate, which results in an uncertainty on the normalization of the templates, whereas other uncertainties change the normalization and the shape of the templates. All uncertainties considered in this analysis are listed below.

Experimental Uncertainties

- **Luminosity (rate):** The uncertainty assigned to the luminosity measurement is 2.5% [170].
- **Pileup (shape+rate):** The pileup corrections are subject to uncertainties as well [169]. These are estimated by shifting the minimum-bias cross section, with which the correction weights are obtained, by 4.6% [169].
- **b tagging (shape+rate):** The b-tagging scale factors have systematic uncertainties [149]. This uncertainty is split into several sources. Two parameters describe the uncertainty introduced by c jets. Furthermore, statistical and systematic uncertainties on the LF and HF scale factors are taken into account separately. The statistical uncertainty describes the fluctuations of the scale factors and the systematic uncertainty includes the impurities of the control samples.
- **Lepton efficiencies (shape+rate):** The systematic uncertainties of the inputs to the lepton scale factors are taken into account by modifying each of them, i.e., tracking, trigger, isolation and the ID flags [166–168]. These systematically shifted distributions account for the statistical and systematic uncertainties of the scale factors.
- **Statistical uncertainty of MC simulated samples (shape+rate):** Since the MC simulation samples used in any analysis are of finite size, the templates obtained from them are subject to statistical uncertainties. These uncertainties can be derived with the Barlow-Beeston method [80]. Here, the implementation of the Combine software package is used [74, 75], which introduces a nuisance

parameter for each bin and region for all processes combined. This reduces the computing effort compared to one nuisance parameter per process, bin and region.

- **Jet energy resolution (shape+rate):** The jet energy resolution (JER) is different in data and simulation. Therefore, the jets in the simulated samples are smeared to match the data. There are two methods for the smearing procedure: the scaling method and the stochastic smearing. The scaling method uses generator-level particles to obtain a scale factor. Hence, it is not possible for jets that cannot be matched to generator-level particles. The stochastic smearing does not depend on generator information as it scales the four-momenta of the jets with a stochastic scale factor. In order to combine the benefits of both models, a hybrid method is introduced. It uses the scaling method for all jets with matched generator-level particles and the stochastic method for the remaining jets. As the fluctuations in the transverse momenta of the jets causes events migrating in and out of the regions, the whole analysis is rerun with samples, for which the JER is systematically shifted upwards and downwards. The templates obtained from these samples are then used to estimate the systematic uncertainty in the final fit [171].

- **Jet energy scale (shape+rate):** The correction to the jet energy scale (JES) introduces a corresponding uncertainty [142–144]. This leads to variations in the transverse momentum of the jets. Hence, the same procedure as introduced for the JER is applied: the analysis is rerun for systematically varied JES corrections. This analysis features processes that produce several jets. Therefore, this uncertainty is of high importance. In order to describe the JES uncertainty more accurately, it is split into 26 different sources. These account for, e.g., the reference scales, extrapolation to high transverse momenta, jet flavor, η dependence, different methods for the scale determination and statistical uncertainties. A full list of these sources can be found in Ref. [172]. For each of them, systematically upwards and downwards varied samples are generated and the analysis is performed on them. This provides templates for each of the sources to be considered in the final fit.

- **Unclustered energy (shape+rate):** The momentum of the particles used to compute missing transverse momentum carry uncertainties as well [151]. These momenta are varied upwards and downwards and therefore the missing transverse momentum is varied as well to take these uncertainties into account. As the event selection involves the missing transverse momentum, the analysis is rerun on dedicated samples to account for migration effects.

Theoretical Uncertainties

- **Parton distribution functions and strong coupling constant (rate):** The uncertainties introduced by the parton distribution functions to the analysis are given as rate uncertainties. All uncertainties stem from theory calculations of the respective processes. For the tH signal processes, the uncertainty on the qg PDF is assumed to be 3.7% and 6.1% for the tHq and tHW processes, respectively. For the single top quark and tZq production, it is set to 4%. The gg PDF uncertainty is assumed to be 3.6% for the $t\bar{t}H$ process and 4.2% for the $t\bar{t}$ process. The uncertainty on the qq PDF for the $t\bar{t}Z$, $t\bar{t}W$, and Z+jets production is estimated to be 2.0%.

- **Flavor composition of additional quarks in top quark pair production (rate):** Since the flavor composition of the $t\bar{t}$+jets process is hard to model in theory, a conservative estimate of the uncertainty of 50% on the nominal cross section for each $t\bar{t}$+jets categories, see Section 5.2.1, is introduced. This uncertainty is constrained in the fit, especially by the fit of the dileptonic BR.

- **Transverse momentum of top quarks (shape+rate):** Since the transverse momentum of top quarks is not well modeled, a correction has been developed [173–177]. Instead of applying the correction, it is implemented as an uncertainty for $t\bar{t}$ events. Therefore, an additional weight is introduced to tune the distribution of the top quark transverse momentum. This results in a template for

the applied correction. Since the uncertainty is estimated by applying the correction versus not applying the correction, its impact on the signal strength is one-sided by construction.

- **Factorization and renormalization scales (shape+rate):** For the simulated samples a certain factorization and renormalization scale is set. This influences the properties of the simulated events. In order to estimate the influence of these scales, both are varied separately during the simulation to 50% and 200% of their nominal value [178]. The variation that yields the highest difference in the observables is chosen, this leads to a conservative estimate of the uncertainty. This is done for all processes separately, except for tW single top quark production, where a rate uncertainty of 3% is assumed as the variations are not available for these simulated samples.

6.6. Jet Assignment

In the analysis presented in this chapter, jet assignments under the tHq, tHW and $t\bar{t}$ hypothesis are performed with the methods introduced in Section 5.3. For the tHq and tHW jet assignments, the training is performed for the ITC scenario as the simulated samples are produced for this scenario. All other scenarios are obtained by applying event-specific weights. As a result of the uncertainty introduced by the reweighting, the overall performance is better for only one training on the ITC scenario compared to a specific training for each coupling point. The training is performed with events of the 3m region and the JA BDTs are applied to the events of the 3m and 4m region to obtain three event-specific jet assignments (one under each hypothesis). As the training is performed twice for each of the hypotheses, each of the following distributions exists twice: one distribution of category A events and one for category B events. Since the split is introduced for technical reasons only, the two distributions are almost identical, any deviations stem from statistical fluctuations. Therefore, only the combined distributions are shown. In contrast to the published analysis, not only one random background permutation is used in the training but all permutations are used. This results in a more robust assignment.

The distributions of correct and wrong assignments for the most important variables of the assignment can be found in Fig. 6.6 for the tHq jet assignment, in Fig. 6.7 for the tHW jet assignment and in Fig. 6.8 for the $t\bar{t}$ assignment. The modeling of these variables can only be checked after the application of the JA BDTs since the variables are calculated during the assignment process. The distribution of the variables in simulation and data can be found in Appendix A.3. As the tHW jet assignment requires one more jet compared to the tHq and $t\bar{t}$ hypotheses, the outputs for events with too few jets are set to a default value, which lies outside of the corresponding distribution. This is visible in the figures as a peak in the leftmost bin.

A complete list of the input variables to the JA BDTs can be found in Tables 5.2, 5.3, and 5.5. The variables ranked by their importance in the training for each of the hypotheses can be found in the Appendix A.4. The distributions of the JA BDT outputs for the test and training data sets can be found in Fig. 6.9. No sign of overtraining is found for any of the hypotheses. The distributions for the best JA BDT output on data and simulation can be found in Fig. 6.10. All variables show good agreement between data and simulation.

In order to evaluate the performance of the jet assignment, the average efficiency, as defined in Section 5.3.5, is computed. For the tHq jet assignment, the total efficiency using the JA BDT (χ^2 method) is 61.5% (51.3%). The total efficiency is 46.1% (24.5%) for the tHW jet assignment. For the $t\bar{t}$ jet assignment, the total efficiency is 51.1% (33.2%). The efficiencies for assigning individual final-state particles correctly can be found in Appendix A.5. The BDT-based approach is significantly better than the χ^2 based method under all three hypotheses.

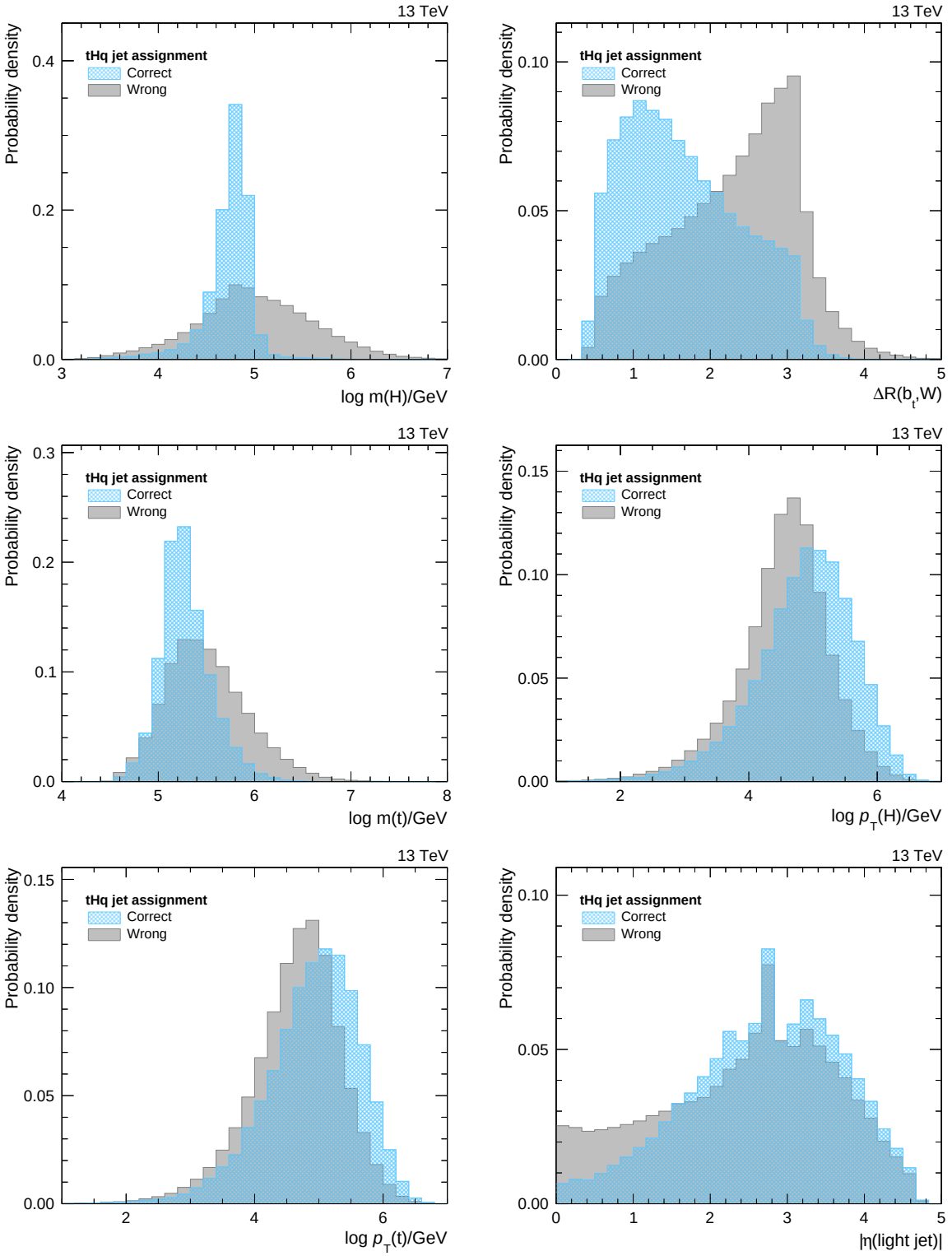

Figure 6.6.: Distributions for the most important variables of the tHq jet assignment: the distributions are shown for correct and wrong assignments for category A and B events combined. A description of the variables can be found in Table 5.2.

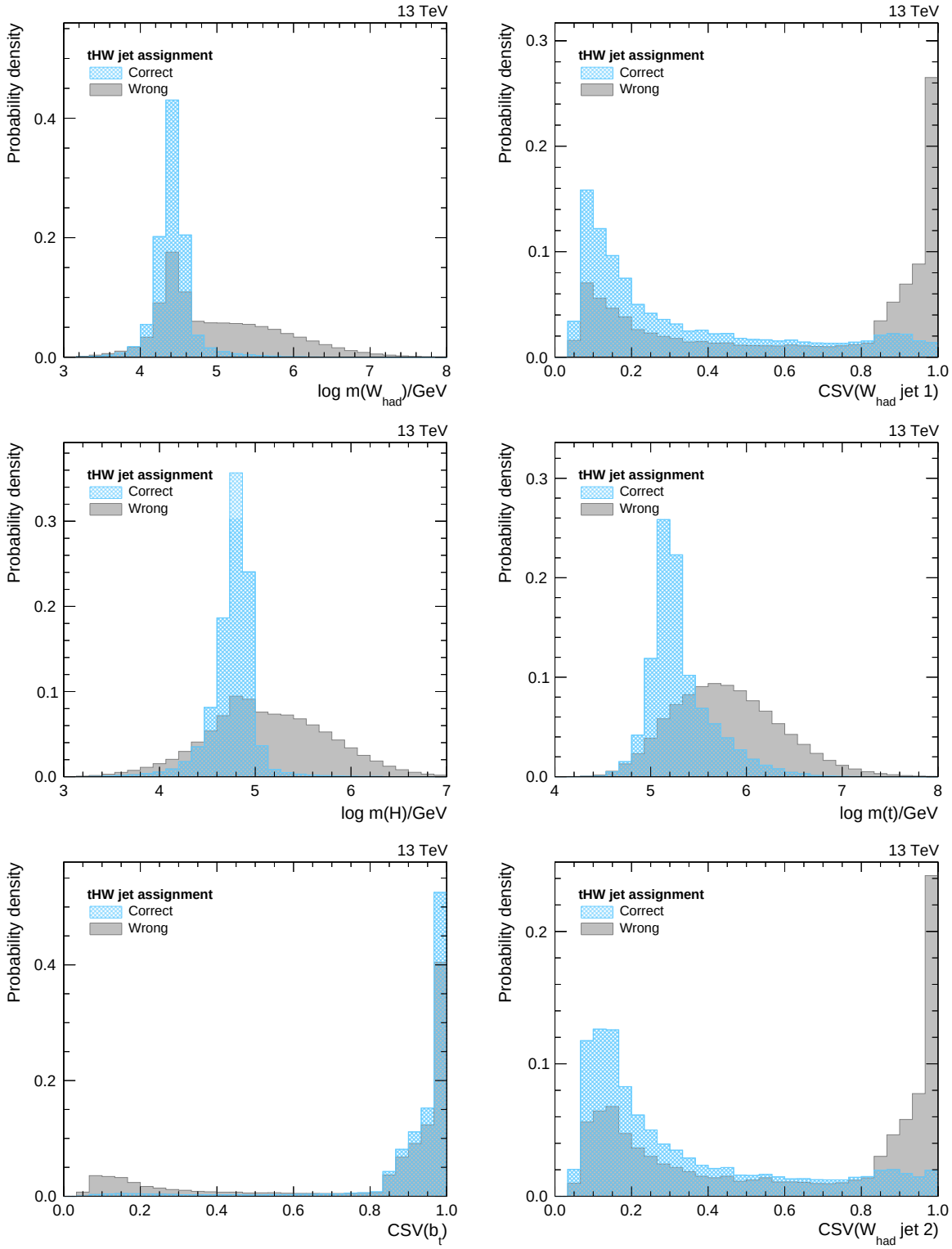

Figure 6.7.: Distributions for the most important variables of the tHW jet assignment: the distributions are shown for correct and wrong assignments for category A and B events combined. A description of the variables can be found in Table 5.3.

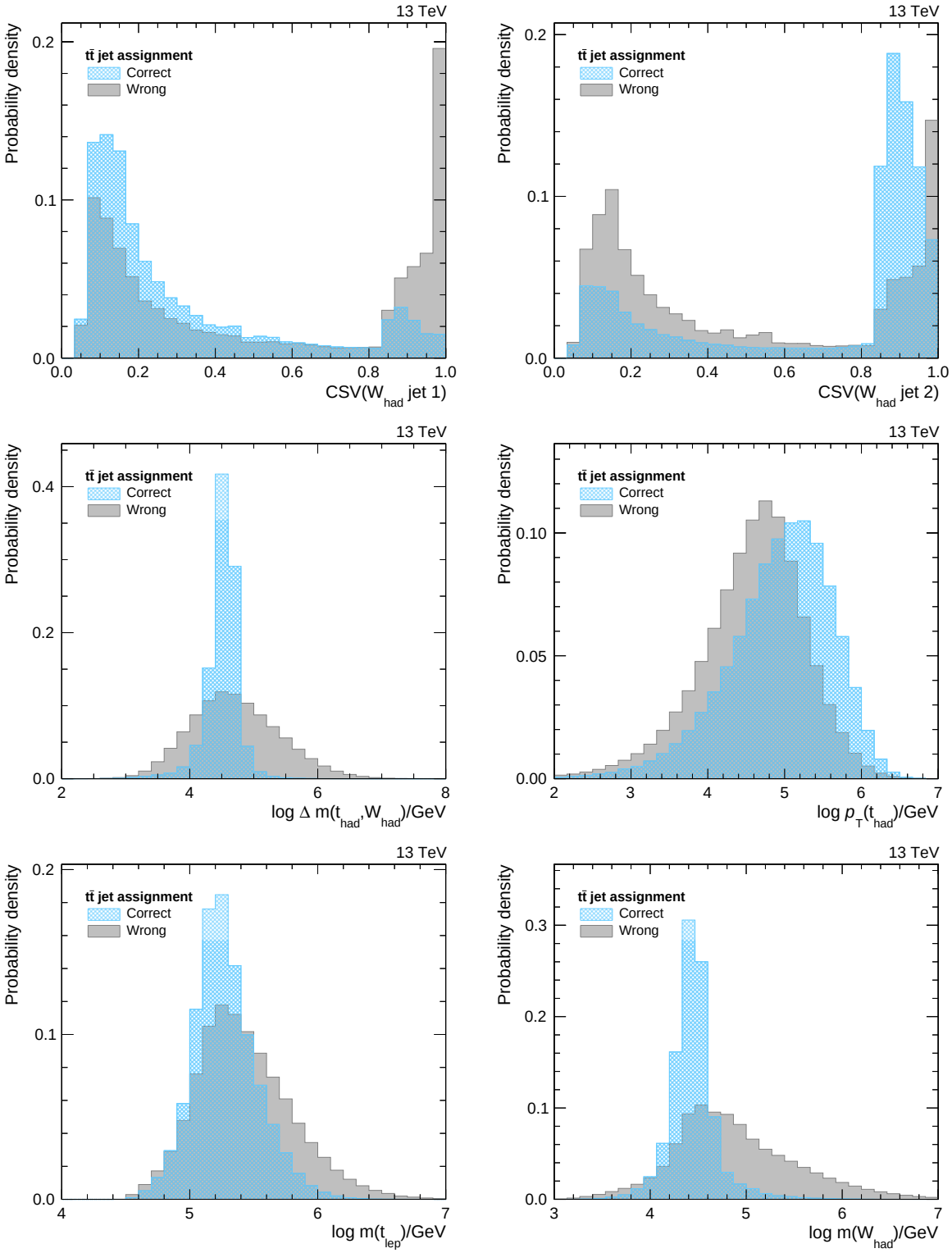

Figure 6.8.: Distributions for the most important variables of the tt̄ jet assignment: the distributions are shown for correct and wrong assignments for category A and B events combined. A description of the variables can be found in Table 5.5.

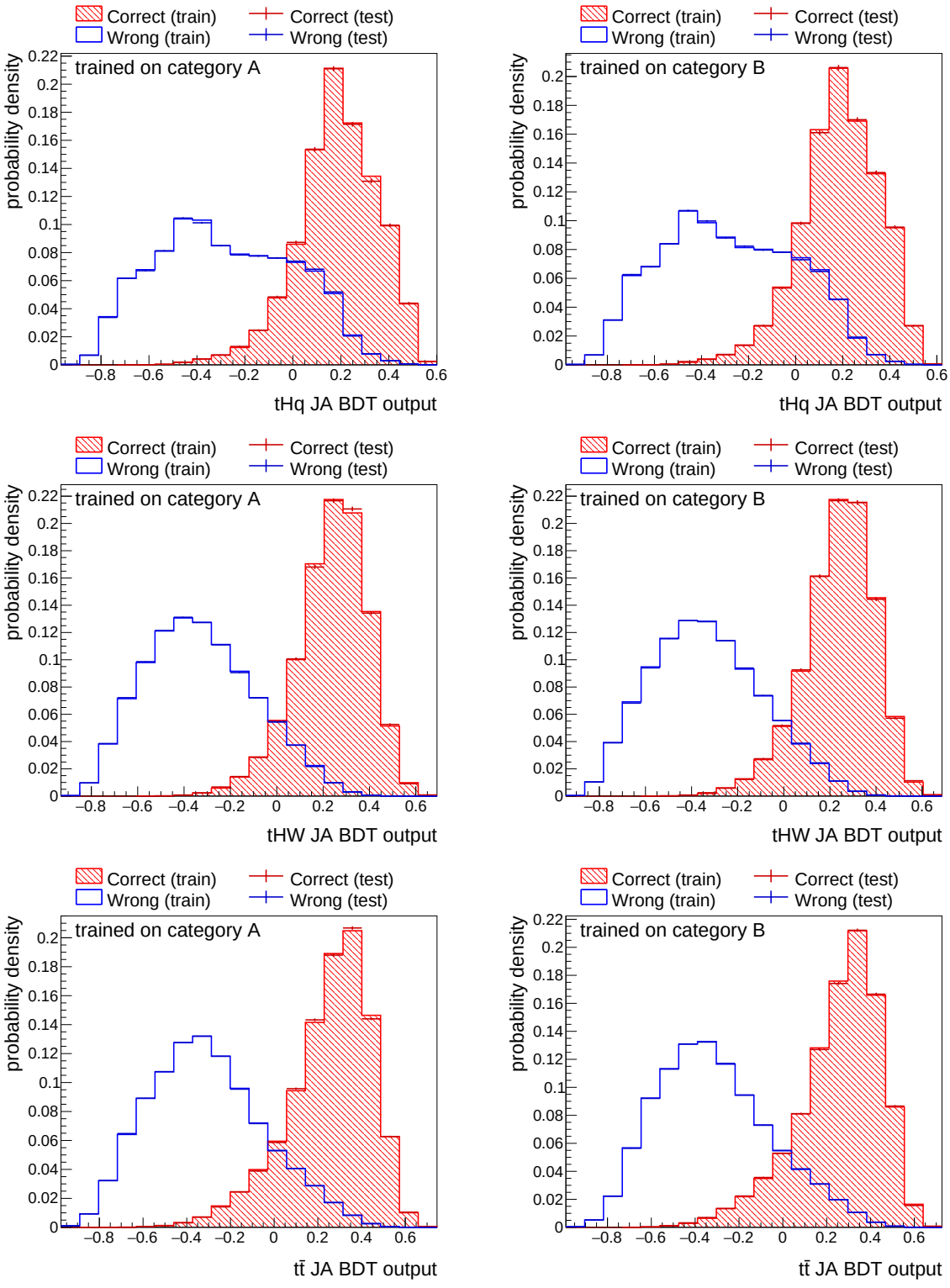

Figure 6.9.: Overtraining check for JA BDTs: the distributions on the left show the training on category A events and testing on category B. The right plots show the training on category B events and testing on category A. All distributions show good agreement between the outputs for the test and training data sets. No indication for overtraining is found. Since the wrong assignment contains the same events several times with different assignments, these events are not uncorrelated. Hence, a Kolmogorow-Smirnow test does not provide a good figure of merit, therefore it is not calculated here.

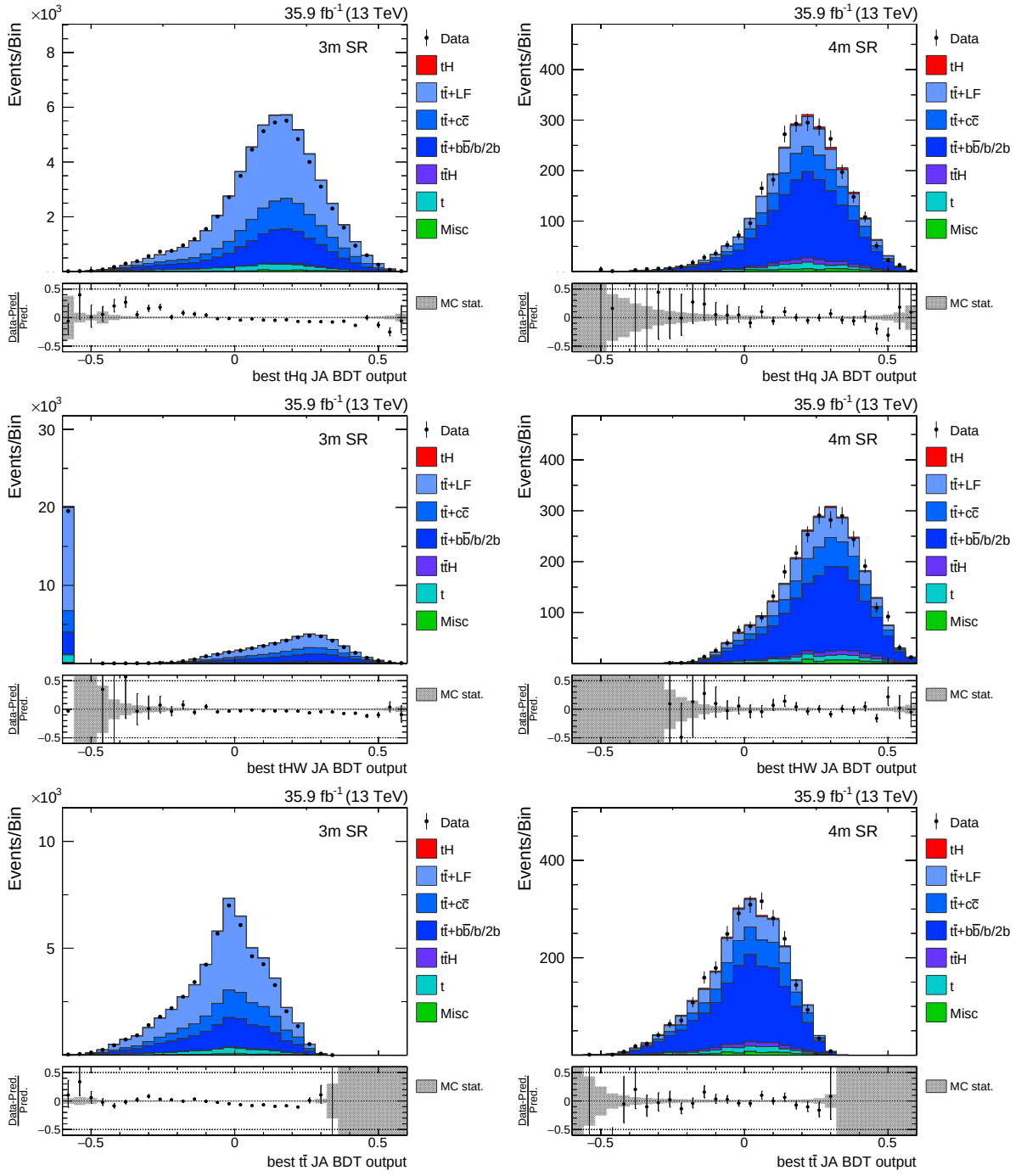

Figure 6.10.: Modeling of the best JA BDT output: the distributions show all events of the 3m region, i.e., category A and B events combined, on the left and the distributions for the 4m region on the right. The JA BDT outputs for the tHq, tHW and $t\bar{t}$ assignment are shown. The leftmost bin in the distribution of the best tHW JA BDT output contains events, for which the jet requirements of the assignment are not fulfilled. Since the 4m region has more jets, the tHW jet assignment is always able to produce a non-default output.

Table 6.4.: Settings used in the training of the SC BDT: these settings are used for the BDT separating tHq and tHW signal events from $t\bar{t}$ background events. The parameters are defined in Ref. [87].

Parameter	Value
`NTrees`	900
`MinNodeSize`	1%
`MaxDepth`	4
`nCuts`	15
`BoostType`	AdaBoost
`AdaBoostBeta`	0.3
`SeparationType`	GiniIndex

6.7. Classification of Events

There are two different types of classification BDTs trained within this analysis: signal classification (SC) BDTs and flavor classification (FC) BDTs. The SC BDTs are used to discriminate signal from background processes in the 3m and 4m signal regions in order to reach the highest sensitivity. The FC BDTs are trained and applied in the dileptonic BR to discriminate $t\bar{t}$+b$\bar{b}$/b/2b events from $t\bar{t}$+lf events in order to constrain the uncertainties on the $t\bar{t}$+jets production.

6.7.1. Signal Classification

The signal processes studied in this analysis are very rare compared to the background processes. In order to obtain the best possible results, an optimal discrimination between signal and background events is desirable. This can be achieved by combining the discrimination power of several variables employing a multivariate method. Here, BDTs are utilized. The SC BDTs are trained to separate tHq and tHW signal events from $t\bar{t}$ background events. This setup is chosen since the $t\bar{t}$ production is the main background process and tHq and tHW are the dominant signal processes within this analysis. Similar to the jet assignment, the ITC scenario is chosen in the training for the signal events. The classification BDT output is then reweighted to obtain the distributions for all other coupling scenarios.

The training is performed in the 3m region and the BDT is then applied to the simulation and data samples of the 3m and 4m regions as this has been studied to result in the best overall separation. Similar to the JA BDTs, the samples are randomly split in categories A and B, where each of the categories contains about half of the events. The BDT trained on category B events is applied to category A events and vice versa. Hence, all events can be used in the final fit. The sets are split into 80% for the training of the SC BDT and 20% for the testing.

The SC BDT uses global variables, as well as variables obtained by the jet assignment, as described in Section 5.3. The variables obtained by the jet assignment provide dedicated information about final-state particles, such as the top quark or the Higgs boson. In principle, the properties of the reconstructed Higgs boson, e.g., the mass, is closer to the real Higgs boson properties for signal events than for background events. The BDT configuration can be found in Table 6.4 and a full list of the variables can be found in Table 6.5.

The distributions of the most important input variables for signal and background events can be found in Fig. 6.11. As these are all jet assignment variables, their modeling is shown in Appendix A.3. The distributions for category A and B look alike, therefore only the category A distributions are shown. The classification BDT output for test and training events for the training on both categories are shown in Fig. 6.12. The classification BDT output in the 3m and 4m region for data and simulation can be found in Fig. 6.13.

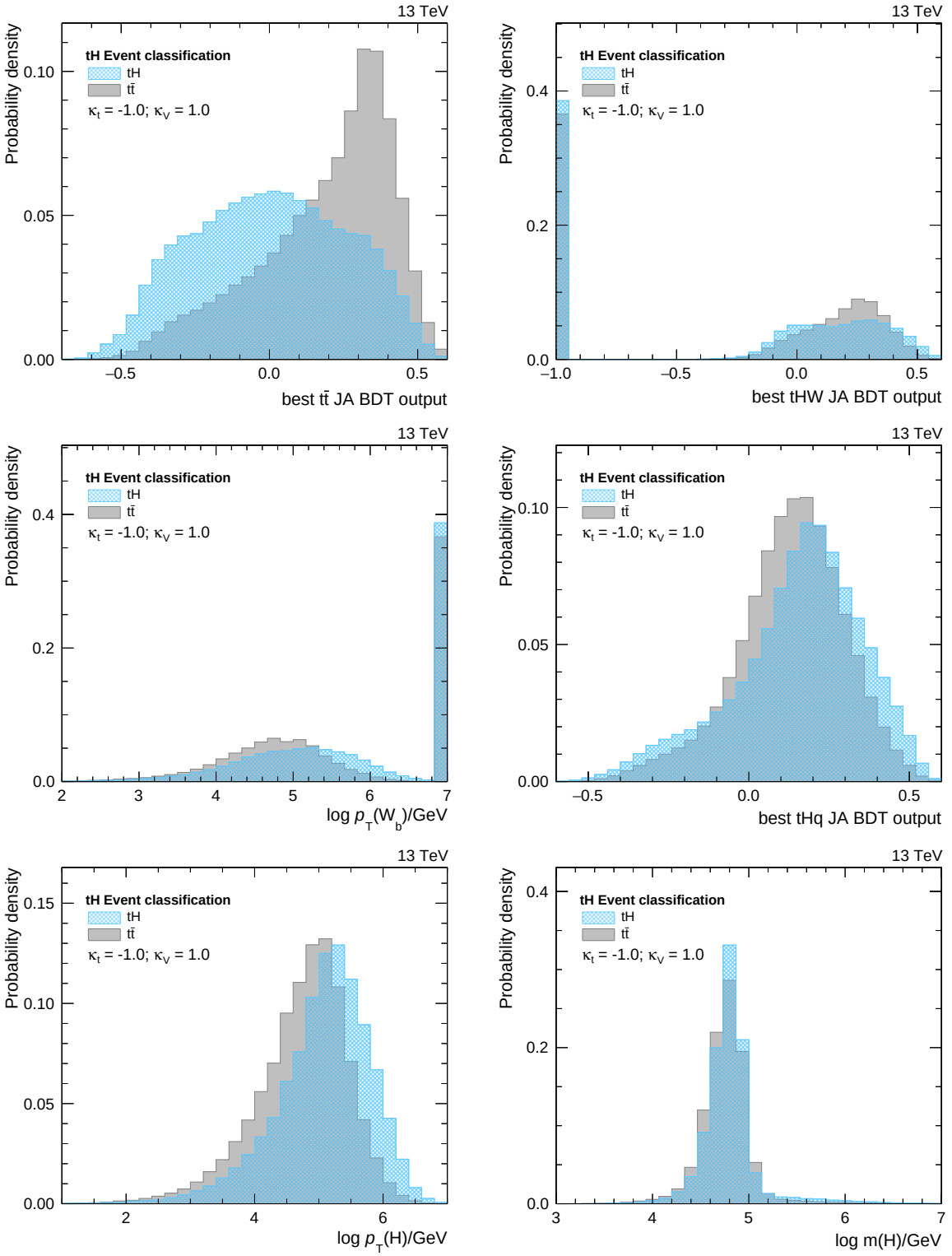

Figure 6.11.: Distributions for the most important variables of the classification BDT: the distributions are shown for signal and background processes of category A events. The distributions for categories A and B look alike, therefore only the category A distributions are shown.

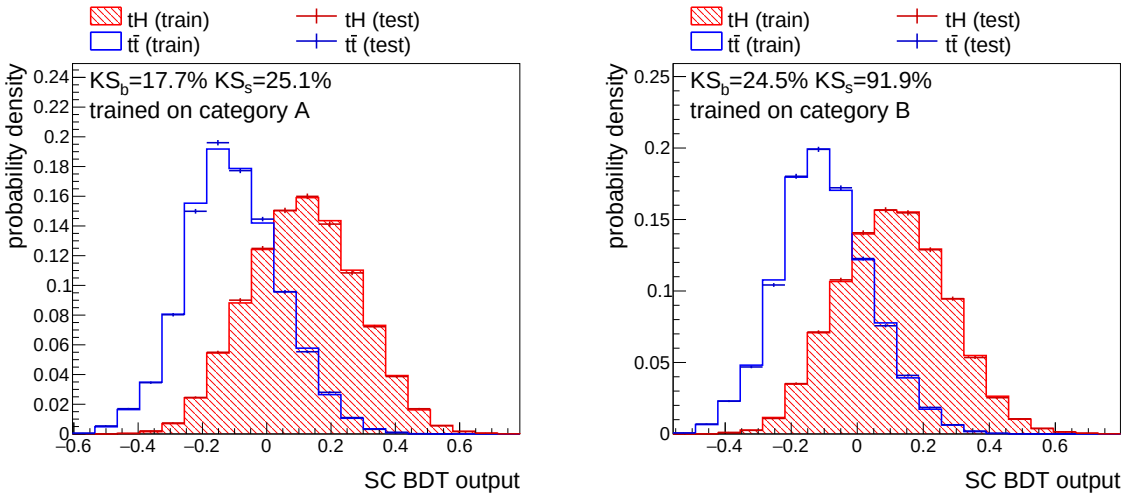

Figure 6.12.: Overtraining check for SC BDTs: the distribution on the left shows the training on category A events and testing on category B. The right plot shows the training on category B events and testing on category A. Both distributions show good agreement between the outputs for the test and training data sets. No indication for overtraining is found.

As the best JA BDTs are very important to the signal classification, further studies have been made to enclose more information on these outputs [180]. For every event there is a huge number of combinations leading to many different JA BDT outputs for each hypothesis, of which only the best output is stored. Nonetheless, the distribution of the remaining outputs contains further information, which can be used to discriminate between signal and background events. One option is to evaluate the second best JA BDT output as well: in case of evaluating the tHq assignment on a tHq event, the best assignment should be clearly better than the second best, i.e., the best and second best JA BDT should differ significantly. For non-tHq events, this difference should be smaller as all assignments are incorrect since there cannot be a correct assignment under the tHq hypothesis for, e.g., a $t\bar{t}$ event.

Furthermore, the shape of all JA BDT outputs under a certain hypothesis for an event can result in additional discrimination power. The shape can be described by, e.g., the asymmetry of the JA BDT output as defined in Ref. [180]. Nonetheless, the study shows that these additional variables describing the JA BDTs do not improve the expected limits. Hence, these variables are not implemented.

Another study is carried out in order to compare the SC BDT to a binary classification DNN [181]. The DNN also includes some further variables and is optimized to obtain the best separation power. This results in slightly better expected limits for some of the coupling ratios and in worse limits for other points. On average no improvement is found. Therefore, the BDT is kept as classification method.

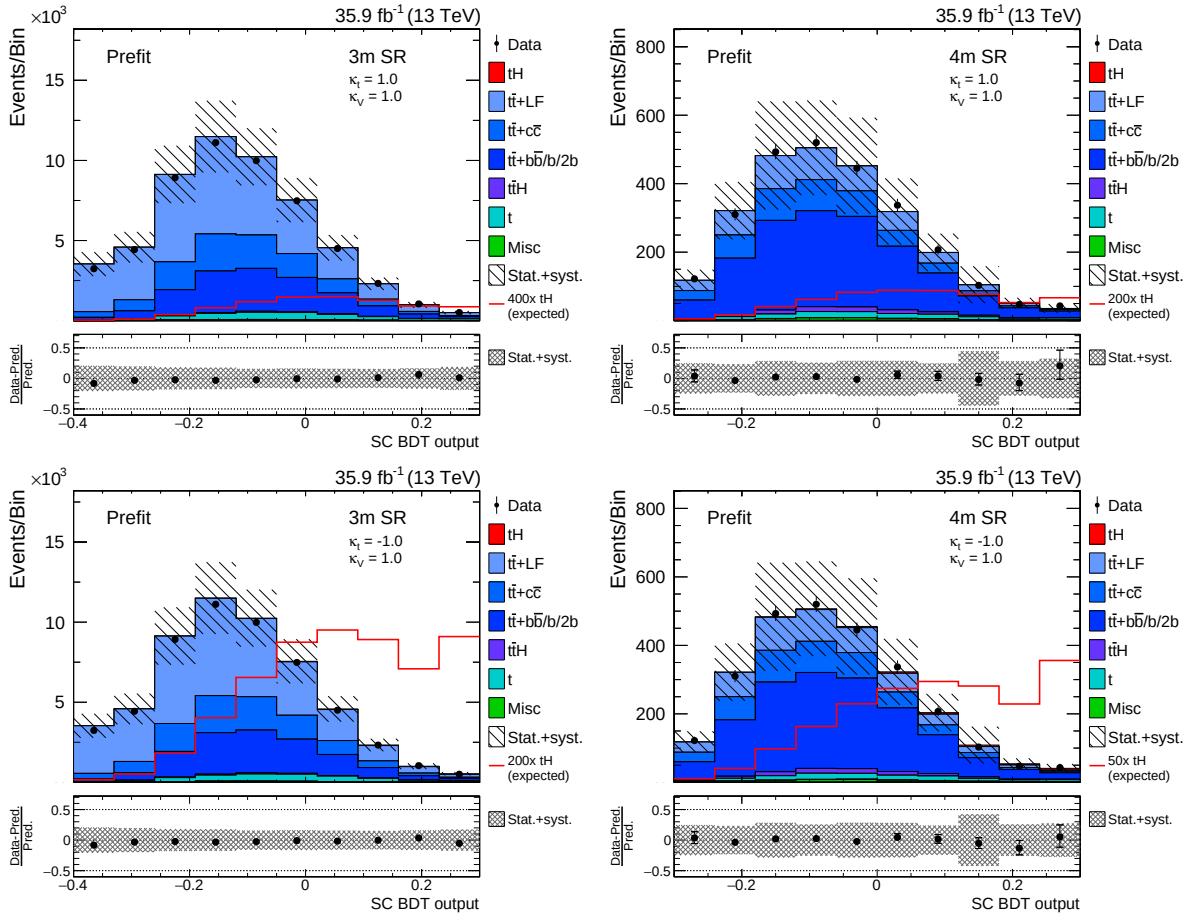

Figure 6.13.: Prefit output distributions of the SC BDTs in the 3m and 4m signal regions: the distribution on the left shows the SC BDT output in the 3m region, the 4m region is shown on the right. The distributions in the top row show the SM case and the bottom row shows the ITC case. The distributions have the same range and number of bins as used in the final fit. The leftmost bin contains all events with a BDT output lower than its upper edge and the rightmost bin contains all events with a BDT output larger than its lower edge. The signal shape is indicated as red line.

Table 6.5.: Input variables used in the SC BDT: the variables are grouped into global and assignment specific variables. Within their respective group, the variables are ranked by their importance as defined in Section 3.5.

Variable	Importance	Description
		global variables
aplanarity	0.05	aplanarity of the event [179]
log m3/GeV	0.04	invariant mass of the three hardest jets in the event
q(ℓ)	0.02	electric charge of the lepton
		tHq jet assignment variables
best tHq JA BDT output	0.07	best tHq JA BDT output
log p_T(H)/GeV	0.07	transverse momentum of the reconstructed Higgs boson
log m(H)/GeV	0.06	invariant mass of the reconstructed Higgs boson
$\lvert\eta$**(light jet)**$\rvert$	0.04	absolute pseudorapidity of the jet assigned to the light quark
CSV($b_{H,1}$)	0.04	output of the b-tagging algorithm for the first jet assigned to the Higgs boson
cos θ(b_t, ℓ)	0.03	cosine of the angle between the jet assigned to the bottom quark and the charged lepton
CSV($b_{H,2}$)	0.03	output of the b-tagging algorithm for the second jet assigned to the Higgs boson
$\lvert\eta$**(t)** $-\ \eta$**(H)**$\rvert$	0.03	absolute difference of the pseudorapidities of the reconstructed top quark and the reconstructed Higgs boson
log p_T(light jet)/GeV	0.03	transverse momentum of the jet assigned to the light quark
		tHW jet assignment variables
best tHW JA BDT output	0.08	best tHW best JA BDT output
log m(W_b)/GeV	0.07	invariant mass of the reconstructed additional W boson
log p_T(W_b)/GeV	0.05	transverse momentum of the reconstructed additional W boson
		$t\bar{t}$ jet assignment variables
best $t\bar{t}$ JA BDT output	0.09	best $t\bar{t}$ best JA BDT output
log m(t_{had})/GeV	0.04	invariant mass of the reconstructed hadronically decaying top quark
CSV(W_{had} jet 1)	0.04	output of the b-tagging algorithm for the first jet assigned to the hadronically decaying W boson
ΔR (W_{had})	0.06	ΔR between the two jets assigned to the W boson of the hadronically decaying top quark
CSV(W_{had} jet 2)	0.04	output of the b-tagging algorithm for the second jet assigned to the hadronically decaying W boson

Table 6.6.: Settings used in the training of the FC BDT: these settings are used for the BDT separating $t\bar{t}$+lf events from $t\bar{t}$+$b\bar{b}$/b/2b events. The parameters are defined in Ref. [87].

Parameter	Value
`NTrees`	600
`MinNodeSize`	2%
`MaxDepth`	2
`nCuts`	16
`BoostType`	AdaBoost
`AdaBoostBeta`	0.3
`SeparationType`	GiniIndex

6.7.2. Flavor Classification

The training of the flavor classification (FC) BDTs is performed in the dileptonic BR. Similar to the signal classification, the training is performed twice, i.e., on category A events and on category B events. The flavor classification does not aim at classifying signal events but on discriminating between different $t\bar{t}$+jets contributions. The goal is to obtain different shapes for the contributions in order to constrain the uncertainty on these contributions from data by adding the dileptonic BR in the final fit. Therefore, a training is performed to discriminate $t\bar{t}$+lf events as signal, where lf stands for up, down and strange quarks and gluons, from $t\bar{t}$+$b\bar{b}$/b/2b events as background. An in depth study of the flavor classification in the context of the tH analysis can be found in Ref. [165].

The training is performed with the BDT setup as shown in Table 6.6. The input variables for the FC BDT are listed in Table 6.7. The shapes of the input variables for signal and background events are shown in Fig. 6.14 and their modeling is shown in Appendix A.6. The FC BDT output on test and training samples for both categories is shown in Fig. 6.15. The output of the FC BDT and its modeling is shown in Fig. 6.16.

The charm vs. light and charm vs. bottom discriminators are introduced to handle the $t\bar{t}$+$c\bar{c}$ contribution, which is not considered in the training. The use of these input variables results in a shape of this contribution that lies between the light flavor and the b jet contribution. Hence, the output of the FC BDT is able to provide different shapes for the different $t\bar{t}$+jets contributions. These are used in the final fit to constrain the respective uncertainties.

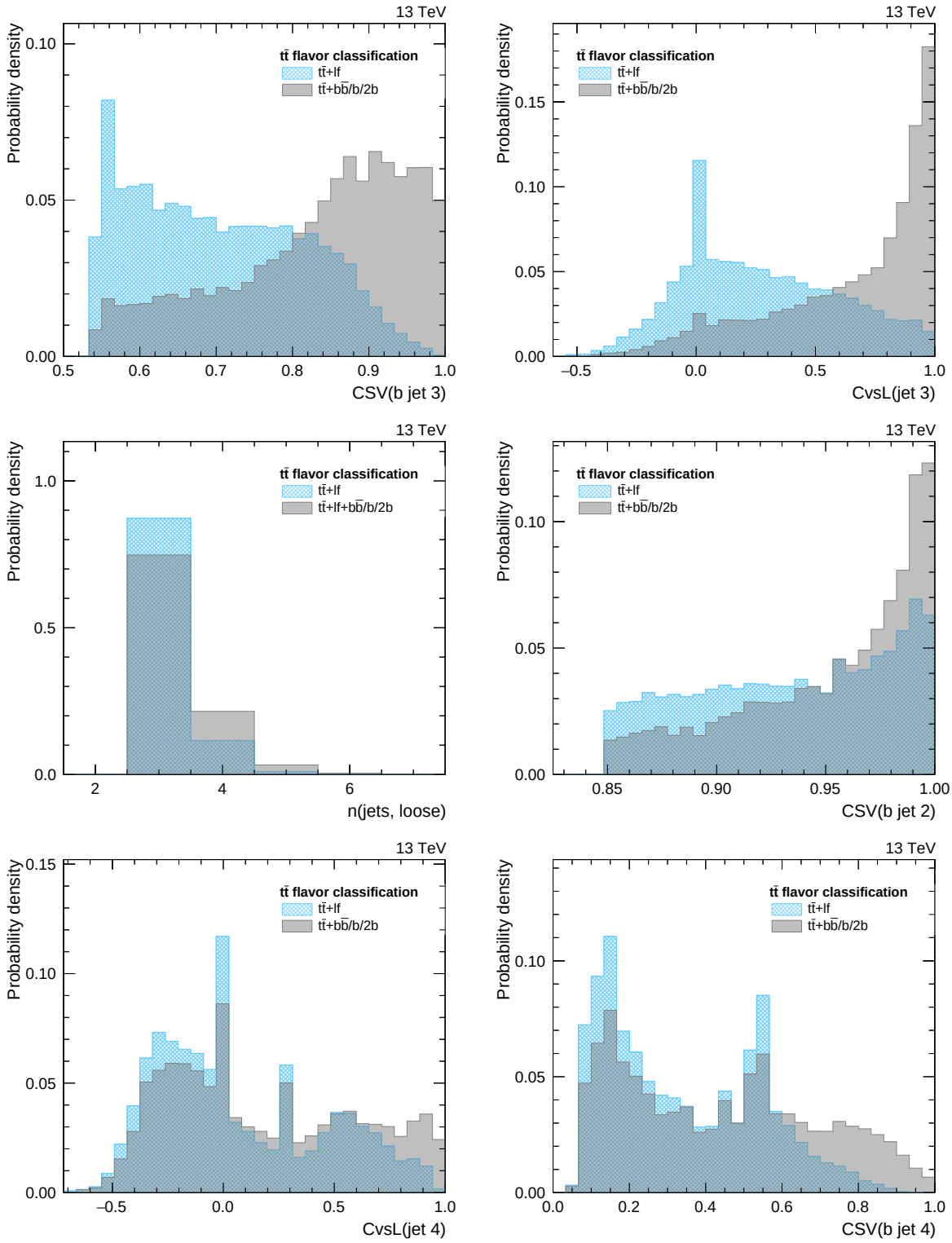

Figure 6.14.: Distributions for the most important variables of the FC BDT: the distributions are shown for $t\bar{t}$+lf and $t\bar{t}$+$b\bar{b}$/b/2b processes of category A events. The distributions for categories A and B look alike, therefore only the category A distributions are shown. The peaks in the CSV output distributions originate from requirements on the CSV output in the even selection.

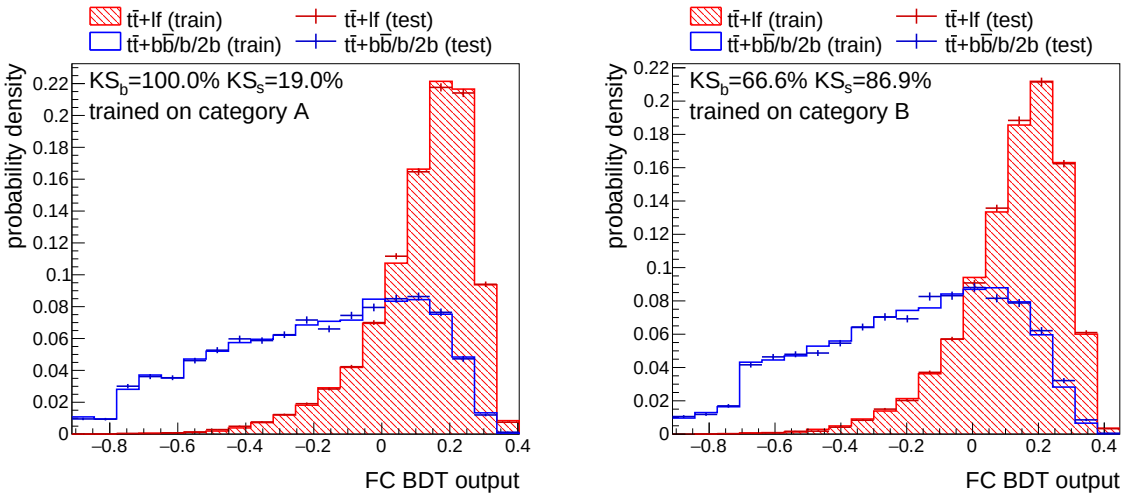

Figure 6.15.: Overtraining check for FC BDTs: the distribution on the left shows the training on category A events and testing on category B. The right plot shows the training on category B events and testing on category A. Both distributions show good agreement between the outputs for the test and training data sets. No indication for overtraining is found.

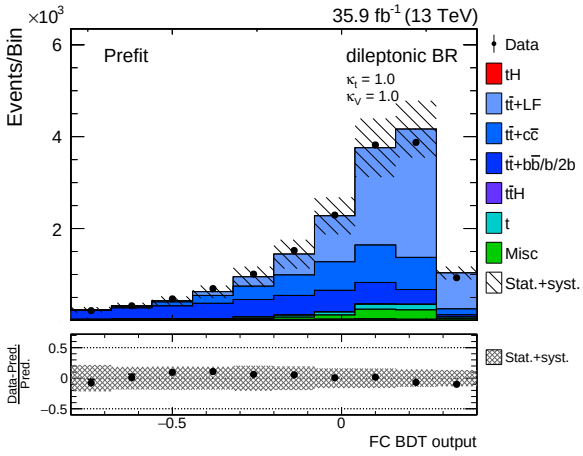

Figure 6.16.: Prefit distribution of the FC BDT in the dileptonic BR: the figure on the left shows the FC BDT output in the dileptonic BR on simulation and data for the SM case. As a result of the low signal contribution in the dileptonic BR, the distributions for other coupling scenarios look alike. The distribution has the same range and number of bins as used in the final fit. The leftmost bin contains all events with a BDT output lower than its upper edge and the rightmost bin contains all event with a BDT output larger than its lower edge.

Table 6.7.: Input variables used in the FC BDT: The variables are ranked by their average importance in the trainings on category A and B events. The definition of the importance can be found in Section 3.5.

Variable	Importance	Description
CSV(b jet 3)	0.20	output of the b tagging algorithm for the jet with the third highest CSV value
CvsL(jet 3)	0.18	output of the charm vs. light tagging algorithm for the jet with the third highest transverse momentum
n(jets, tight)	0.15	number of jets passing the tight b tag working point
CSV(b jet 2)	0.13	output of the b tagging algorithm for the jet with the second highest CSV value
CvsL(jet 4)	0.18	output of the charm vs. light tagging algorithm for the jet with the fourth highest transverse momentum
CSV(b jet 4)	0.10	output of the b tagging algorithm for the jet with the fourth highest CSV value
CvsB(jet 3)	0.08	output of the charm vs. bottom tagging algorithm for the jet with the third highest transverse momentum
n(jets, loose)	0.04	number of jets passing the loose b tag working point

Table 6.8.: Limits on the tH signal strength: upper limits on the combined signal strength of the tH processes are set for the SM and ITC scenarios. The signal events are reweighted to represent the respective scenario.

Scenario	μ_{obs}	μ_{exp}	±1 std. dev.	±2 std. dev.
SM	53	35	[23, 53]	[17, 77]
ITC	2.9	2.0	[1.4, 3.0]	[1.0, 4.5]

6.8. Results of the Higgs Boson to $b\bar{b}$ Decay Channel

This analysis features three different results: limits on the signal strength of tH production in the SM and ITC scenarios, limits on cross section times branching ratio for different κ_t/κ_V coupling ratios and limits on the CP-mixing angle α. For all of them the outputs of the SC BDT in the 3m and 4m signal region and the FC BDT in the dileptonic BR are fitted simultaneously.

For the limits on the signal strength $\mu = \sigma_{obs}/\sigma_{pred}$ of the tH production, the tHq and tHW processes are considered as signal, whereas all other processes, including $t\bar{t}H$, are considered as background. The shapes of the BDTs are reweighted to represent the SM and ITC scenarios, respectively. The signal strength of the tH processes is fitted with one common signal strength modifier. The upper observed (expected) limit on SM-like tH production at 95% C.L. is 53 (35) and for the ITC scenario 2.9 (2.0). Both results are compatible within the uncertainties of the respective expectations as listed in Table 6.8.

The upper limits on the cross section times branching ratio for different coupling ratios κ_t/κ_V use the tHq and tHW processes as signal, too. In addition, the $t\bar{t}H$ process is treated as signal, since it adds sensitivity to κ_t. The BDT output shape is reweighted to obtain shapes representing the specific coupling ratios for all three signal processes. The limits are shown in Fig. 6.17. The results for the individual scenarios are listed in Appendix A.7. This analysis constrains the κ_t to $[-2.5, 4.0]$ for an SM-like coupling between the Higgs boson and the W boson, i.e., $\kappa_V = 1$. The postfit distributions of the BDT outputs for the SM and ITC case are shown in Fig. 6.18. The impacts of the systematic uncertainties for the SM and the ITC scenario can be found in Appendix A.8.

Again, the tHq, tHW and $t\bar{t}H$ processes are treated as signal in the study of the CP-mixing angle α as all of them are sensitive to modifications of this angle. The BDT output shape is reweighted to obtain

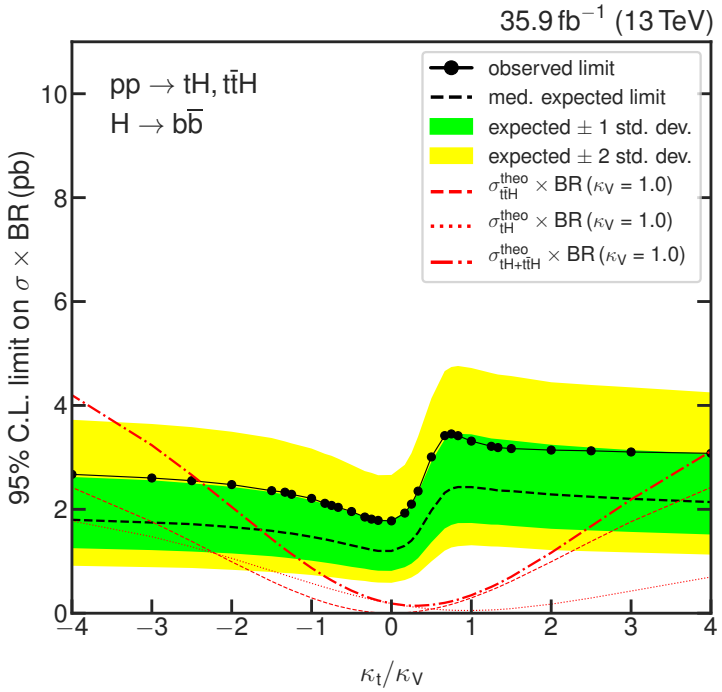

Figure 6.17.: Limits on the cross section times branching ratio for different κ_t/κ_V-ratios: the limits and the corresponding theory predictions are shown for different κ_t/κ_V-ratios. Scenarios, where the observed limit is below the theory prediction, are excluded, i.e., values below $\kappa_t = -2.5$ and above $\kappa_t = 4$ assuming $\kappa_V = 1$.

shapes representing the specific α values for all three signal processes. Here, the kinematic variations of the tt̄H process are taken into account for the first time. Previous analyses [98, 159] only considered variations in the cross section but not in the kinematic properties. The results are shown in Fig. 6.19. Tables listing the results for the individual scenarios can be found in Appendix A.7. All points are compatible with the expectation; no scenario can be excluded.

A comparison of the results for the coupling ratio as presented in this thesis and the results obtained without the improvements implemented in this thesis, namely the b-jet energy regression, the two-folded training and the additional simulated signal events, is shown in Fig. 6.20. The improved analysis results in better limits for all studied ratios. The average improvement over the previous method is 31% based on expected limits assuming the same amount of data. About 10% of the improvement stem from the improved energy resolution of b jets. The increased number of events in the training of the BDTs yields the remaining improvement.

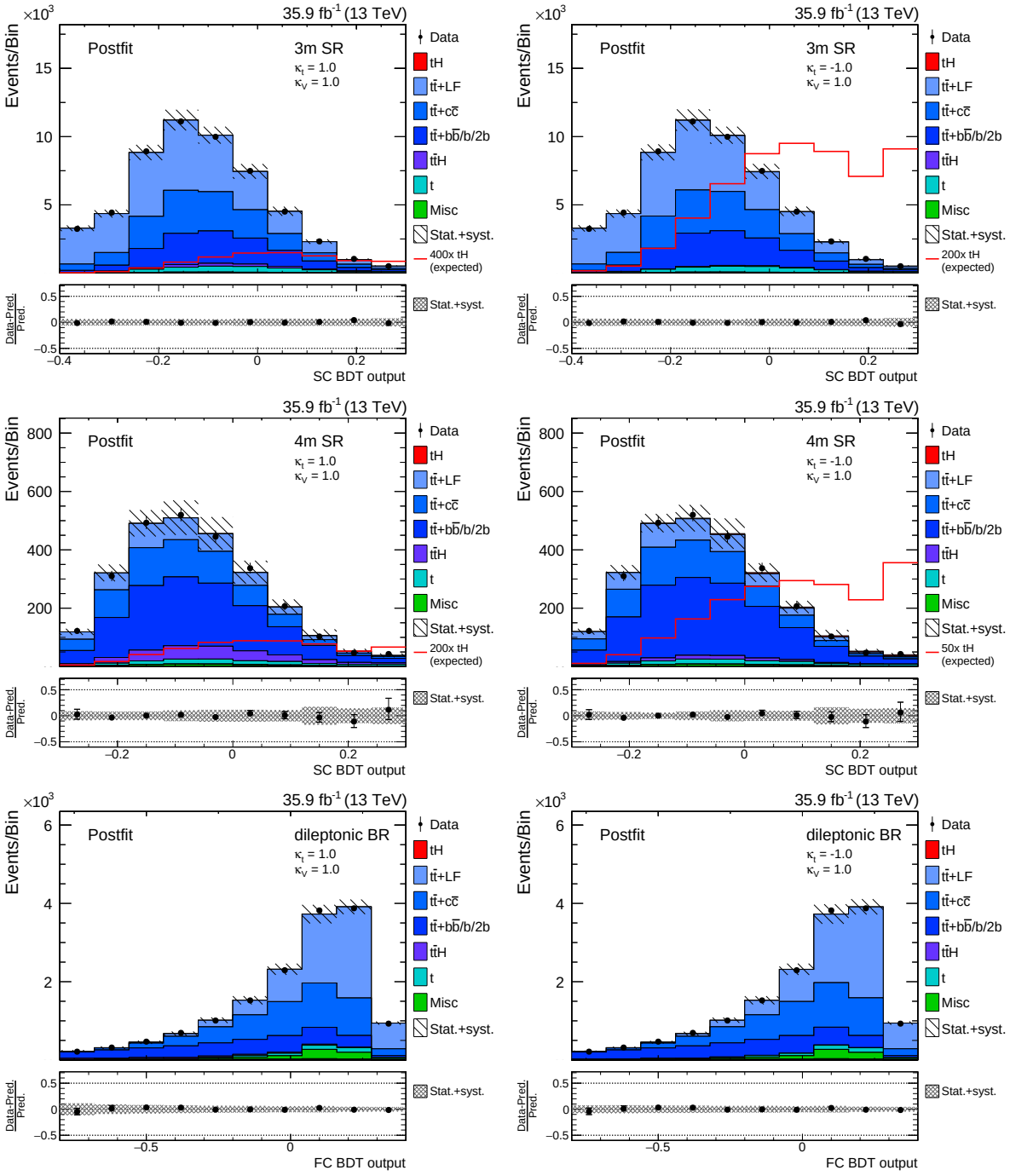

Figure 6.18.: Postfit distributions of the BDT outputs: the figures on the left show the SM case and the figures on the right show the ITC case. The distributions have the same range and number of bins as used in the final fit. The leftmost bin contains all events with a BDT output lower than its upper edge and the rightmost bin contains all event with a BDT output larger than its lower edge.

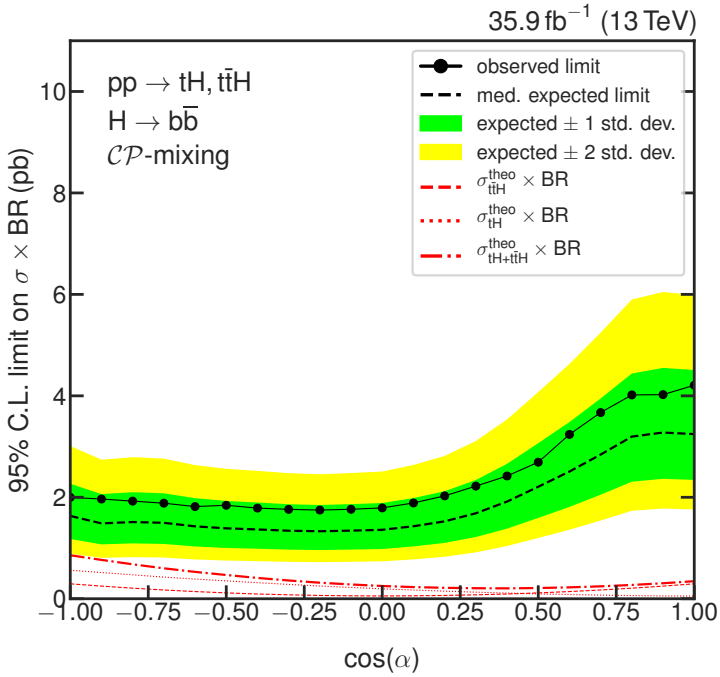

Figure 6.19.: Limits on the cross section times branching ratio for different cos(α) values: the limits and the corresponding theory predictions are shown for different cos(α) values. The SM scenario corresponds to $\cos(\alpha) = 1$ and the ITC corresponds to $\cos(\alpha) = -1$. None of the scenarios can be excluded. The limits as shown here are calculated using a dedicated $t\bar{t}H$ sample that contains event weights to obtain the different CP-scenarios. This simulated sample has fewer events than the sample used in the rest of the analysis. Therefore, the limits for the SM and ITC scenarios are slightly worse than in the κ_t/κ_V scan. However, both limits for the SM case and for the ITC case are compatible within the uncertainties.

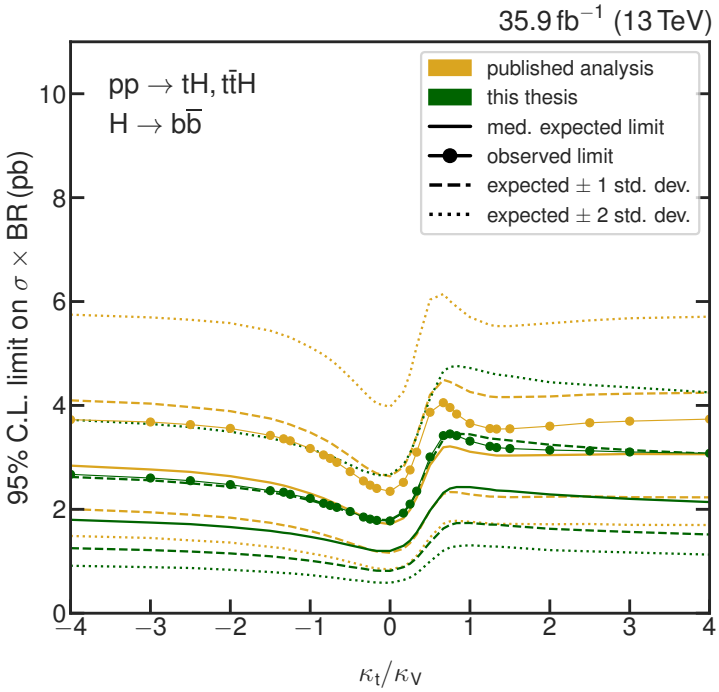

Figure 6.20.: Comparison of limits on the cross section times branching ratio for different κ_t/κ_V-ratios: a comparison of the limits computed in the published analysis [2] and the improved limits of this thesis. The observed, as well as the expected limits, are significantly improved.

6.9. Combined Measurement of Different Higgs Boson Decay Channels

In addition to the Higgs boson decay into a pair of bottom quarks [1] as presented in the previous chapter, the single top quark production in association with a Higgs boson is studied in the multilepton decay channel [160]. Here, the analysis with the results as given in the paper is used as input, i.e., the improvements, as discussed above, are not implemented as the improvements were not available at the time of the publication. In order to derive the most sensitive limits on the Higgs boson couplings and the SM-like tH production, the results of the $b\bar{b}$ and multilepton decay channels are combined. Finally, the analysis results of the Higgs boson decay into a pair of photons [161] are reinterpreted and also used in the combined result. The analyses are performed on the data set taken by the CMS experiment in 2016 with an integrated luminosity of 35.9 fb^{-1}. The combined result is published by the CMS collaboration in Ref. [2].

6.9.1. $b\bar{b}$ Channel

The Higgs boson to $b\bar{b}$ decay channel is already described above. The analysis strategy and workflow for the published analysis are similar to those described above, except for the two-fold BDT training and the b-jet energy regression. Hence, only the postfit SC and FC BDT outputs are shown in Fig. 6.21 as obtained from the combined fit.

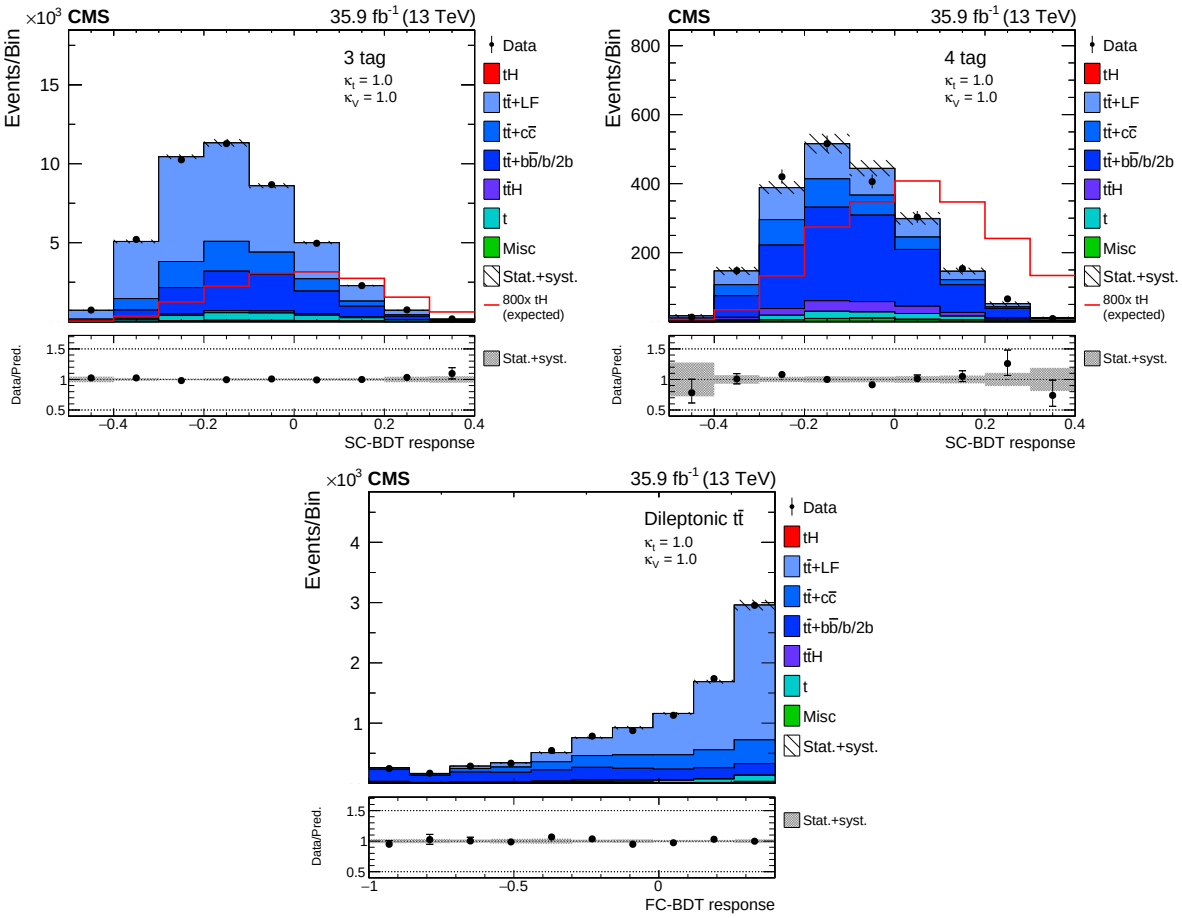

Figure 6.21.: Postfit distributions of the SC and FC BDTs: the distributions are shown for the two signal regions at the top and the dileptonic BR at the bottom. The shape of the SM-like tH signal, enhanced by a factor of 800, is shown by the red line in the distributions for the signal regions. Taken from Ref. [2].

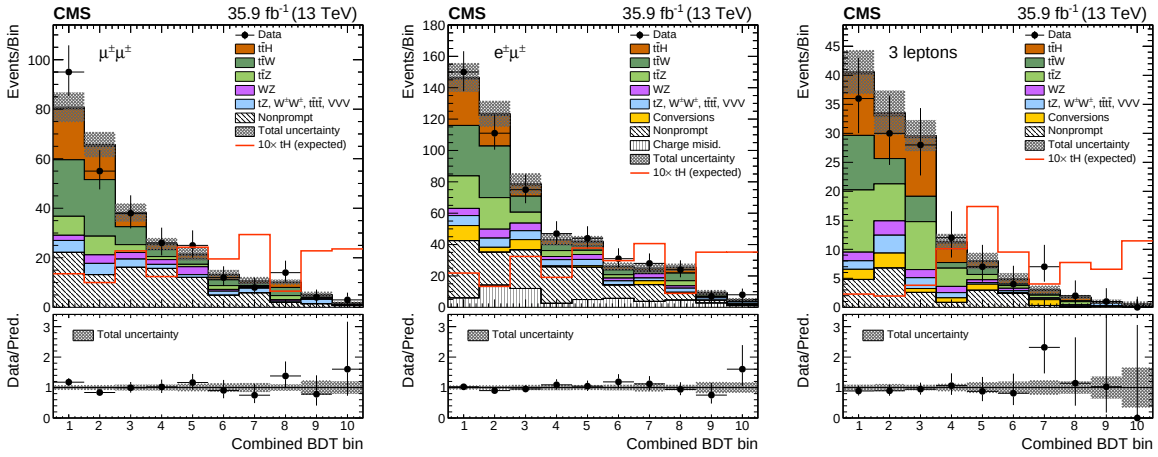

Figure 6.22.: Postfit distributions of the combined BDT bins: the distributions are shown for all three signal regions. The shape of the SM-like tH signal, enhanced by a factor of ten, is shown by the red line. Taken from Ref. [2].

6.9.2. Multilepton Channel

The multilepton channels target events, where the top quark decays leptonically and the Higgs boson decays either into τ leptons or vector bosons. Since this decay channel is not directly part of this thesis, only a brief summary is given here. For more details, see Refs. [2, 160].

The events with three charged leptons or two leptons with identical charge are of special interest as they feature a comparatively low background contribution. The main background processes to this analysis are $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}H$ production as these can yield the required prompt leptons. The $t\bar{t}$+jets production with non-prompt leptons is another important background process.

The analysis utilizes three orthogonal signal regions, all of which require at least one b-tagged jet and at least one untagged jet. The three lepton region ($\ell\ell\ell$, $\ell = e, \mu$) requires exactly three leptons fulfilling the respective tight lepton ID. The two same sign channels, $\mu^{\pm}\mu^{\pm}$ and $e^{\pm}\mu^{\pm}$ require exactly two tight muons or a tight electron and a tight muon, respectively.

About 1% of the events passing the event selection stem from tH production, when assuming SM-like tH production. Two BDTs are trained to discriminate tHq signal events from background events. For one of the BDTs $t\bar{t}W$ and $t\bar{t}Z$ events are considered as background and for the other BDT, $t\bar{t}$+jets production is considered as background. The events are sorted depending on the output of the two BDTs to obtain a one dimensional histogram with ten bins, which are used in the final fit. The postfit histograms for the combined BDT bins are shown in Fig. 6.22.

The most important uncertainties of the multilepton channels are the lepton selection efficiencies, the normalization of non-prompt backgrounds and the scale of the $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}H$ production.

6.9.3. Diphoton Channel

Similar to the multilepton channels, the Higgs boson to diphoton channel is not directly part of this thesis and therefore it is only briefly described here. Further details on this channel can be found in Refs. [2, 161].

As there is no dedicated tH analysis in the diphoton channel, the two $t\bar{t}H$ categories of the diphoton analysis [161] are reinterpreted. These categories are tailored to enrich the $t\bar{t}H$ process with leptonic and hadronic top quark decays, respectively. As the $t\bar{t}H$ process is similar to the tHq and tHW signal processes studied within this chapter, the $t\bar{t}H$ categories also have contributions from these processes. The diphoton analysis requires two prompt photons with high transverse momentum. For the leptonic $t\bar{t}H$ channel, at least one well-separated and well-reconstructed lepton is required. Furthermore, at least

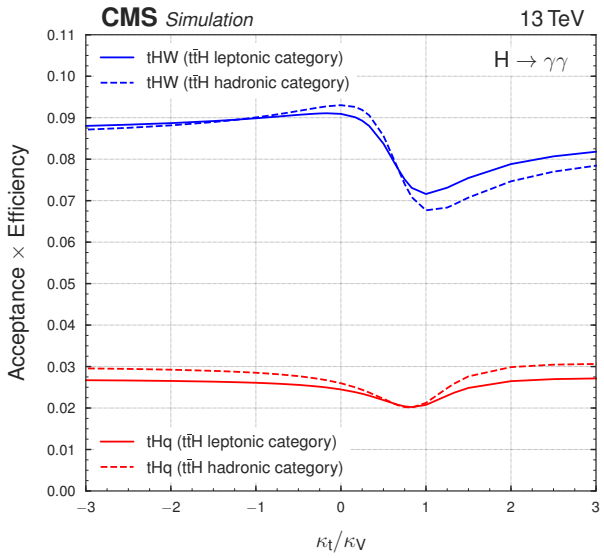

Figure 6.23.: Acceptance for the tHq and tHW signal processes in the diphoton channel: the acceptance and efficiency change depending on which coupling ratio is chosen as a result of the kinematic variations. The behavior for the leptonic and hadronic $t\bar{t}H$ categories is slightly different as indicated by the solid and dashed lines. Taken from Ref. [2].

two jets, at least one of which fulfilling the medium b-tag requirements, are demanded. In the hadronic $t\bar{t}H$ channel, at least three jets, where at least one of them is loosely b tagged, are required.

The different coupling scenarios are considered by computing their impact on the acceptance and efficiency as shown in Fig. 6.23. The diphoton invariant mass is modeled with a sum of Gaussian functions. The background is described using a data-driven profiling method. The diphoton mass peak is used in the final fit to set limits on the tH production and the coupling ratio.

6.9.4. Results

The signal yields are derived by fitting the discriminator outputs of the $b\bar{b}$ and multilepton channels and the invariant diphoton mass. Additionally, a maximum likelihood scan is performed to obtain limits on different coupling ratios. The channels can be combined straightforwardly as their event selections are mutually exclusive. Uncertainties present in all channels, are taken to be fully correlated. These include uncertainties on b tagging, integrated luminosity, and theory predictions.

The scan of the likelihood ratio $\mathcal{L}(\kappa_t)/\mathcal{L}(\hat{\kappa}_t)$ is derived for an SM-like Higgs boson to vector boson coupling, i.e., $\kappa_V = 1$ using tH and $t\bar{t}H$ production as signal. Here, $\hat{\kappa}_t$ is the best fit value of κ_t. The scan is shown in Fig. 6.24. The expected values are derived from an Asimov data set and the observed values are taken from a fit to data. An excess of n standard deviations corresponds to a value of n^2 on the y axis; a 95% interval covers values below 3.84.. The observed (expected) results favor the SM case over the ITC case by about 1.5 (4.0) standard deviations. Values outside of [-0.9,0.5] and [1.0,2.1] are excluded, with an expectation to exclude values outside of $[-0.5, 1.6]$ at 95% confidence level.

The limits on tH cross section times branching ratio are derived by fitting the data with tHq and tHW as signal. These limits are derived for fixed points of κ_t. The limits are shown in Fig. 6.25. The deviation between data and expectation around $\kappa_t = 0$ occurs as the $t\bar{t}H$ contribution would be very small for these points, but the data actually favors a higher than expected contribution. Nonetheless, a fit with independently floating signal strengths for SM-like tH and $t\bar{t}H$ production yield best fit values compatible with the SM prediction within two standard deviations as shown in Fig. 6.26.

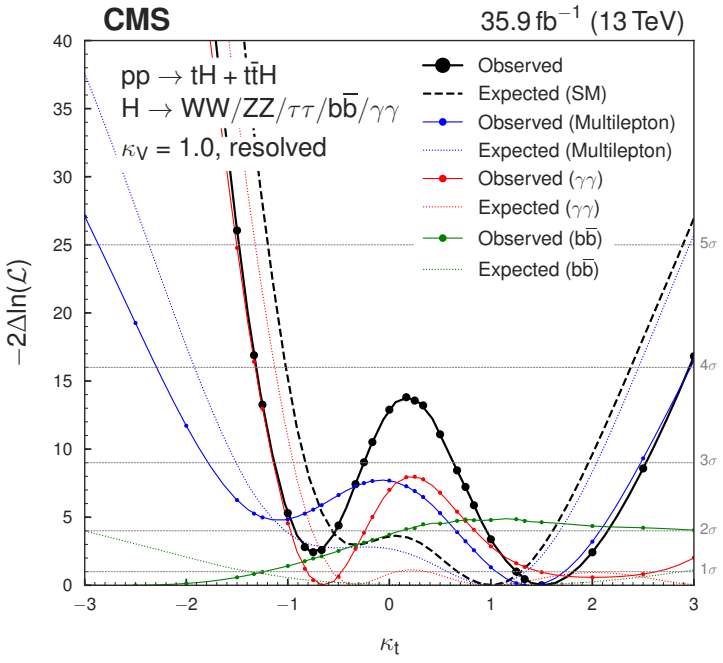

Figure 6.24.: Likelihood ratio scan: the data points are shown in black for the combined result; the individual results are shown in blue, red and green. The expectations are shown as dashed lines. Taken from Ref. [2].

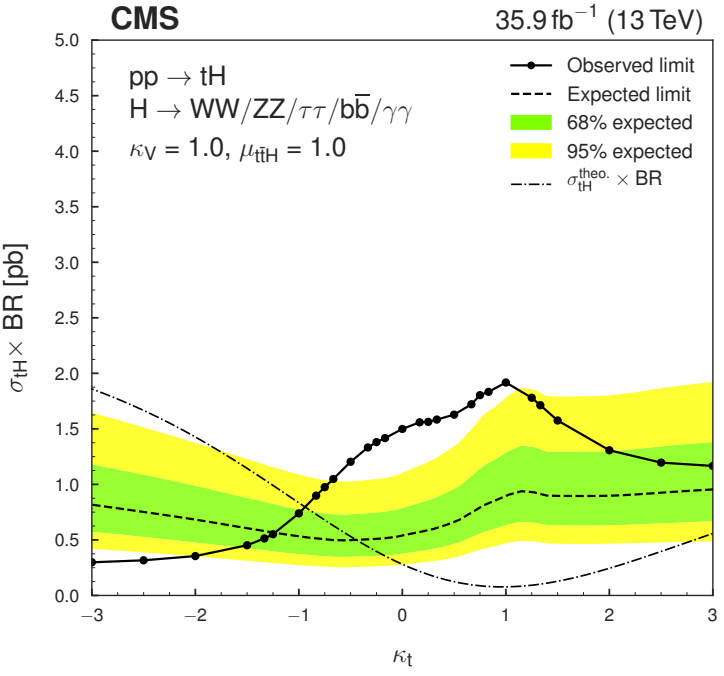

Figure 6.25.: Limits on tH production: the observed limits are shown as dots and the expectation as a dashed line. The expectation is calculated on a data set without any tH contribution, but including a κ_t-dependent $t\bar{t}H$ contribution. The tH contribution is allowed to float in the fit, whereas the $t\bar{t}H$ contribution is fixed to the expected SM event yield. Taken from Ref. [2].

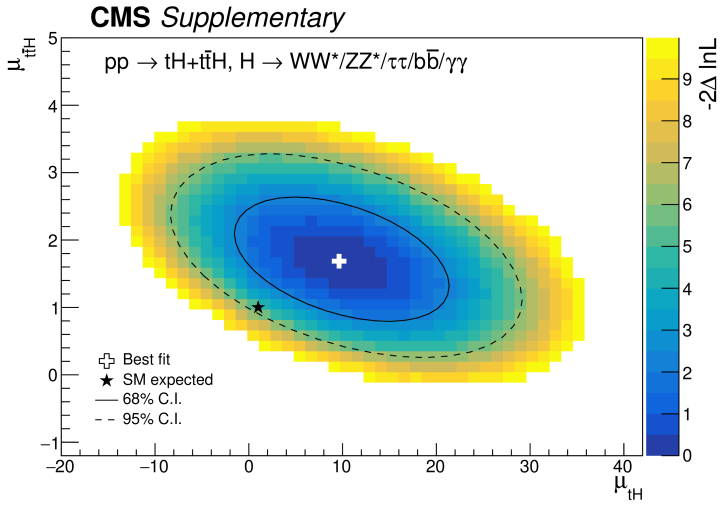

Figure 6.26.: Simultaneous fit of the tH and tt̄H signal strength: the best fit value is shown as a white cross and the SM expectation is depicted as a black cross. The best fit value and the SM prediction are compatible within two standard deviations. Taken from the additional figures of Ref. [2].

6.10. Summary and Outlook

The Higgs boson production in association with a single top quark or in association with a top quark pair is well suited to study the top quark Yukawa coupling. This chapter presented results based on data taken by the CMS experiment in 2016. For the Higgs boson to bb̄ decay channel, upper exclusion limits on the SM-like tH production, different κ_t/κ_V coupling ratios and CP-mixing angles are derived. The upper observed (expected) limit on SM-like tH production at 95% C.L. is 53 (35) and for the ITC scenario 2.9 (2.0). All results are compatible with the SM. The analysis presented in this chapter results in an average improvement of 30% for the expected limits in the bb̄ channel compared to the previously published version [2]. This leads to a constraint of κ_t to $[-2.5, 4.0]$ assuming $\kappa_V = 1$ using only the bb̄ decay channel. For the first time, the bb̄ channel alone is able to exclude certain coupling scenarios. The improvement originates from the implementation of the b-jet energy regression, the simulation of additional signal events and the introduction of a two-fold approach for the BDT trainings.

The combination of the diphoton, multilepton and bb̄ channels yields results that constrain κ_t to values of $[-0.9, 0.5]$ and $[1.0, 2.1]$ assuming $\kappa_V = 1$, i.e., an SM-like Higgs boson to W boson coupling. Hence, the ITC scenario is excluded at 95% C.L.. The simultaneous measurement of the tH and tt̄H signal strength are compatible with the SM.

The next step is to analyze not only the data taken in 2016 by the CMS experiment but the whole data taken in the years 2016 to 2018 using the improvements as introduced in this chapter. In order to use the results of the analysis in a future combination of all Higgs boson related results, it is advantageous to avoid any overlap with other analyses. Therefore, a combined tH and tt̄H analysis targeting a phase space that has the highest sensitivity to both of these processes is desirable. Studies targeting such an analysis are presented in the following chapter.

7. Investigation of Higgs Boson Couplings in tH and $t\bar{t}H$ Events with $H \to b\bar{b}$ with the 2016–2018 Data Set

This chapter presents studies for a combined analysis of the tH and $t\bar{t}H$ production processes using the data taken in the years 2016, 2017, and 2018 by the CMS experiment, corresponding to integrated luminosities of $35.9\,\text{fb}^{-1}$, $41.5\,\text{fb}^{-1}$, and $59.7\,\text{fb}^{-1}$, respectively. The Higgs boson is required to decay into a pair of bottom quarks. The top quark is required to decay leptonically for the tH process. For the $t\bar{t}H$ process, one of the top quarks is required to decay leptonically and the other one to decay hadronically. The analysis presented here is the successor of two analyses: the tH analysis performed with data taken in 2016 [1] as discussed in the previous chapter and the $t\bar{t}H$ analysis with data taken in 2016–2017 [182]. The two analyses are combined in order to avoid any partial overlap of the two analyses. This allows to use the results of this analysis to be part of a combined result of an all-embracing Higgs boson study, similar to the previous combined Higgs boson measurements [41, 42].

The BDT-based jet assignment as performed in the tH analysis has proven to be beneficial, hence it is used in this combination. This was not the case for the $t\bar{t}H$ analysis. Furthermore, the training is performed in a two-fold approach to increase the number of events used for the training and the actual analysis. With respect to the tH analysis, the classifier is changed as well. Here, a DNN with multiple output nodes is used.

Similar to the analysis presented in the previous chapter, jet assignments under different hypotheses as described in Section 5.3 are performed to obtain event specific variables. These variables are, together with global variables, used to train the multi-classification DNNs to separate different signal and background contributions. The outputs of the nodes provide templates, which are used in the final fit. The background nodes are used to constrain the uncertainties of the background processes during the fitting procedure.

7.1. Analysis Strategy

As this analysis is the successor of the tH analysis on the 2016 data set, the signal and background processes are the same: the tHq, tHW and $t\bar{t}H$ processes as described in Section 5.1 are considered as signal. The $t\bar{t}$ process is the main background process. Further contributions come from single top quark production and other processes of minor impact. The background topology is described in Section 5.2.

The study of the Higgs boson coupling to top quarks and W bosons uses a grid of $-3 \leq \kappa_t \leq 3$ and $\kappa_V = 0.5, 1, 1.5$. A complete list of the scenarios and their corresponding cross sections can be found in Appendix A.1. Based on these points, limits on the κ_t/κ_V ratio are set. Additionally, limits on the CP-mixing angle α are set.

Another result of this analysis is the measurement of the SM cross section of the $t\bar{t}H$ process. Furthermore, a limit on SM-like tH production is set.

7.2. Event Selection

The events are selected to obtain signal regions targeting signal-like event topologies as described in Section 5.1. All three signal processes produce four jets originating from bottom quarks and at least

Table 7.1.: Trigger requirements for all three years studied: each event has to pass at least one of the triggers. All triggers require exactly one tight muon or electron.

Year	Channel	Trigger
2016	μ	HLT_IsoMu24_v*
	μ	HLT_IsoTkMu24_v*
	e	HLT_Ele27_WPTight_Gsf_v*
2017	μ	HLT_IsoMu27_v*
	e	HLT_Ele32_WPTight_Gsf_L1DoubleEG_v*
	e	HLT_Ele28_eta2p1_WPTight_Gsf_HT150_v*
2018	μ	HLT_IsoMu24_v*
	e	HLT_Ele32_WPTight_Gsf_v*
	e	HLT_Ele28_eta2p1_WPTight_Gsf_HT150_v*

Table 7.2.: Lepton and jet requirements: the selection criteria are shown for the three years studied. Only objects fulfilling these criteria are considered in the analysis.

	electron		muon		jet	
	$\lvert\eta\rvert$	p_T	$\lvert\eta\rvert$	p_T	$\lvert\eta\rvert$	p_T
2016	< 2.4	$> 29\,\text{GeV}$	< 2.4	$> 26\,\text{GeV}$	$< 2.4(4.7)$	$> 30(40)\,\text{GeV}$
2017	< 2.4	$> 30\,\text{GeV}$	< 2.4	$> 29\,\text{GeV}$	$< 2.4(4.7)$	$> 30(40)\,\text{GeV}$
2018	< 2.4	$> 30\,\text{GeV}$	< 2.4	$> 26\,\text{GeV}$	$< 2.4(4.7)$	$> 30(40)\,\text{GeV}$

one additional jet. Since this analysis targets leptonic decays, one tight electron or muon is required. The triggers used within this analysis are listed in Table 7.1 for all three years. As already discussed for the tH analysis presented in the previous chapter, one of the b jets often fails the jet requirements. Consequently, two signal regions are defined: one with exactly three b-tagged jets and one with at least four b-tagged jets. Additionally, a region with at least three central jets, at least three b-tagged jets and at least one forward jet is introduced. The medium b-tag working point, corresponding to a mistag rate of 1%, is used for all three years. Only jets in the central region ($|\eta| < 2.4$) can be b-tagged. The criteria are summarized in Table 7.2. Events containing tauons are only considered indirectly if these further decay into electrons or muons. The event selection for the three signal regions is summarized in Table 7.3.

Table 7.3.: Selection criteria for the signal regions: each event has to pass the respective lepton, p_T^{miss}, and jet selections. The signal regions are defined by the number of jets (j), number of b-tagged jets (t), and number of forward jets with $|\eta| > 2.4$ (f).

	$\geq$ 4j, 3t, 0f	$\geq$ 4j, $\geq$ 4t, 0f	$\geq$ 3j, $\geq$ 3t, $\geq$ 1f
# tight leptons	1	1	1
# additional loose leptons	0	0	0
p_T^{miss}	$> 20\,GeV$	$> 20\,GeV$	$> 20\,GeV$
# jets with $\|\eta\| < 2.4$	≥ 4	≥ 4	≥ 3
# jets with $\|\eta\| > 2.4$	0	0	≥ 1
# b-tagged jets	3	≥ 4	≥ 3

7.3. Systematic Uncertainties

Two classes of systematic uncertainties are considered, i.e., theoretical and experimental. The experimental uncertainties cover uncertainties on luminosity, b tagging, pileup, jet energy resolution and scale, and lepton efficiencies. The theoretical uncertainties include the uncertainties associated with pdfs, initial-state and final-state radiation, and the factorization and renormalization scales. A detailed description of these uncertainties can be found in Section 6.5.
The unclustered energy is not considered in this analysis as its impact is negligible. For the analysis presented in the previous chapter, it is not within the 30 uncertainties with the highest impact on the signal strength as shown in Appendix A.8.

7.4. Jet Assignment

Similar to the tH analysis, a jet assignment under different hypotheses is performed as described in Section 5.3. As this analysis is dedicated to study tH and $t\bar{t}H$ events, jet assignments are performed under tHq, tHW, and $t\bar{t}H$ signal hypotheses and under a $t\bar{t}$ background hypothesis. The assignment is trained separately for each of the three years using the events of the three signal regions of the respective year. Since the figures for the three years studied within this analysis differ only sightly, the histograms for 2017 are shown in Fig. 7.1, Fig. 7.2, Fig. 7.3, and Fig. 7.4 for the assignments under the tHq, tHW, $t\bar{t}H$, and $t\bar{t}$ hypothesis, respectively. The figures for 2016 and 2018 are shown in Appendix B.1. A complete list of all variables for the specific JA BDTs ranked by their importance in the training can be found in Appendix B.2.

The distributions of the JA BDTs on the training data set and on an independent test data set are shown in Fig. 7.5 and 7.6 for 2017. The corresponding distributions for 2016 and 2018 can be found in Appendix B.4. The distributions for the test and training data set agree within their uncertainties, i.e., no overtraining is observed.

The BDT-based jet assignment results in a significant improvement over the χ^2-based approach, which was used in previous $t\bar{t}H$ analyses for an assignment under the $t\bar{t}H$ hypothesis. For the $t\bar{t}H$ hypothesis, the BDT-based approach is more than twice as efficient as the χ^2-based approach. The assignment efficiencies are listed in Table 7.4 for both approaches and all three years.

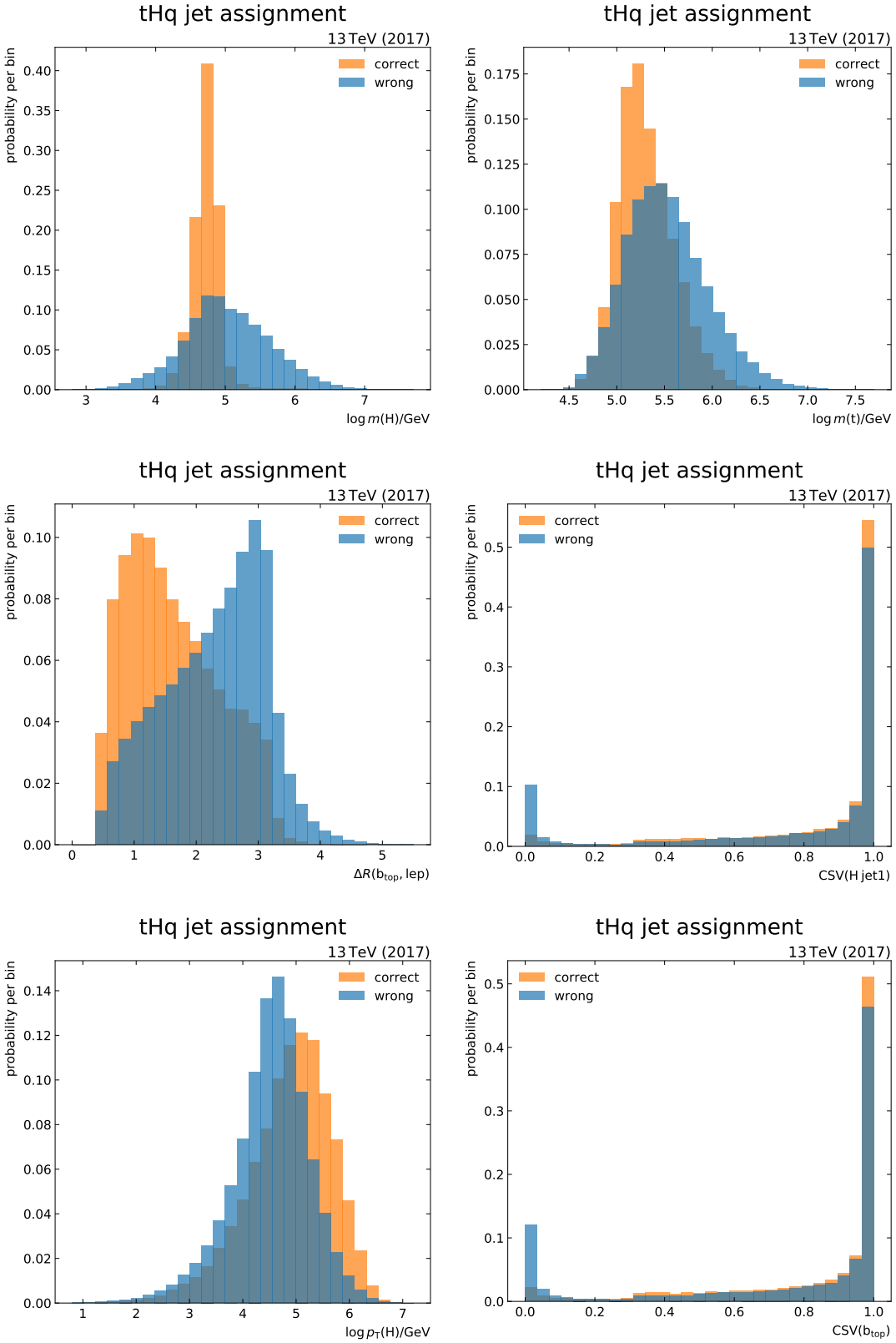

Figure 7.1.: Distributions for the most important variables of the tHq jet assignment: the distributions are shown for correct and wrong assignments for category A and B events combined. A description of the variables can be found in Table 5.2.

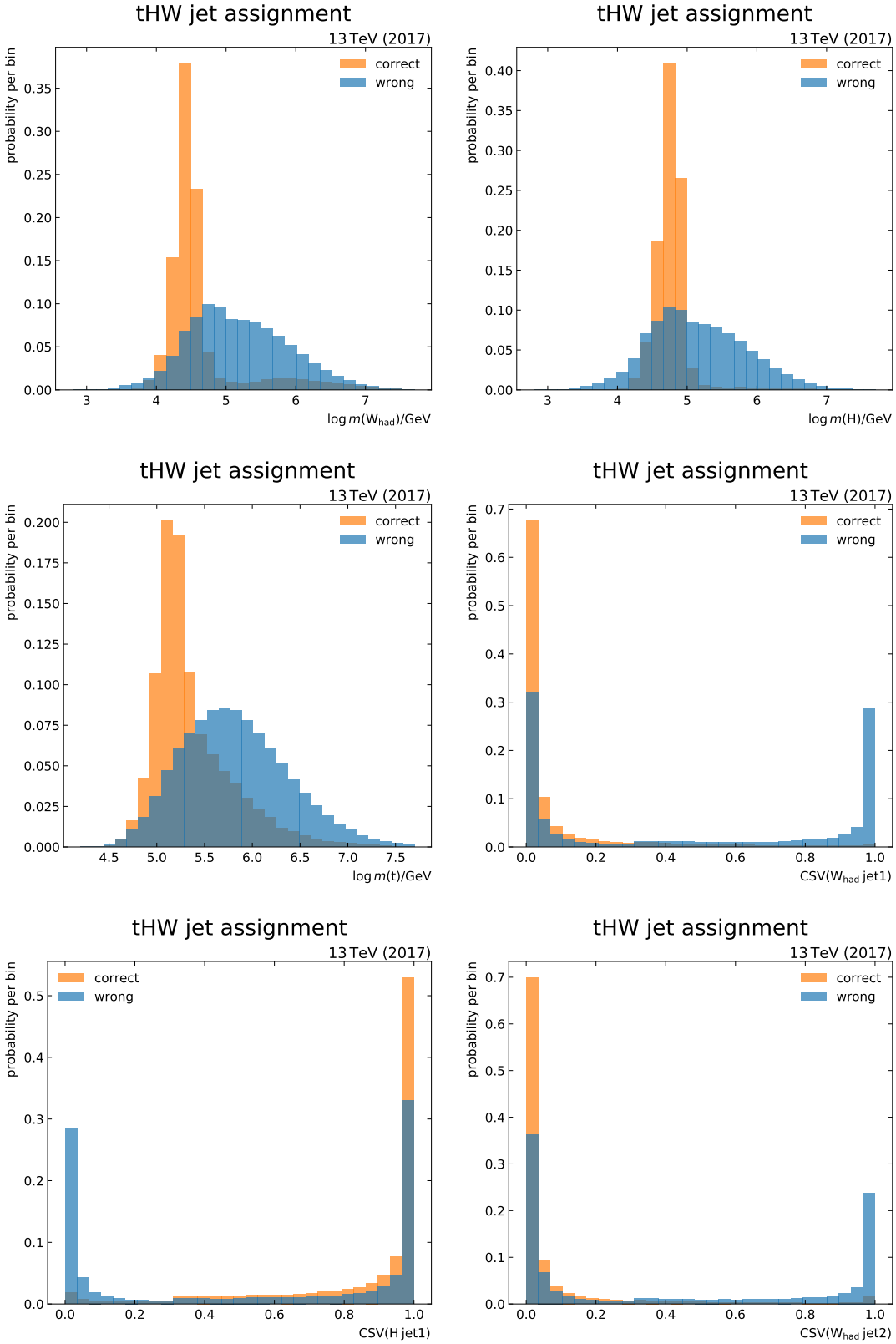

Figure 7.2.: Distributions for the most important variables of the tHW jet assignment: the distributions are shown for correct and wrong assignments for category A and B events combined. A description of the variables can be found in Table 5.3.

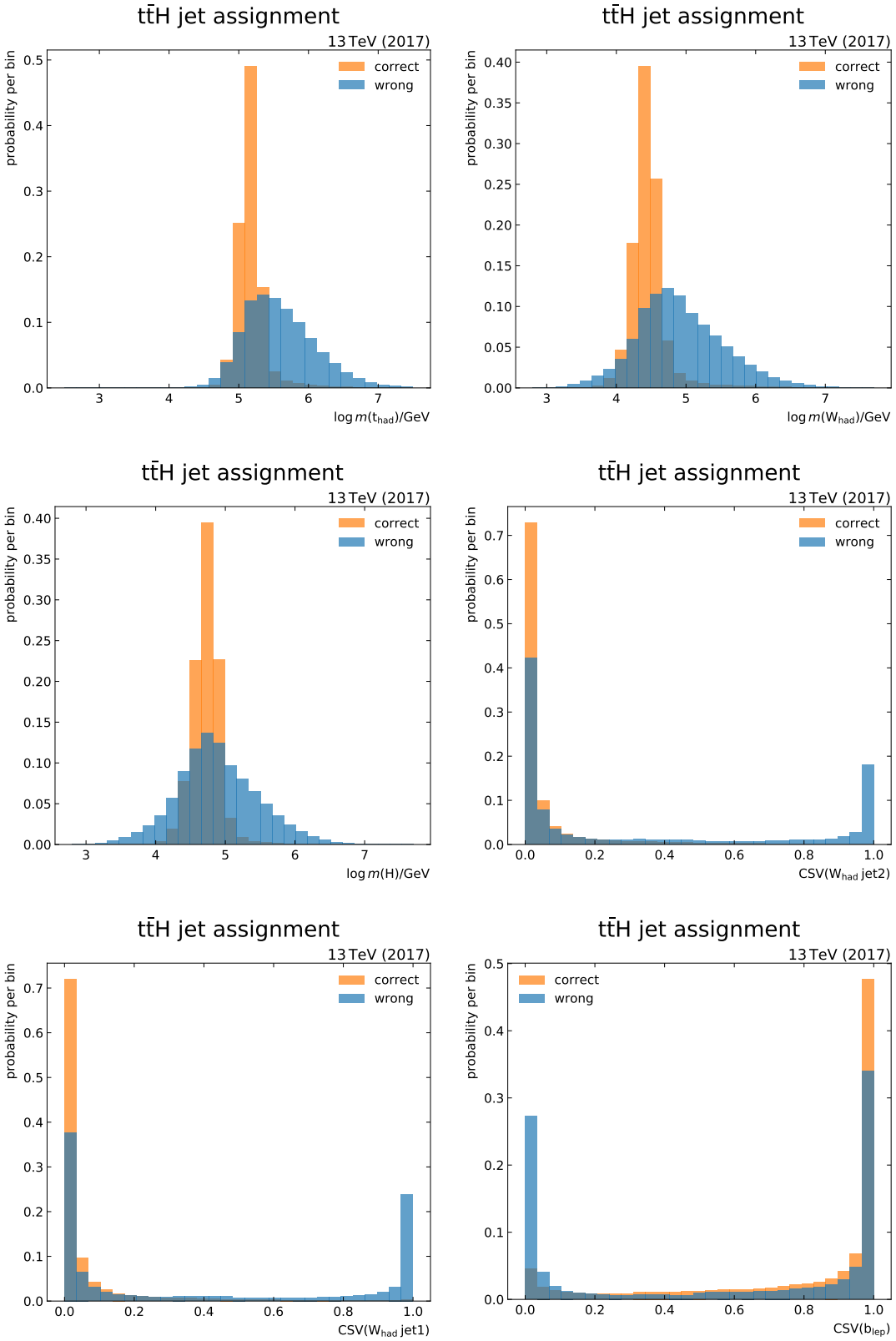

Figure 7.3.: Distributions for the most important variables of the $t\bar{t}H$ jet assignment: the distributions are shown for correct and wrong assignments for category A and B events combined. A description of the variables can be found in Table 5.4.

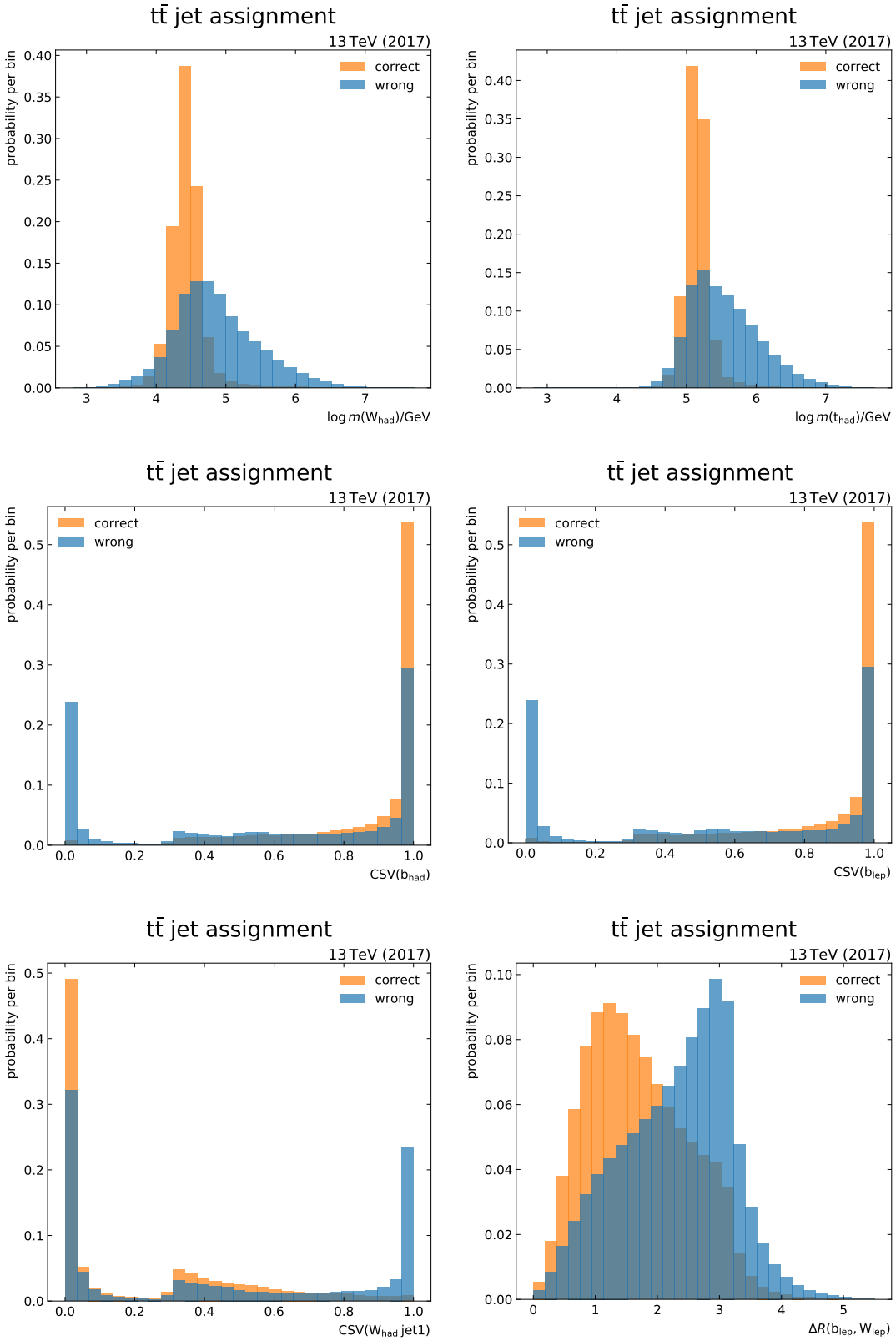

Figure 7.4.: Distributions for the most important variables of the $t\bar{t}$ jet assignment: the distributions are shown for correct and wrong assignments for category A and B events combined. The CSV output distributions for the jets assigned to the bottom quarks of the top quark decays are very similar. Nonetheless, both variables provide valuable input to the jet assignment. A description of the variables can be found in Table 5.5.

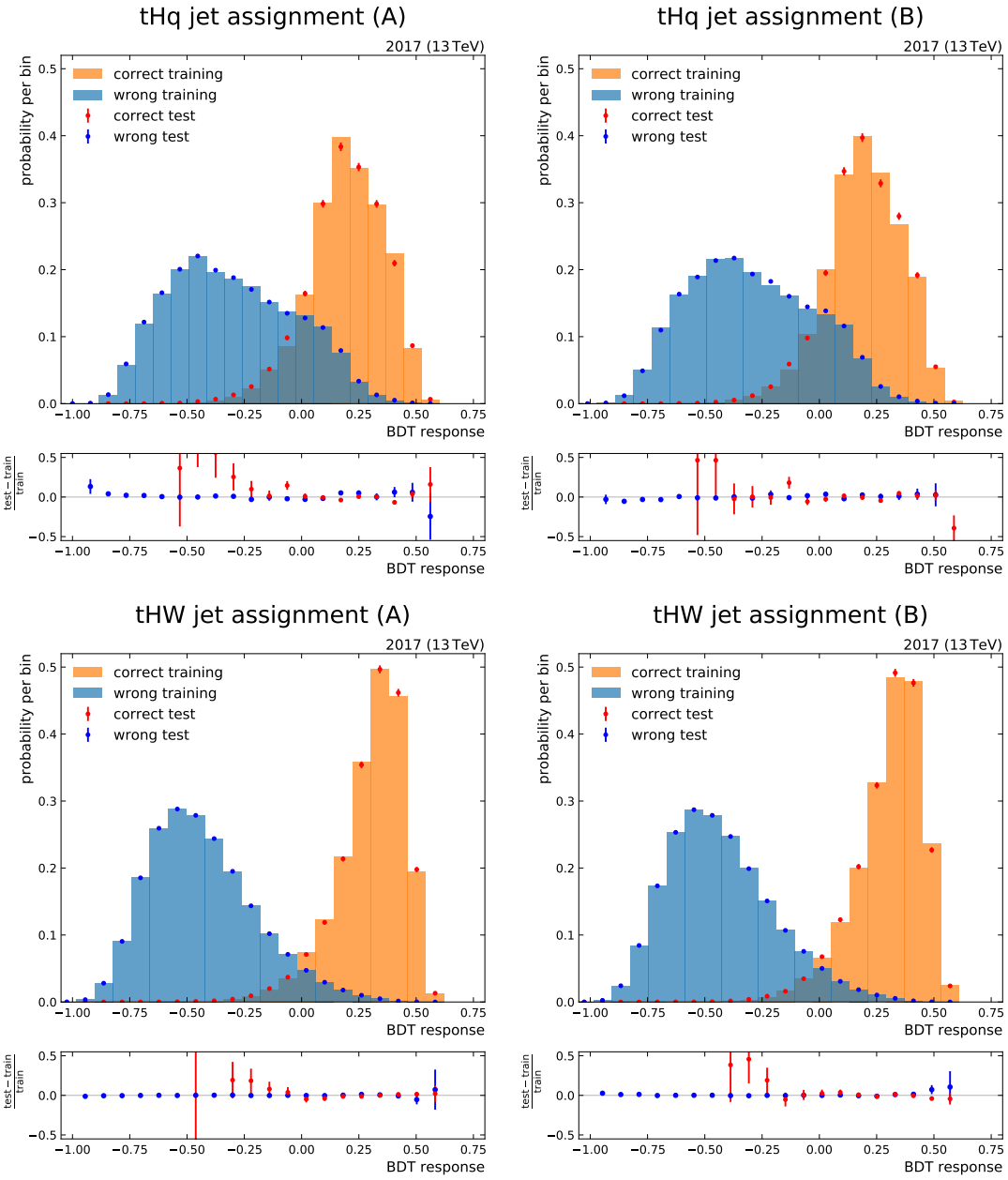

Figure 7.5.: Overtraining check for tHq (top) and tHW (bottom) JA BDTs: the distributions on the left show the BDT response for the training on category A events and testing on category B. The right plots show the BDT response for the training on category B events and testing on category A. All distributions show good agreement between the outputs for the test and training data sets. No indication for overtraining is found. Since the wrong assignment contains the same events several times with different assignments, these events are not uncorrelated. Hence, a Kolmogorow-Smirnow test does not provide a good figure of merit, therefore it is not calculated here.

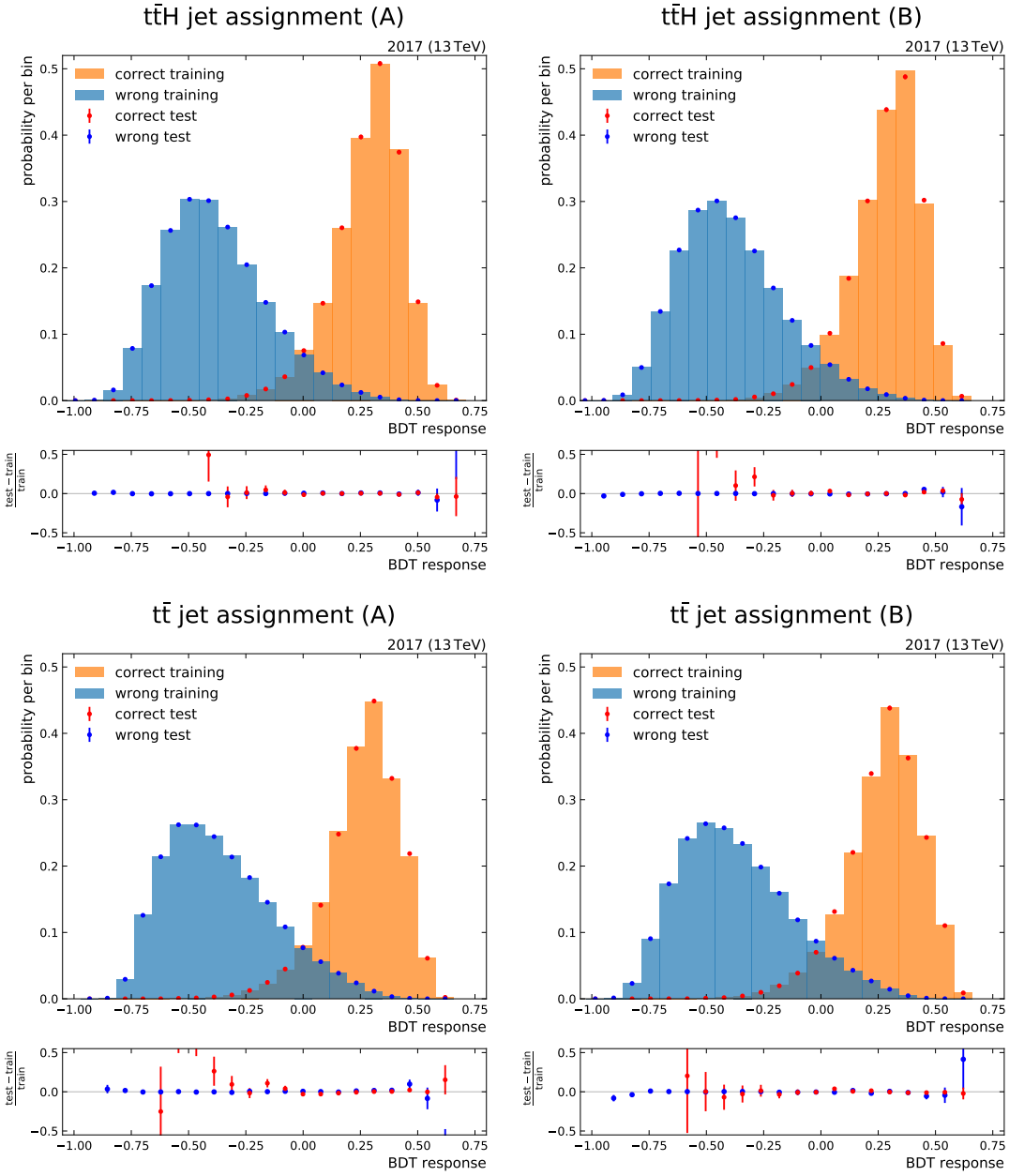

Figure 7.6.: Overtraining check for tt̄H (top) and tt̄ (bottom) JA BDTs: the distributions on the left show the BDT response for the training on category A events and testing on category B. The right plots the BDT response for show the training on category B events and testing on category A. All distributions show good agreement between the outputs for the test and training data sets. No indication for overtraining is found. Since the wrong assignment contains the same events several times with different assignments, these events are not uncorrelated. Hence, a Kolmogorow-Smirnow test does not provide a good figure of merit, therefore it is not calculated here.

Table 7.4.: Efficiencies of the jet assignments: the efficiencies of the tHq, tHW, $t\bar{t}H$ and $t\bar{t}$ jet assignment as defined in Section 5.3.5 are listed for all three years for the BDT-based (χ^2-based) approach. The χ^2-based efficiencies are listed for comparison. Here only the total efficiencies are listed, the efficiencies for the individual final-state particles can be found in Appendix B.3.

	2016	2017	2018
tHq	57% (43%)	58% (42%)	58% (42%)
tHW	51% (28%)	51% (26%)	52% (26%)
$t\bar{t}H$	45% (20%)	44% (18%)	43% (18%)
$t\bar{t}$	67% (36%)	66% (32%)	66% (32%)

7.5. Event Classification

The signal processes are rare compared to the background processes, after applying the event selection. In order to increase the significance, an event classification is performed. This event classification combines the information of several variables, i.e., global event variables and variables obtained by the jet assignments, to achieve a good separation. In contrast to the analysis presented in the previous chapter, the classification used in this analysis does not only classify the events into one single signal and background category but splits these categories further. Here, a DNN with the following output nodes is used: tHW, tHq, $t\bar{t}H$, ttb_bb, which targets $t\bar{t}$+b and $t\bar{t}$+$b\bar{b}$ events, tt2b, ttcc and ttlf. Minor background processes are not considered in the training. The sum of all outputs is normalized to one. The splitting of the different background nodes allows to obtain different shapes for the different $t\bar{t}$+jets categories. This yields in the same benefits as obtained by introducing the dileptonic background region in the analysis presented in the previous chapter: the rate uncertainties of these contributions are constrained in the final fit.

The classification DNN consists of three layers with 100 neurons each using the exponential linear unit as activation function. After each layer, a dropout regularization of 50% is implemented. The cross entropy is used as loss function. Three classification DNNs are trained, one in each signal region. The DNNs as presented here are trained with the simulated events for the year 2017. The confusion matrix for the training in the $\geq 4\text{j}, 3\text{t}, 0\text{f}$ region is shown in Fig. 7.7. The matrices for the remaining regions are shown in Appendix B.5. The final analysis is planned to use DNNs that is trained on the simulated samples of all three years combined.

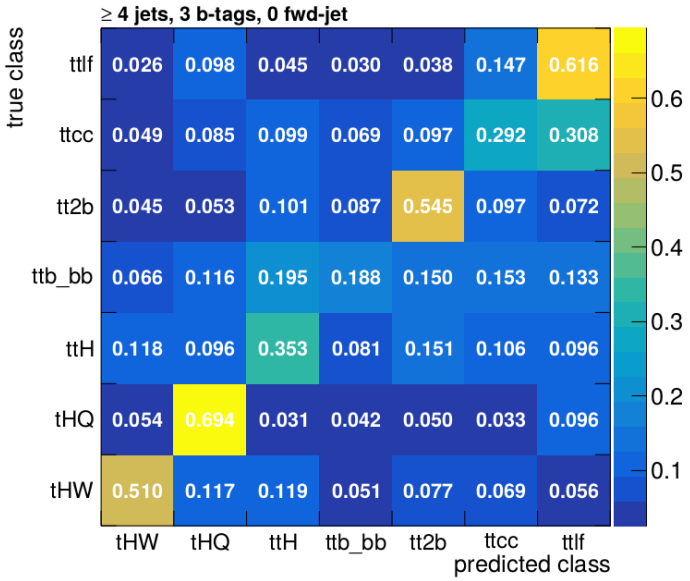

Figure 7.7.: Confusion matrix of the classification DNN for the $\geq$ 4j, 3t, 0f region: the matrix shows the predicted class for a given true class, e.g., 51% of the tHW events are classified as tHW. The DNN shows especially good results for the tHq and tHW signal processes, as well as for the $t\bar{t}$lf and $t\bar{t}$2b background processes.

7.6. Expected Results

The results presented in the following are derived using the 2017 simulation only, i.e., an Asimov dataset [81] is utilized in the fit. The expected signal strength for the $t\bar{t}H$ process in the $H \to b\bar{b}$ decay channel is $\mu = 1.0 \pm 0.49$ for 2017 only. In order to obtain the significance for the three years combined, the luminosity is scaled to 140 fb^{-1}. The expected signal strength for the whole three year period is $\mu = 1.0 \pm 0.26$, which corresponds to an expected significance of 3.9 standard deviations. In order to increase the significance, the results of the semileptonic channel are combined with the dileptonic and fully-hadronic channels. This results in an expected significance of 4.4 standard deviations for the three years combined. The impacts of the systematic uncertainties can be found in Appendix B.6.

The following results are obtained using the semileptonic channel. The expected upper limits on the SM-like tH production signal strength modifier are 14.6 $[10.1, 21.7]$ when assuming the luminosity of 2017 only, the values in brackets correspond to ± 1 standard deviation. Compared to the expected results of the previous chapter this is an improvement of about 58% on a data set of comparable size. The improved limit is a result of the multi-classification DNN and the use of the DeepCSV b-tagging algorithm. When scaling the luminosity to 140 fb^{-1}, which corresponds to the 2016–2018 data set, the expected limit is 8.3 $[5.7, 12.2]$, which is an improvement to the expected limits of the 2016 analysis as presented in the previous chapter of about 76%. The expected limits for the κ_t/κ_V coupling ratio scaled to the luminosity of the full data set are shown in Fig. 7.8. A comparison of the analysis presented in this chapter with the analysis on the 2016 data set is shown in Fig. 7.9. The expected upper exclusion limits scaled to 140 fb^{-1}are improved by 79% on average compared to the analysis presented in Chapter 6 and about 86% compared to the published analysis [2].

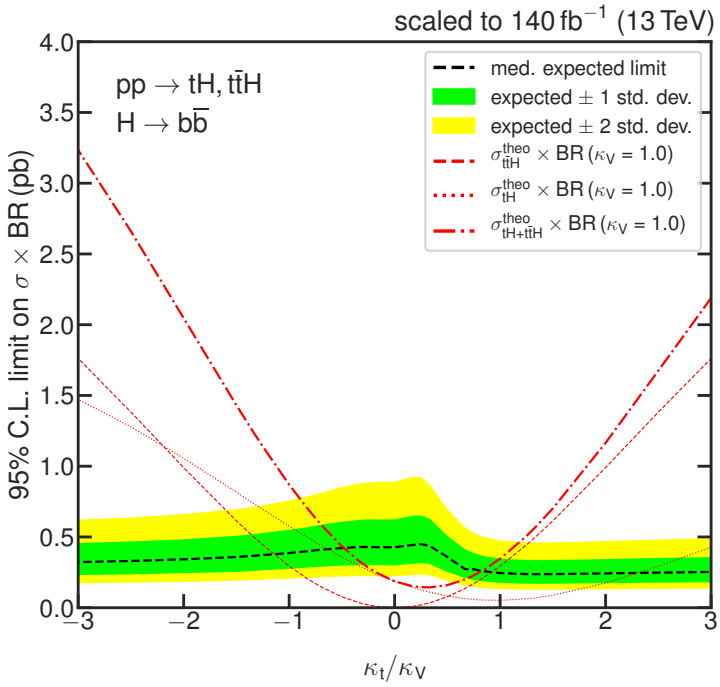

Figure 7.8.: Expected limits on cross section times branching ratio for different coupling ratios: the luminosity is scaled to 140 fb^{-1}. The analysis is expected to be sensitive to the regions, where the expected limits are below the theory prediction.

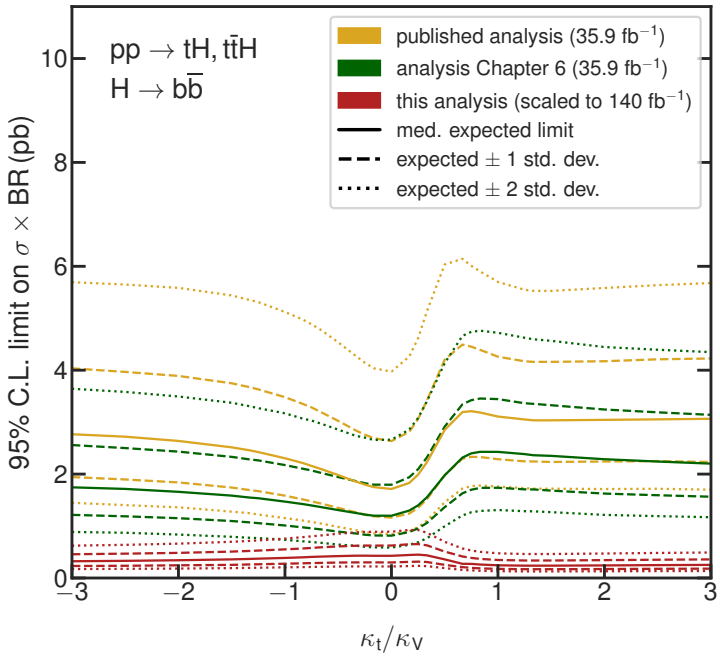

Figure 7.9.: Comparison of expected limits on cross section times branching ratio for different coupling ratios: the limits of the published analysis [2], the improved analysis as presented in Chapter 6 and the analysis presented in this chapter. The analysis of the 2016–2018 data set results in better limits than the previous analyses for each coupling ratio.

8. Conclusion and Outlook

An investigation of the top quark Yukawa coupling in Higgs boson production in association with top quarks with data taken by the CMS experiment has been described. The analysis investigates the sign and magnitude of the top quark Yukawa coupling directly. In this thesis, the $H \to b\bar{b}$ decay channel is investigated. The search for Higgs boson production in association with a single top quark is carried out in the multileptons [160] and diphoton [161] decay channels as well. Within the scope of this thesis, these channels are combined with the $H \to b\bar{b}$ decay channel to achieve a higher sensitivity. As a consequence, the coupling modifier κ_t is constrained to values of $[-0.9, 0.5]$ and $[1.0, 2.1]$ at 95% C.L. assuming $\kappa_V = 1$, i.e., the ITC scenario is excluded at 95% C.L.. The combined analysis without the improvements as introduced in this thesis is published in Ref. [2].

In addition to the published analysis [1, 2], several improvements have been introduced in the analysis of $H \to b\bar{b}$ decay channel: the energy resolution of jets originating from bottom quarks is enhanced via a b-jet energy regression method, the number of simulated events available in the training of the machine learning methods is increased through a two-fold training approach and additional signal events are simulated. This leads to an average improvement compared to the published analysis by about 30% in the $H \to b\bar{b}$ decay channel on the same data set. The analysis of the $H \to b\bar{b}$ decay channel is able to constrain the coupling modifier κ_t to $[-2.5, 4.0]$ assuming $\kappa_V = 1$, i.e., an SM-like Higgs to W boson coupling. The upper observed (expected) limit on the signal strength for the SM-like tH production at 95% C.L. is 53.3 (35.1). In addition to the published analysis, this thesis features upper exclusion limits on a possible CP-odd contribution to the Higgs boson coupling to top quarks. For the first time, the kinematic variations of the $t\bar{t}H$ process introduced by the modified coupling are taken into account.

The $t\bar{t}H$ production is directly sensitive to the magnitude of the top quark Yukawa coupling, hence it is included as another signal process for the coupling scan in the analyses mentioned above. The event selection is optimized for Higgs boson production associated with a single top quark, though. In order to optimize the sensitivity to both processes, a combined analysis of Higgs boson production in association with a single top quark or a top quark pair is beneficial. Furthermore, this enables an all-embracing Higgs boson combination similar to the analysis presented in Ref. [41, 42] to incorporate these results as there are no data events considered twice. Consequently, a study for such an analysis with the 2016–2018 data set is presented in Chapter 7 including the improvements developed within this thesis. The expected significance for SM-like $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel is 4.4 standard deviations for a combined study of fully-leptonic, semileptonic and fully-hadronic decays of the top quark pair. The analysis of the coupling modifiers, as well as the upper limit on SM-like tH production, is performed in the semileptonic channel since this channel yields the highest sensitivity to Higgs boson production in association with a single top quark. The expected upper limits on the SM-like tH production at 95% C.L. is $\mu_{exp} = 8.3$, when assuming the luminosity of the 2016–2018 data set. This is an improvement over the already improved analysis with the 2016 data set of about 76%. The expected upper exclusion limits for the coupling scan are improved by 79% on average compared to the analysis presented in Chapter 6 and about 86% compared to the published analysis [2], when assuming the luminosity of the 2016–2018 data set.

As a next step, the analysis has to be carried out on the full simulation data set. In addition, the analysis of the tH process can be extended to the dileptonic and fully-hadronic $t\bar{t}$ decay channels to improve the sensitivity. After that, the unblinding step can be performed and the analysis is applied to data. This will result in a measurement of the $t\bar{t}H$ and tH signal strengths, upper exclusion limits for different Higgs boson coupling ratios, as well as on a possible CP-mixing of the Higgs boson coupling.

These results will provide further insight to whether the Higgs boson properties are compatible with the SM predictions. Especially the tH process gives valuable information on the relative sign between the top quark Yukawa coupling and the coupling of the Higgs boson to the W boson in a direct measurement of the ratio of these couplings. The direct measurement of the $t\bar{t}H$ and tH processes is furthermore beneficial as it is not sensitive to any loop contributions in the production or decay of the Higgs boson.

Once the analysis of the $t\bar{t}H$ and tH processes is finished, the results can be used to obtain the Higgs boson properties in a simultaneous fit of all Higgs boson related analyses carried out by the CMS collaboration on the 2016–2018 data set. This will provide measurements of the Higgs boson couplings to bosons, leptons and quarks. Finally, the results of the CMS experiment can be combined with the results provided by the ATLAS collaboration to achieve the highest possible precision.

The CMS and ATLAS experiments will take more data from 2021 onwards. This gives the possibility for observations of rare production processes like the tH production or rare decays as the Higgs boson decay into a pair of charm quarks or muons.

A. Search for tH production

This appendix provides further information to the tH production as presented in Chapter 6.

A.1. Cross Sections of Signal Processes

The cross sections as a function of κ_V and κ_t can be found in Table A.1 for the tHq process and in Table A.2 for the tHW process.

Table A.1.: Production cross sections for tHq at $\sqrt{s} = 13\,\text{TeV}$: the cross sections are obtained from simulations using MadGraph5_aMC@NLO at NLO in the 4F scheme. The uncertainties to the cross sections correspond to scale variations in %. Source: [183].

κ_t	κ_V	σ (pb)	κ_t	κ_V	σ (pb)	κ_t	κ_V	σ (pb)
−3.0	0.5	$2.260^{+1.9}_{-2.7}$	−3.0	1.0	$2.991^{+2.1}_{-3.1}$	−3.0	1.5	$3.845^{+2.6}_{-3.2}$
−2.0	0.5	$1.160^{+2.0}_{-2.9}$	−2.0	1.0	$1.706^{+2.6}_{-3.2}$	−2.0	1.5	$2.371^{+2.5}_{-3.6}$
−1.5	0.5	$0.748^{+2.1}_{-3.1}$	−1.5	1.0	$1.205^{+2.5}_{-3.6}$	−1.5	1.5	$1.784^{+2.7}_{-3.9}$
−1.25	0.5	$0.573^{+2.1}_{-3.0}$	−1.25	1.0	$0.987^{+2.6}_{-3.4}$	−1.25	1.5	$1.518^{+2.8}_{-3.9}$
−1.0	0.5	$0.457^{+2.3}_{-3.3}$	−1.0	1.0	$0.793^{+2.7}_{-3.9}$	−1.0	1.5	$1.287^{+3.0}_{-4.3}$
−0.75	0.5	$0.300^{+2.5}_{-3.5}$	−0.75	1.0	$0.621^{+2.9}_{-4.1}$	−0.75	1.5	$1.067^{+3.1}_{-4.4}$
−0.5	0.5	$0.198^{+2.8}_{-3.9}$	−0.5	1.0	$0.457^{+3.2}_{-4.4}$	−0.5	1.5	$0.874^{+3.4}_{-4.7}$
−0.25	0.5	$0.119^{+3.1}_{-4.6}$	−0.25	1.0	$0.351^{+3.5}_{-5.0}$	−0.25	1.5	$0.703^{+3.6}_{-5.0}$
0.0	0.5	$0.062^{+3.8}_{-5.6}$	0.0	1.0	$0.248^{+3.9}_{-5.5}$	0.0	1.5	$0.558^{+3.8}_{-5.4}$
0.25	0.5	$0.028^{+5.0}_{-7.1}$	0.25	1.0	$0.169^{+4.4}_{-6.2}$	0.25	1.5	$0.454^{+4.2}_{-6.1}$
0.5	0.5	$0.018^{+4.2}_{-6.7}$	0.5	1.0	$0.113^{+5.0}_{-7.1}$	0.5	1.5	$0.334^{+4.6}_{-6.5}$
0.75	0.5	$0.030^{+1.4}_{-2.9}$	0.75	1.0	$0.081^{+5.7}_{-7.6}$	0.75	1.5	$0.256^{+5.2}_{-7.2}$
1.0	0.5	$0.066^{+1.0}_{-3.6}$	1.0	1.0	$0.071^{+4.1}_{-6.7}$	1.0	1.5	$0.200^{+5.7}_{-7.6}$
1.25	0.5	$0.124^{+0.9}_{-3.7}$	1.25	1.0	$0.084^{+2.3}_{-4.6}$	1.25	1.5	$0.167^{+5.5}_{-7.5}$
1.5	0.5	$0.205^{+0.8}_{-3.7}$	1.5	1.0	$0.120^{+1.2}_{-2.9}$	1.5	1.5	$0.159^{+4.1}_{-6.7}$
2.0	0.5	$0.454^{+1.0}_{-3.6}$	2.0	1.0	$0.260^{+1.0}_{-3.6}$	2.0	1.5	$0.211^{+2.0}_{-3.9}$
3.0	0.5	$1.177^{+1.2}_{-3.2}$	3.0	1.0	$0.821^{+0.8}_{-3.7}$	3.0	1.5	$0.589^{+0.9}_{-3.7}$

Table A.2.: Production cross sections for tHW at $\sqrt{s} = 13\,\text{TeV}$: the cross sections are obtained from simulations using MadGraph5_aMC@NLO at NLO in the 5F scheme. The uncertainties to the cross section correspond to scale variations in %. Source: [183].

κ_t	κ_V	σ (pb)	κ_t	κ_V	σ (pb)	κ_t	κ_V	σ (pb)
-3.0	0.5	$0.514^{+2.3}_{-3.0}$	-3.0	1.0	$0.641^{+2.3}_{-2.7}$	-3.0	1.5	$0.783^{+2.1}_{-2.1}$
-2.0	0.5	$0.255^{+2.3}_{-2.8}$	-2.0	1.0	$0.346^{+2.2}_{-2.5}$	-2.0	1.5	$0.456^{+2.1}_{-2.1}$
-1.5	0.5	$0.159^{+2.3}_{-2.8}$	-1.5	1.0	$0.253^{+2.1}_{-2.2}$	-1.5	1.5	$0.329^{+1.9}_{-1.8}$
-1.25	0.5	$0.120^{+2.2}_{-2.5}$	-1.25	1.0	$0.188^{+2.0}_{-2.0}$	-1.25	1.5	$0.275^{+1.9}_{-1.6}$
-1.0	0.5	$0.087^{+2.1}_{-2.3}$	-1.0	1.0	$0.147^{+2.0}_{-1.8}$	-1.0	1.5	$0.224^{+1.9}_{-1.5}$
-0.75	0.5	$0.059^{+2.0}_{-2.1}$	-0.75	1.0	$0.110^{+2.0}_{-1.7}$	-0.75	1.5	$0.180^{+1.8}_{-1.3}$
-0.5	0.5	$0.037^{+1.9}_{-1.8}$	-0.5	1.0	$0.080^{+1.7}_{-1.4}$	-0.5	1.5	$0.141^{+1.6}_{-1.2}$
-0.25	0.5	$0.020^{+1.8}_{-1.3}$	-0.25	1.0	$0.055^{+1.6}_{-1.1}$	-0.25	1.5	$0.108^{+1.6}_{-1.2}$
0.0	0.5	$0.009^{+1.6}_{-1.3}$	0.0	1.0	$0.036^{+1.5}_{-1.2}$	0.0	1.5	$0.081^{+1.5}_{-1.2}$
0.25	0.5	$0.004^{+2.1}_{-2.0}$	0.25	1.0	$0.022^{+1.6}_{-1.5}$	0.25	1.5	$0.059^{+1.5}_{-1.4}$
0.5	0.5	$0.004^{+4.6}_{-6.1}$	0.5	1.0	$0.014^{+2.1}_{-2.0}$	0.5	1.5	$0.043^{+1.8}_{-1.7}$
0.75	0.5	$0.010^{+4.7}_{-6.3}$	0.75	1.0	$0.012^{+3.2}_{-3.9}$	0.75	1.5	$0.033^{+2.1}_{-2.0}$
1.0	0.5	$0.021^{+4.0}_{-5.5}$	1.0	1.0	$0.016^{+4.6}_{-6.1}$	1.0	1.5	$0.028^{+2.8}_{-3.0}$
1.25	0.5	$0.038^{+3.7}_{-5.2}$	1.25	1.0	$0.025^{+4.8}_{-5.4}$	1.25	1.5	$0.029^{+3.6}_{-4.7}$
1.5	0.5	$0.061^{+3.5}_{-4.9}$	1.5	1.0	$0.039^{+4.6}_{-6.3}$	1.5	1.5	$0.035^{+4.6}_{-6.0}$
2.0	0.5	$0.125^{+3.0}_{-4.3}$	2.0	1.0	$0.086^{+4.0}_{-5.5}$	2.0	1.5	$0.065^{+4.8}_{-6.5}$
3.0	0.5	$0.317^{+2.8}_{-4.0}$	3.0	1.0	$0.247^{+3.3}_{-4.6}$	3.0	1.5	$0.193^{+4.0}_{-5.6}$

A.2. Kinematic Variations of Signal Processes

The kinematic properties of the final-state particles as a function of κ_t can be found in Fig. A.1 for the tHq process and in Fig. A.2 for the tHW process.

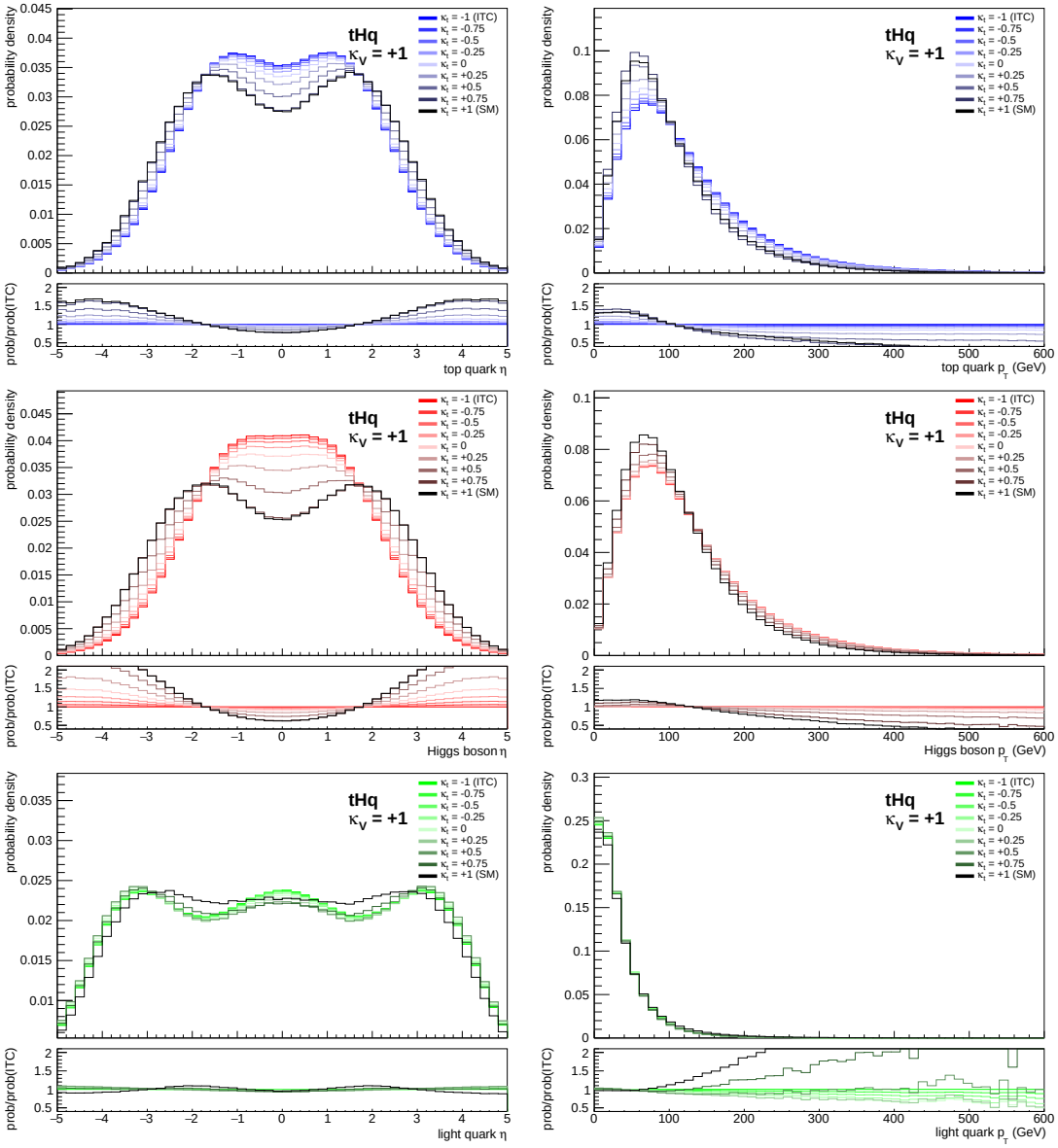

Figure A.1.: Kinematic variations of the tHq final-state particles for different coupling scenarios: the η and p_{T} distributions are shown for the top quark, the Higgs boson and the light quark.

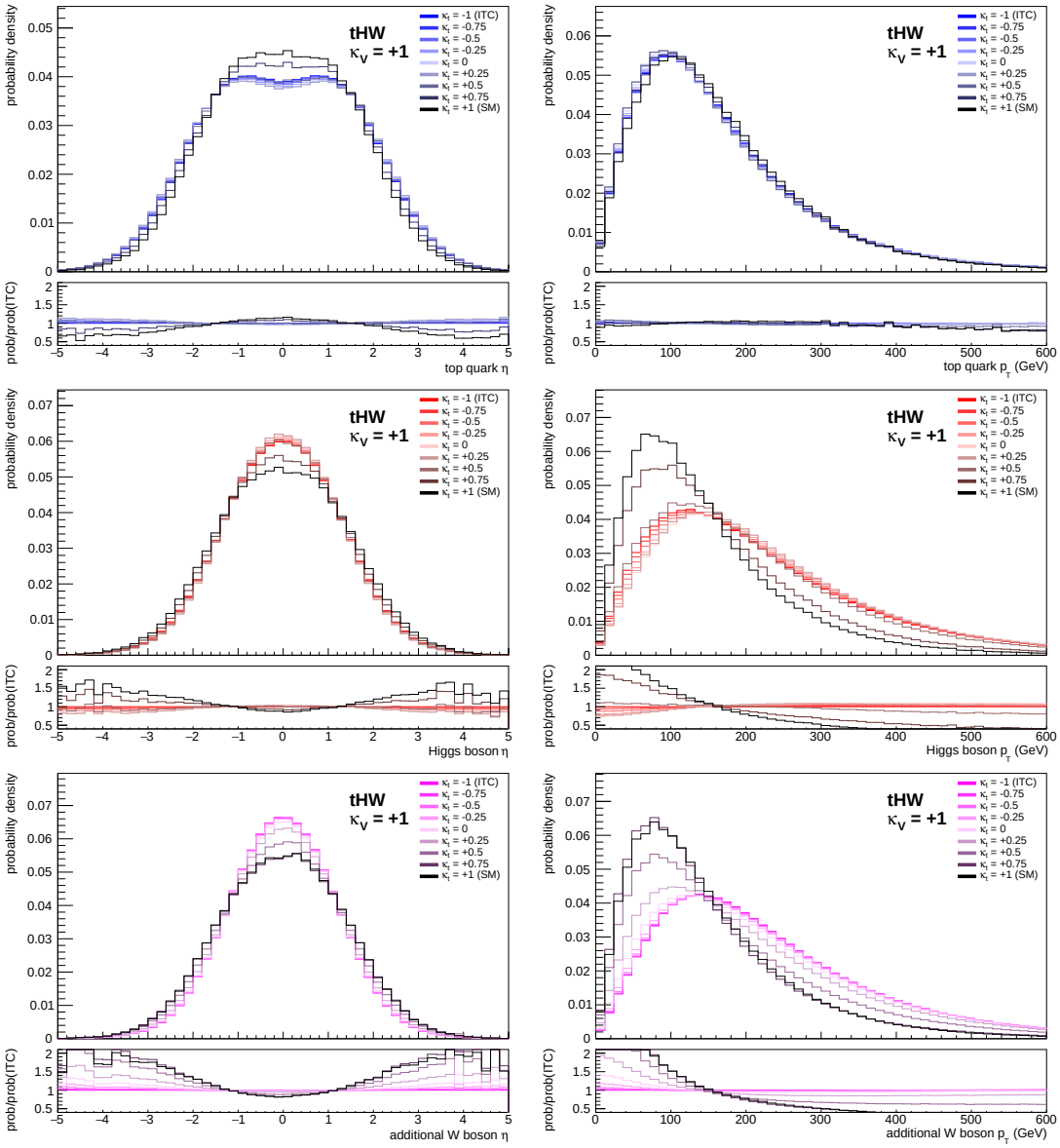

Figure A.2.: Kinematic variations of the tHW final-state particles for different coupling scenarios: the η and p_T distributions are shown for the top quark, the Higgs boson and the additional W boson.

A.3. Input Variables of the Jet Assignment

The distributions of the input variables of the tHq jet assignment are shown in Fig. A.3. The corresponding depictions for the tHW and $t\bar{t}$ jet assignment can be found in Fig. A.4 and Fig A.5, respectively. As the tHW jet assignment requires one more jet compared to the tHq and $t\bar{t}$ hypotheses, the outputs for these events are set to a default value of -999. This is visible in the figures as a peak in the leftmost bin.

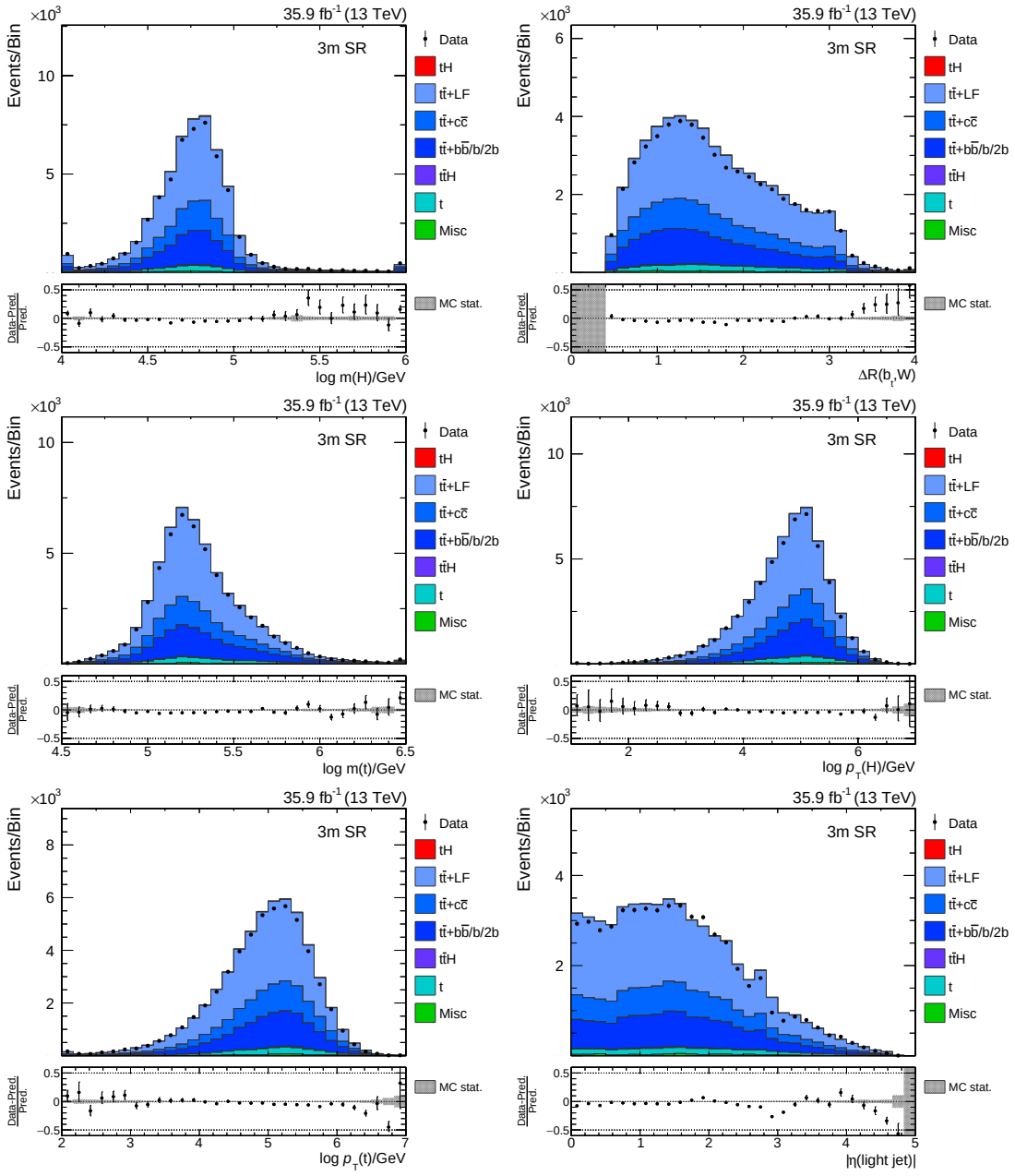

Figure A.3.: Modeling of the most important input variables of for the tHq jet assignment: the distributions show all events of the 3m region, i.e., category A and B events combined.

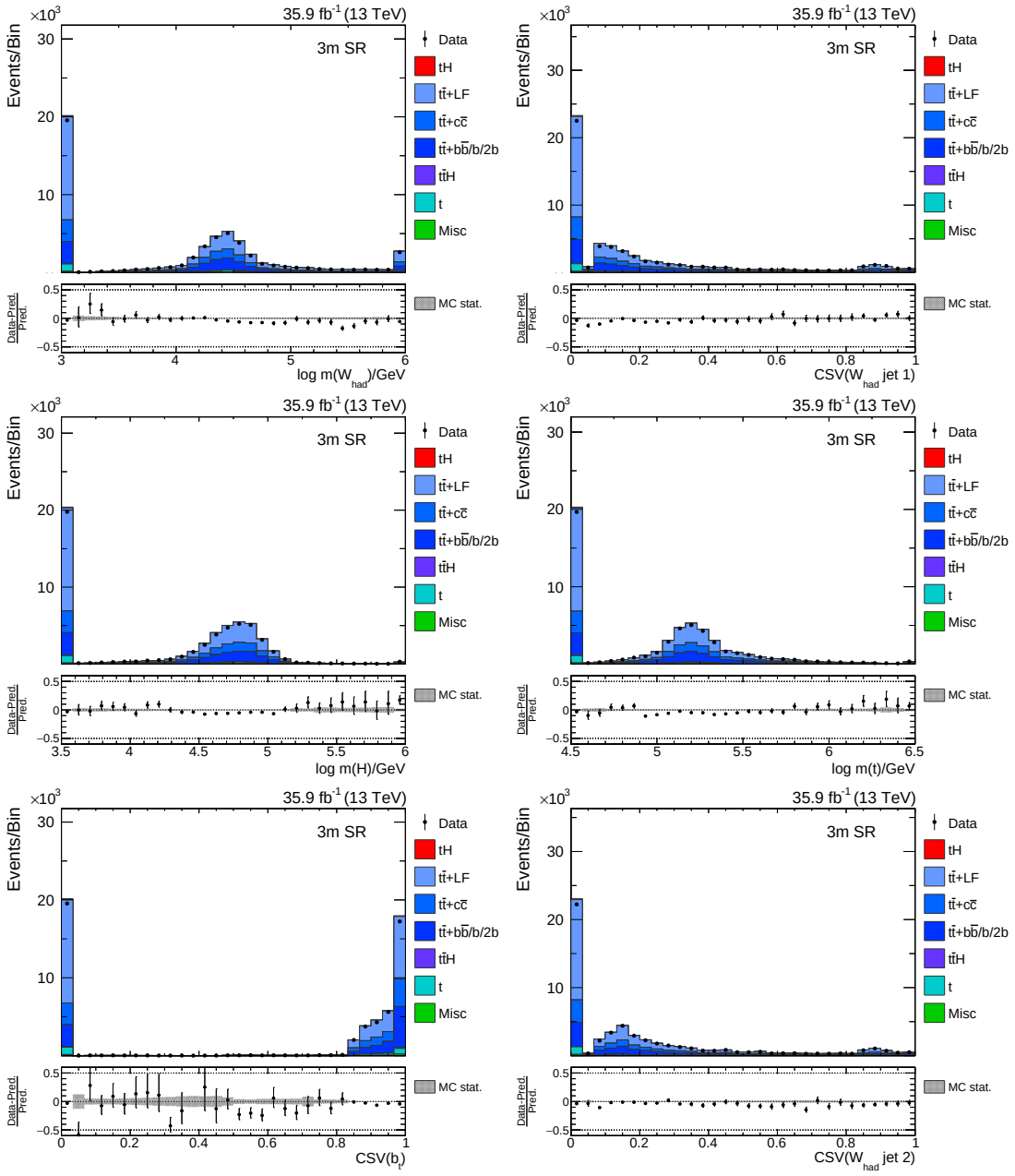

Figure A.4.: Modeling of the most important input variables of for the tHW jet assignment: the distributions show all events of the 3m region, i.e., category A and B events combined.

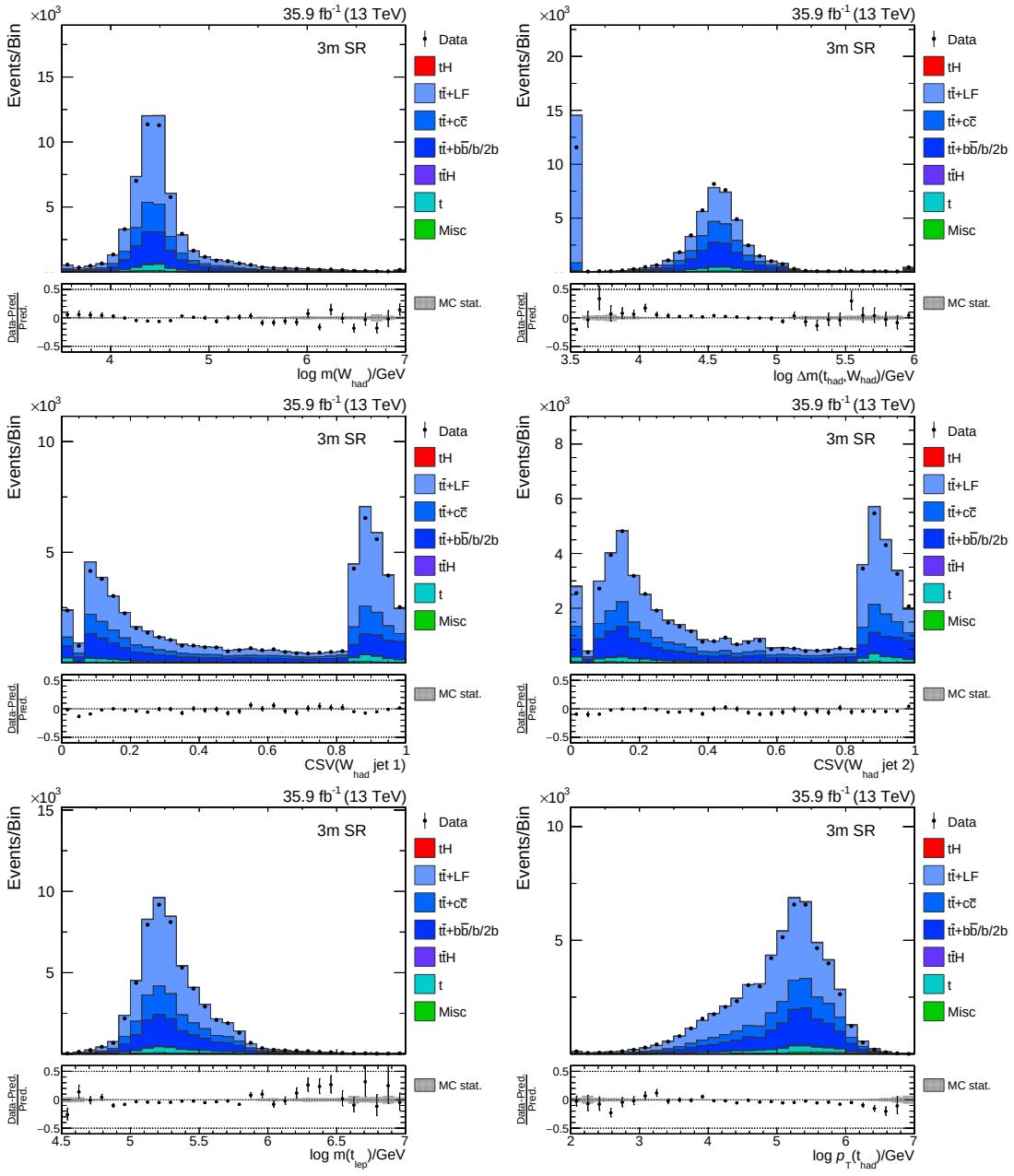

Figure A.5.: Modeling of the most important input variables of for the $t\bar{t}$ jet assignment: the distributions show all events of the 3m region, i.e., category A and B events combined.

Table A.3.: Input variables used in the tHq JA BDTs: these variables describe the properties of the final-state particles, i.e., the Higgs boson, the top quark, and the light quark. For variables with long tails, e.g., the reconstructed mass of the Higgs boson, the logarithm of the variable is used as narrower distributions are better handled by the BDT. The given importance is averaged over the trainings for category A and B events. The definition of the importance can be found in Section 3.5.

Variable	Importance	Description
log m(H)/GeV	0.18	invariant mass of the reconstructed Higgs boson
$\Delta R(b_t,W)$	0.11	ΔR between the jet assigned to the bottom quark of the top quark decay and the W boson
log m(t)/GeV	0.11	invariant mass of the reconstructed top quark
log p_T(H)/GeV	0.08	transverse momentum of the reconstructed Higgs boson
log p_T(t)/GeV	0.06	transverse momentum of the reconstructed top quark
$\lvert\eta$(**light jet**)$\rvert$	0.05	absolute pseudorapidity of the jet assigned to the light quark
CSV(b_t)	0.05	output of the b-tagging algorithm for the jet assigned to the bottom quark of the top quark decay
relative H_T	0.05	scalar sum of the transverse momentum for the jets assigned under the tHq hypothesis divided by the scalar sum of the total transverse momentum (jets, lepton, $\vec{p}_T^{\,miss}$)
ΔR(H)	0.05	ΔR between the two jets assigned to the Higgs boson decay
$\lvert\eta$(**light jet**)$-\eta(b_t)\rvert$	0.05	absolute difference in pseudorapidity of the reconstructed light quark and the reconstructed top quark
ΔR(t, H)	0.04	ΔR between reconstructed top quark and reconstructed Higgs boson
CSV(light jet)	0.04	output of the b-tagging algorithm for the jet assigned to the light quark
CSV($b_{H,2}$)	0.04	output of the b-tagging algorithm for the second jet assigned to the Higgs boson
log p_T(light jet)/GeV	0.04	transverse momentum of the jet assigned to the light quark
$\lvert\eta(b_t)\rvert$	0.03	absolute pseudorapidity of the jet assigned to the bottom quark of the top quark decay
CSV($b_{H,1}$)	0.02	output of the b-tagging algorithm for the first jet assigned to the Higgs boson

A.4. Ranking of Input Variables of the Jet Assignment

The input variables ranked by their average importance in the trainings on dataset A and B are listed in Table A.3 for the tHq jet assignment, in Table A.4 for the tHW jet assignment, and in Table A.5 for the $t\bar{t}$ jet assignment.

Table A.4.: Input variables used in the tHW JA BDTs: these variables describe the properties of the final-state particles, i.e., the Higgs boson, the top quark, and the additional W boson. In order to figure out which of the W bosons decays leptonically, also properties of the hadronically and leptonically decaying W bosons are utilized. For variables with long tails, e.g., the reconstructed mass of the Higgs boson, the logarithm of the variable is used as narrower distributions are better handled by the BDT. The given importance is averaged over the trainings for category A and B events. The definition of the importance can be found in Section 3.5.

Variable	Importance	Description
log m(W_{had})/GeV	0.11	invariant mass of the reconstructed hadronically decaying W boson
CSV(W_{had} jet 1)	0.08	output of the b-tagging algorithm for the first jet assigned to the hadronically decaying W boson
log m(H)/GeV	0.08	invariant mass of the reconstructed Higgs boson
log m(t)/GeV	0.07	invariant mass of the reconstructed top quark
CSV(b_t)	0.07	output of the b-tagging algorithm for the jet assigned to the bottom quark of the top quark decay
CSV(W_{had} jet 2)	0.06	output of the b-tagging algorithm for the second jet assigned to the hadronically decaying W boson
log p_T(H)/GeV	0.05	transverse momentum of the reconstructed Higgs boson
CSV($b_{H,2}$)	0.05	output of the b-tagging algorithm for the second jet assigned to the Higgs boson
CSV($b_{H,1}$)	0.05	output of the b-tagging algorithm for the first jet assigned to the Higgs boson
ΔR(H)	0.05	ΔR between the two jets assigned to the Higgs boson decay
log p_T(W_{had})/GeV	0.05	transverse momentum of the reconstructed hadronically decaying W boson
log p_T(W_b)/GeV	0.04	transverse momentum of the reconstructed additional W boson
cos $\theta(b_t, \ell)$	0.04	cosine of the angle between the leptonically decaying and the jet assigned to the bottom quark from the top quark decay
log m(W_b)/GeV	0.03	invariant mass of the reconstructed additional W boson
p_T(W_{lep})-p_T(W_{had})	0.03	difference in transverse momentum between reconstructed leptonically decaying and hadronically decaying W boson
ΔR(b_t,W_{lep})	0.03	ΔR between the jet assigned to the bottom quark of the top quark decay and the leptonically decaying W boson
relative H_T	0.02	scalar sum of the transverse momentum for the jets assigned under the tHW hypothesis divided by the scalar sum of the total transverse momentum (jets, lepton, $\vec{p}_T^{miss}$)
log p_T(t)/GeV	0.02	transverse momentum of the reconstructed top quark
$\lvert\eta(t)\rvert$	0.02	absolute pseudorapidity of the reconstructed top quark
$\lvert\eta(W_b)\rvert$	0.02	absolute pseudorapidity of the reconstructed additional W boson
$\lvert\eta(b_t)\rvert$	0.02	absolute pseudorapidity of the jet assigned to the bottom quark of the top quark decay
$\lvert\eta(W_{lep}) - \eta(W_{had})\rvert$	0.02	absolute difference in pseudorapidity of the reconstructed leptonically decaying and hadronically decaying W boson

Table A.5.: Input variables used in the $t\bar{t}$ JA BDTs for the analysis presented in Chapter 6: these variables are ranked by their importance in the training. The given importance is averaged over the trainings for category A and B events. The definition of the importance can be found in Section 3.5.

Variable	Importance	Description
log m(W_{had})/GeV	0.18	invariant mass of the two jets assigned to the W boson of the top quark
log Δm(t_{had},W_{had})/GeV	0.16	difference between the invariant masses of the reconstructed hadronically decaying top quark and the reconstructed hadronically decaying W boson
CSV(W_{had} jet 1)	0.16	output of the b-tagging algorithm for the first jet assigned to the hadronically decaying W boson
CSV(W_{had} jet 2)	0.14	output of the b-tagging algorithm for the second jet assigned to the hadronically decaying W boson
log m(t_{lep})/GeV	0.11	invariant mass of the reconstructed leptonically decaying top quark
log p_T(t_{had})/GeV	0.09	transverse momentum of the reconstructed hadronically decaying top quark
log p_T(t_{lep})/GeV	0.07	transverse momentum of the reconstructed leptonically decaying top quark
ΔR (W_{had})	0.06	ΔR between the two jets assigned to the W boson of the hadronically decaying top quark
relative H_T	0.04	scalar sum of the transverse momentum for the jets assigned under the $t\bar{t}$ hypothesis divided by the scalar sum of the total transverse momentum (jets, lepton, $\vec{p}_T^{miss}$)

A.5. Jet Assignment Efficiencies

The efficiencies for the jet assignment under the tHq, tHW and $t\bar{t}$ hypothesis are shown in Table A.6, A.7, and A.8, respectively.

Table A.6.: tHq jet assignment efficiencies: the table lists the efficiencies for the individual particles and the efficiency for assigning all jets correctly for a BDT-based and a χ^2-based approach.

	BDT efficiency	**χ^2 efficiency**
Higgs boson	71.9%	64.7%
Top quark	73.4%	64.3%
Light quark	91.8%	89.1%
Top quark + Higgs boson	65.8%	55.8%
Total	61.5%	51.3%

Table A.7.: tHW jet assignment efficiencies: the table lists the efficiencies for the individual particles and the efficiency for assigning all jets correctly for a BDT-based and a χ^2-based approach.

	BDT efficiency	**χ^2 efficiency**
Higgs boson	73.0%	50.1%
Bottom quark from top quark decay	72.0%	43.1%
Correctly assigning W bosons to decays	74.8%	71.4%
Hadronically decaying W boson	73.0%	44.9%
Top quark	54.7%	33.5%
Total	46.1%	24.5%

Table A.8.: $t\bar{t}$ jet assignment efficiencies: the table lists the efficiencies for the individual particles and the efficiency for assigning all jets correctly for a BDT-based and a χ^2-based approach.

	BDT efficiency	**χ^2 efficiency**
Bottom quark from leptonic top quark decay	68.8%	59.8%
Bottom quark from hadronic top quark decay	68.4%	50.0%
Hadronically decaying W boson	70.5%	49.2%
Hadronically decaying top quark	61.0%	41.5%
Total	51.1%	33.2%

A.6. Input Variables for the Flavor Classification BDT

The distributions of the input variables of the FC BDT are shown in Fig. A.6.

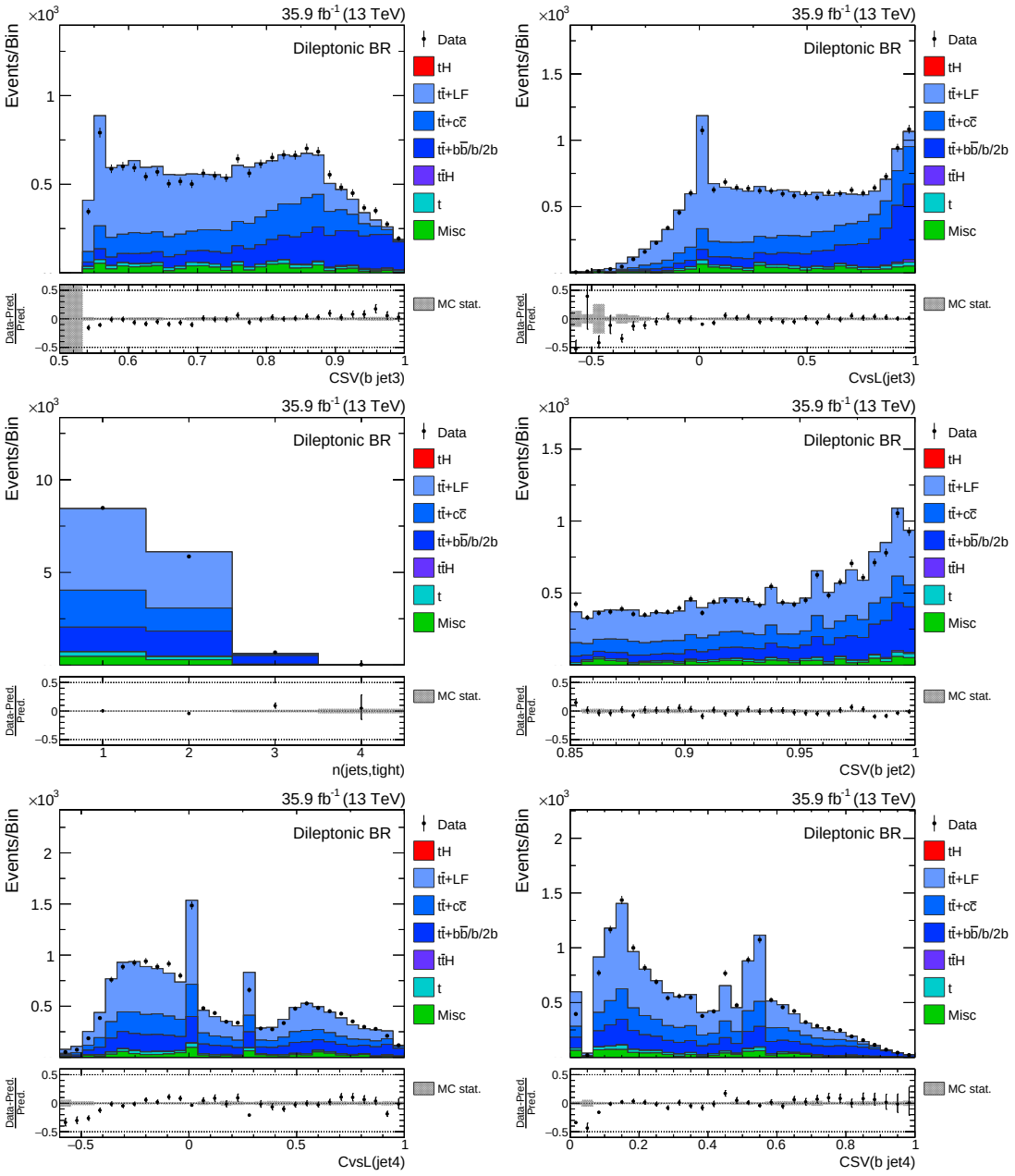

Figure A.6.: Modeling of the most important input variables of for the FC BDT: the distributions show all events in the dileptonic $t\bar{t}$ region, i.e., category A and B events combined.

A.7. Upper Limits on Cross Section Times Branching Ratio

The upper limits for each of the studied points are listed in Table A.9 for the scan of the coupling ratio and in Table A.10 for the CP-mixing scan.

Table A.9.: Upper limits on the combined tH+t$\bar{t}$H cross section times branching ratio: the limits are calculated with the reweighted tH templates, the cross section of the t$\bar{t}$H process is scaled depending on κ_t.

κ_t/κ_V	observed	expected	±1 std. dev.	±2 std. dev.
-6.00	2.75	1.86	[0.95, 1.30]	[2.70, 3.82]
-4.00	2.67	1.80	[0.91, 1.25]	[2.62, 3.72]
-3.00	2.60	1.75	[0.89, 1.22]	[2.56, 3.64]
-2.50	2.55	1.71	[0.87, 1.19]	[2.50, 3.58]
-2.00	2.48	1.66	[0.84, 1.15]	[2.43, 3.49]
-1.50	2.36	1.59	[0.79, 1.10]	[2.33, 3.37]
-1.33	2.33	1.55	[0.77, 1.07]	[2.28, 3.31]
-1.25	2.29	1.54	[0.77, 1.06]	[2.26, 3.27]
-1.00	2.21	1.47	[0.74, 1.01]	[2.17, 3.17]
-0.83	2.12	1.42	[0.71, 0.98]	[2.10, 3.07]
-0.75	2.08	1.40	[0.69, 0.96]	[2.07, 3.03]
-0.67	2.04	1.37	[0.68, 0.94]	[2.03, 2.97]
-0.50	1.96	1.31	[0.64, 0.89]	[1.94, 2.86]
-0.33	1.85	1.25	[0.62, 0.86]	[1.86, 2.74]
-0.25	1.81	1.22	[0.60, 0.83]	[1.82, 2.70]
-0.17	1.79	1.20	[0.59, 0.82]	[1.80, 2.66]
0.00	1.78	1.20	[0.59, 0.82]	[1.80, 2.66]
0.17	1.93	1.30	[0.64, 0.89]	[1.94, 2.86]
0.25	2.10	1.40	[0.69, 0.95]	[2.09, 3.08]
0.33	2.35	1.55	[0.76, 1.06]	[2.32, 3.40]
0.50	3.01	1.98	[1.00, 1.36]	[2.91, 4.14]
0.67	3.42	2.32	[1.20, 1.63]	[3.34, 4.66]
0.75	3.45	2.40	[1.27, 1.70]	[3.42, 4.74]
0.83	3.42	2.43	[1.29, 1.73]	[3.45, 4.76]
1.00	3.31	2.43	[1.31, 1.74]	[3.44, 4.72]
1.25	3.21	2.38	[1.29, 1.71]	[3.36, 4.62]
1.33	3.19	2.36	[1.28, 1.70]	[3.35, 4.60]
1.50	3.17	2.35	[1.27, 1.68]	[3.33, 4.57]
2.00	3.14	2.29	[1.21, 1.62]	[3.24, 4.45]
2.50	3.13	2.24	[1.19, 1.59]	[3.19, 4.39]
3.00	3.10	2.20	[1.17, 1.56]	[3.14, 4.35]
4.00	3.08	2.14	[1.13, 1.52]	[3.07, 4.25]
6.00	3.03	2.09	[1.09, 1.47]	[3.00, 4.16]

Table A.10.: Upper limits on the combined tH+t$\bar{t}$H cross section times branching ratio: the limits are calculated with the reweighted tH and t$\bar{t}$H templates.

cos(α)	observed	expected	$\pm$1 std. dev.	$\pm$2 std. dev.
-1.0	2.00	1.63	[0.89,1.18]	[2.27,3.01]
-0.9	1.97	1.49	[0.81,1.07]	[2.06,2.74]
-0.8	1.92	1.51	[0.82,1.09]	[2.10,2.79]
-0.7	1.88	1.50	[0.81,1.08]	[2.08,2.76]
-0.6	1.82	1.43	[0.77,1.03]	[1.98,2.63]
-0.5	1.84	1.39	[0.75,1.00]	[1.93,2.56]
-0.4	1.79	1.37	[0.74,0.99]	[1.90,2.52]
-0.3	1.76	1.34	[0.73,0.97]	[1.86,2.48]
-0.2	1.75	1.33	[0.72,0.96]	[1.85,2.45]
-0.1	1.76	1.34	[0.73,0.97]	[1.87,2.48]
0.0	1.79	1.36	[0.74,0.98]	[1.89,2.51]
0.1	1.89	1.43	[0.78,1.03]	[1.99,2.64]
0.2	2.03	1.53	[0.83,1.10]	[2.12,2.82]
0.3	2.22	1.69	[0.91,1.22]	[2.34,3.11]
0.4	2.42	1.92	[1.04,1.38]	[2.66,3.53]
0.5	2.69	2.21	[1.20,1.59]	[3.07,4.08]
0.6	3.24	2.51	[1.36,1.81]	[3.48,4.63]
0.7	3.67	2.84	[1.54,2.05]	[3.95,5.25]
0.8	4.02	3.20	[1.73,2.31]	[4.44,5.90]
0.9	4.02	3.28	[1.78,2.36]	[4.55,6.05]
1.0	4.21	3.25	[1.76,2.34]	[4.51,5.99]

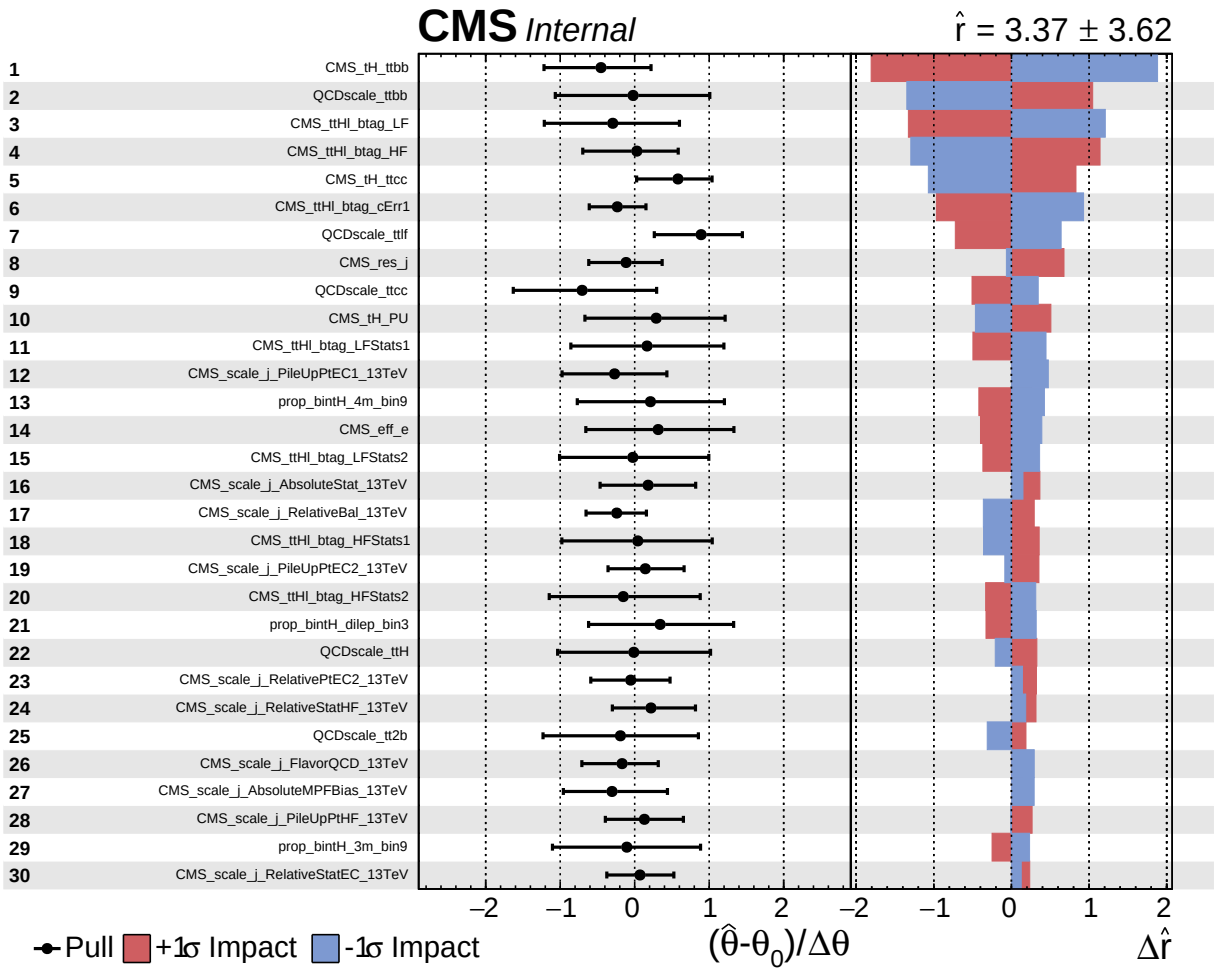

Figure A.7.: Impacts of the systematic uncertainties for the SM scenario: the uncertainties are ranked by their impact on the signal strength modifier.

A.8. Systematic Impacts

The impacts of the systematic uncertainties on the signal strength modifier can be found in Fig. A.7 for the SM case and in Fig. A.8 for the ITC scenario.

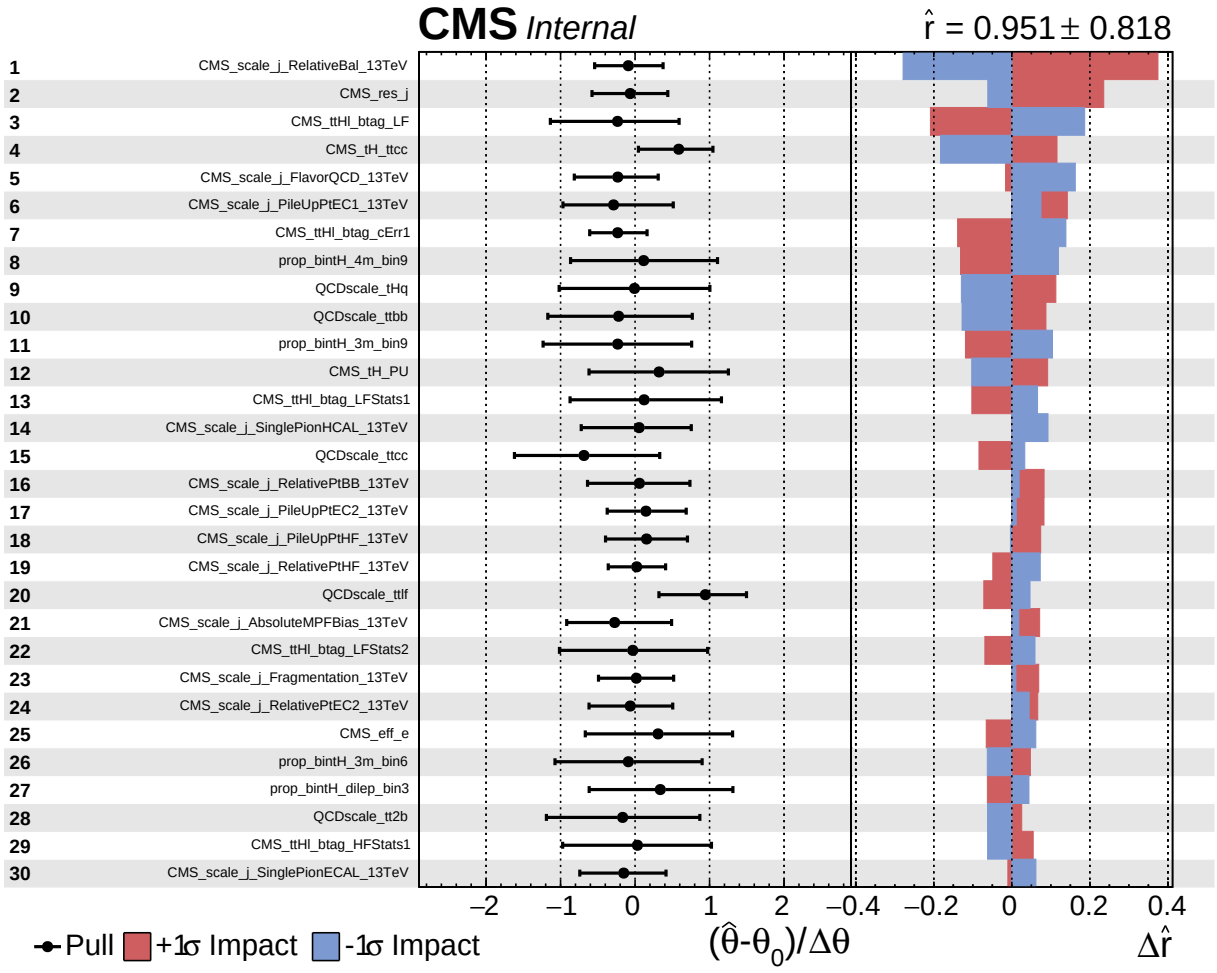

Figure A.8.: Impacts of the systematic uncertainties for the ITC scenario: the uncertainties are ranked by their impact on the signal strength modifier.

B. Combined tH and $t\bar{t}H$ analysis

This appendix shows tables and figures providing additional information to the analysis presented in Chapter 7.

B.1. Input Variables of the Jet Assignment

This section shows figures related to the jet assignment under the tHq, tHW, $t\bar{t}H$ and $t\bar{t}$ hypotheses. The distributions of the most important variables to the jet assignment under the tHq hypothesis can be found in Fig. B.1 for 2016 and Fig. B.5 for 2018. The corresponding distributions for the tHW assignment can be found in Fig. B.2 and Fig. B.6, respectively. Figure B.3 and Fig. B.7 show the distributions for the $t\bar{t}H$ assignment and Fig. B.4 and B.8 depict the variables for the $t\bar{t}$ hypothesis.

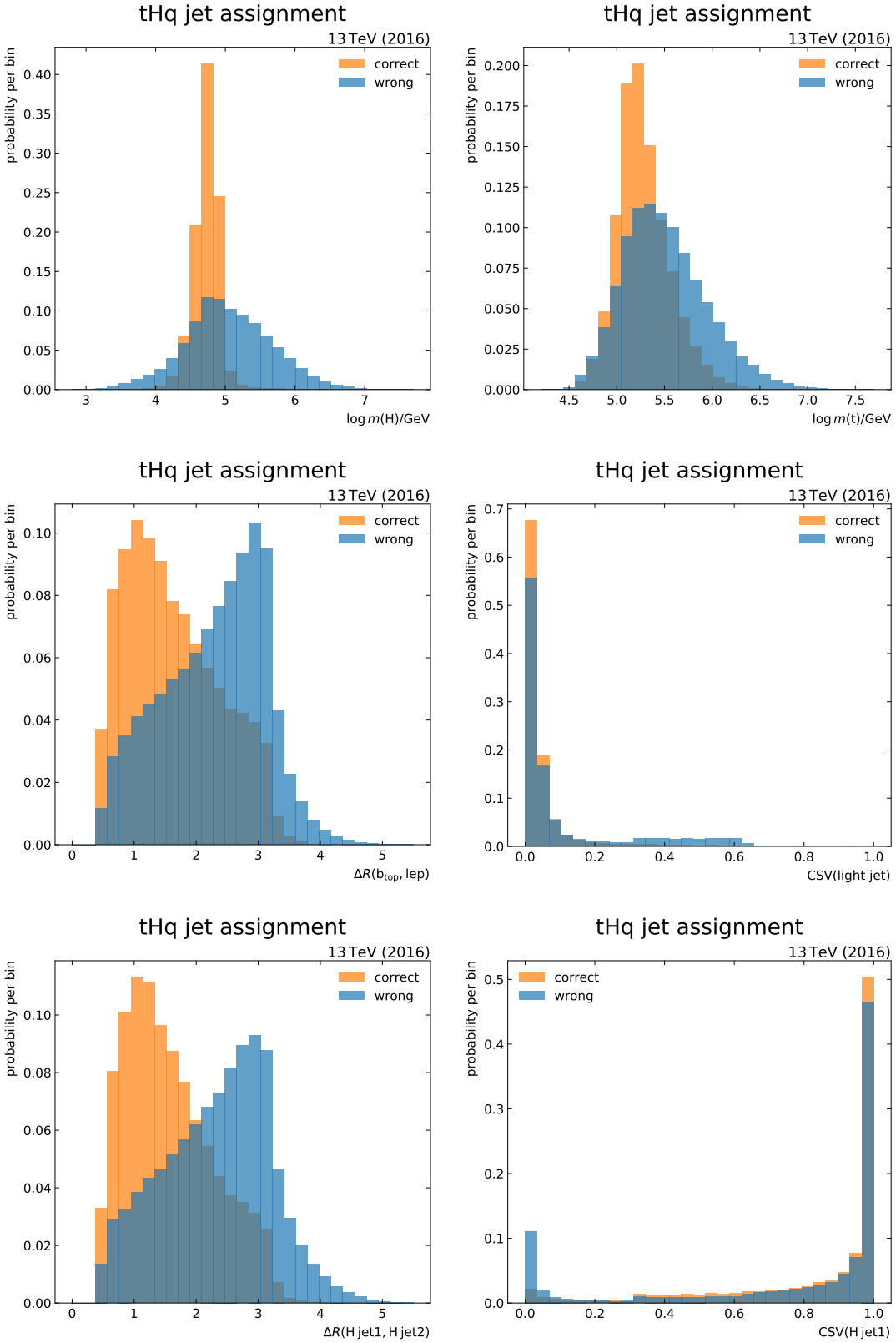

Figure B.1.: Distributions for the most important variables of the tHq jet assignment (2016): the distributions are shown for correct and wrong assignments for category A and B events combined.

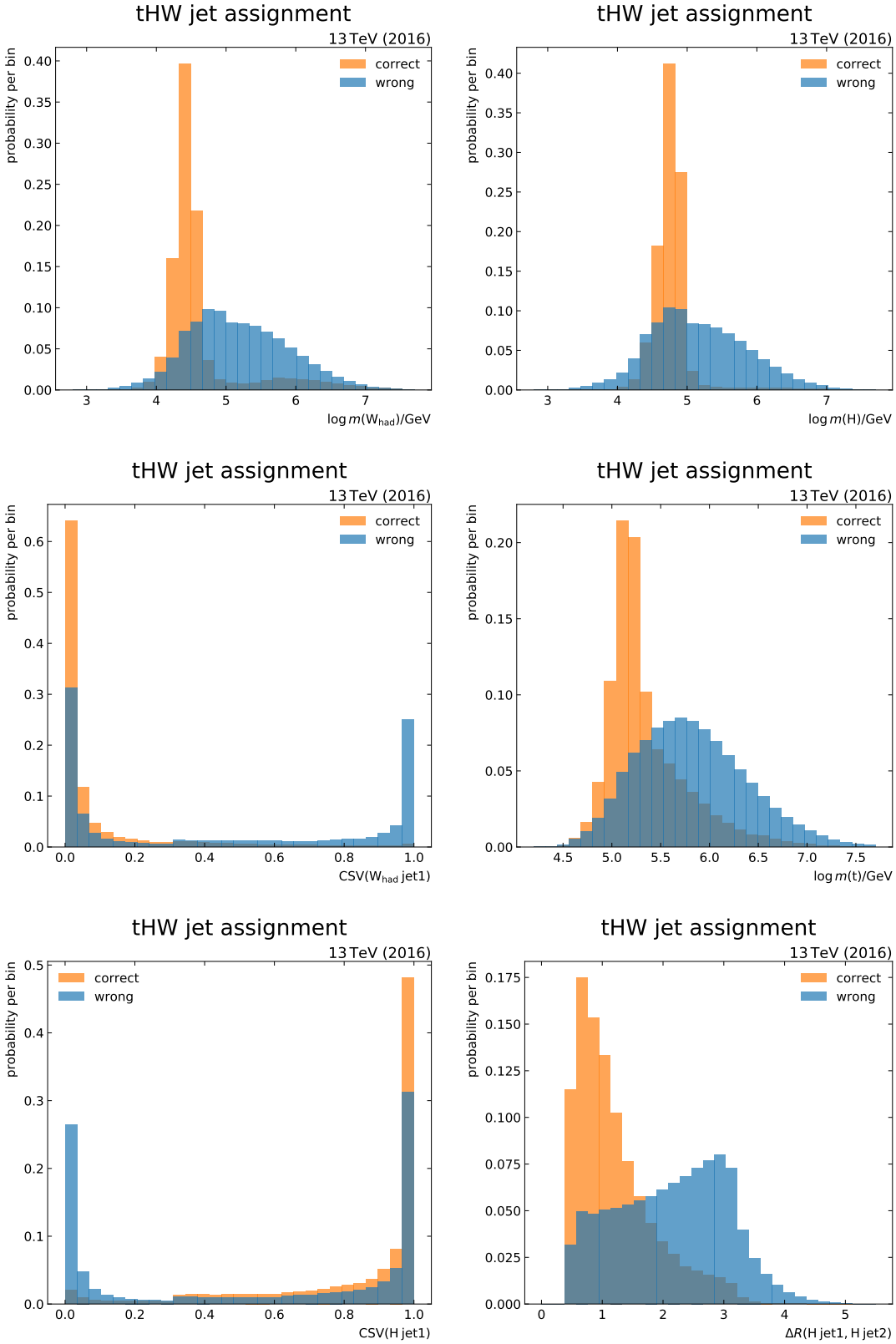

Figure B.2.: Distributions for the most important variables of the tHW jet assignment (2016): the distributions are shown for correct and wrong assignments for category A and B events combined.

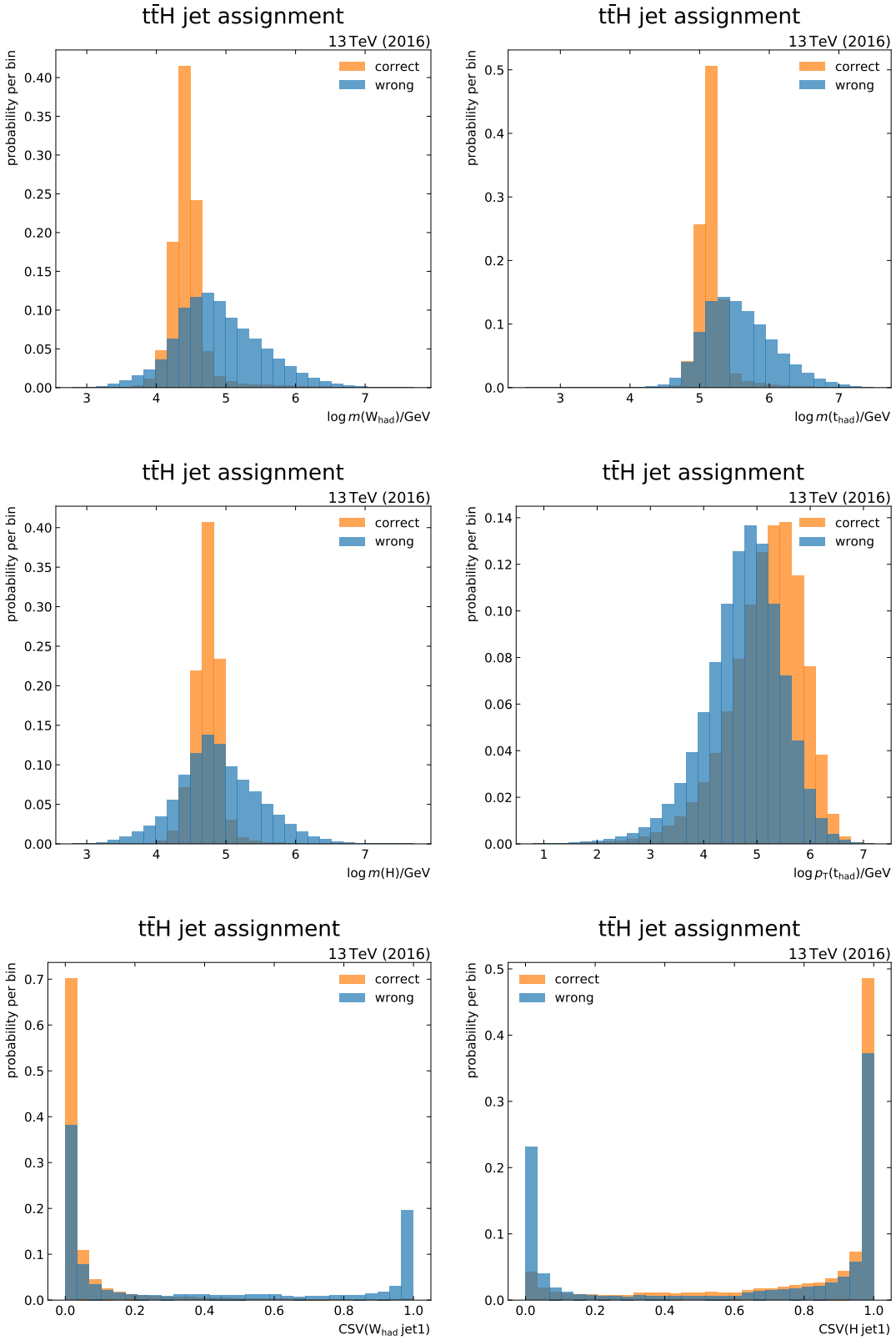

Figure B.3.: Distributions for the most important variables of the tt̄H jet assignment (2016): the distributions are shown for correct and wrong assignments for category A and B events combined.

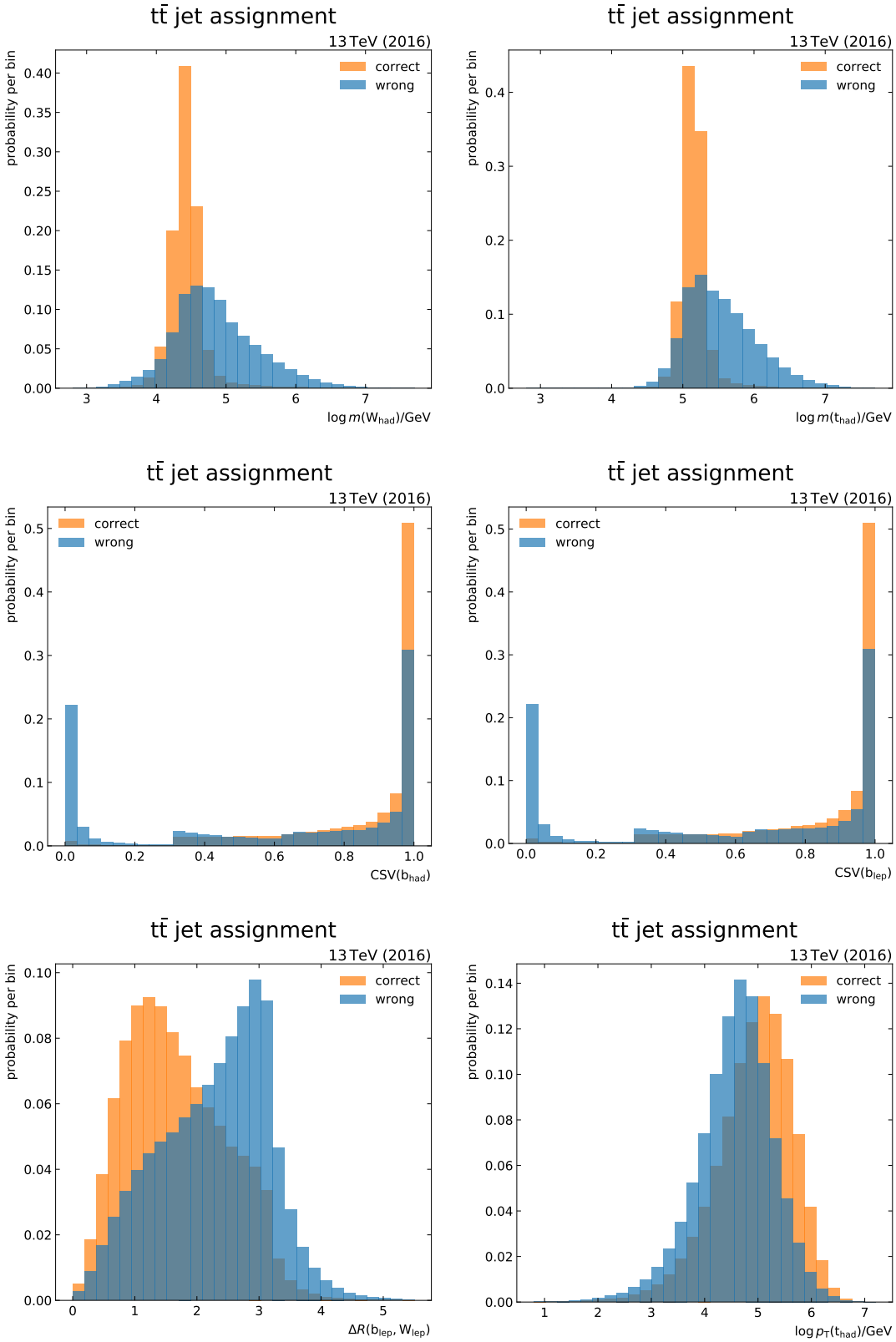

Figure B.4.: Distributions for the most important variables of the $t\bar{t}$ jet assignment (2016): the distributions are shown for correct and wrong assignments for category A and B events combined. The CSV output distributions for the jets assigned to the bottom quarks of the top quark decays are very similar. Nonetheless, both variables provide valuable input to the jet assignment.

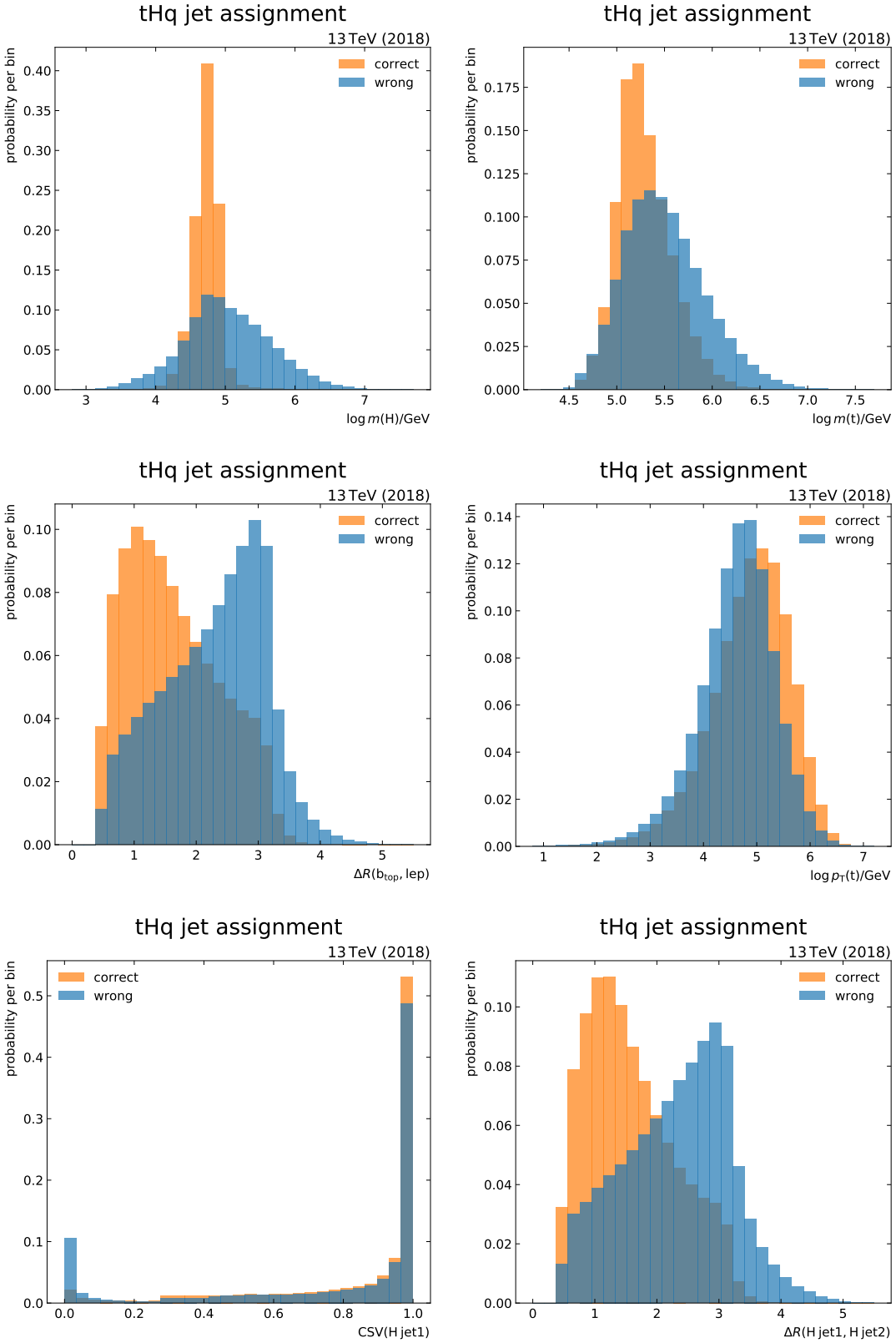

Figure B.5.: Distributions for the most important variables of the tHq jet assignment (2018): the distributions are shown for correct and wrong assignments for category A and B events combined.

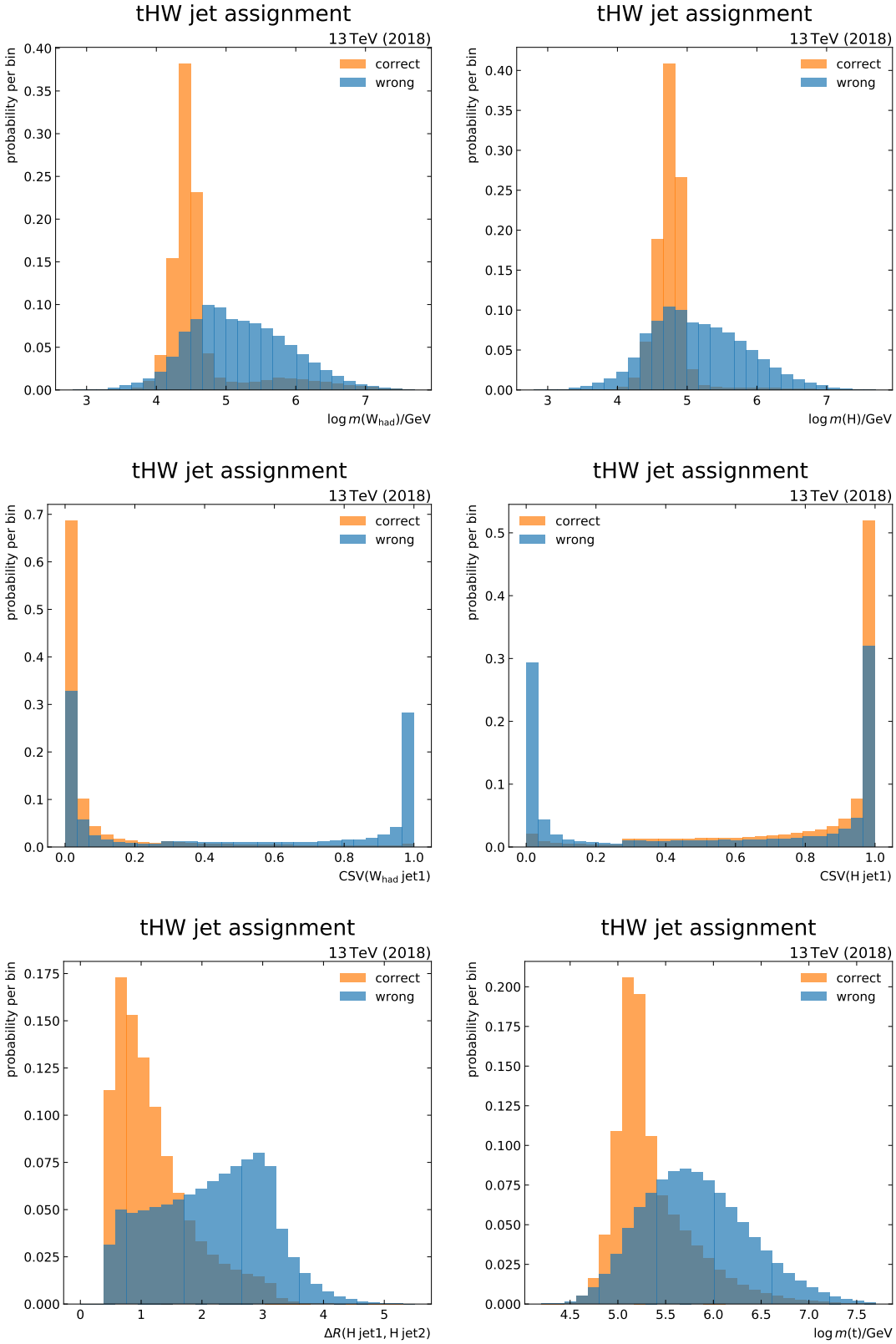

Figure B.6.: Distributions for the most important variables of the tHW jet assignment (2018): the distributions are shown for correct and wrong assignments for category A and B events combined.

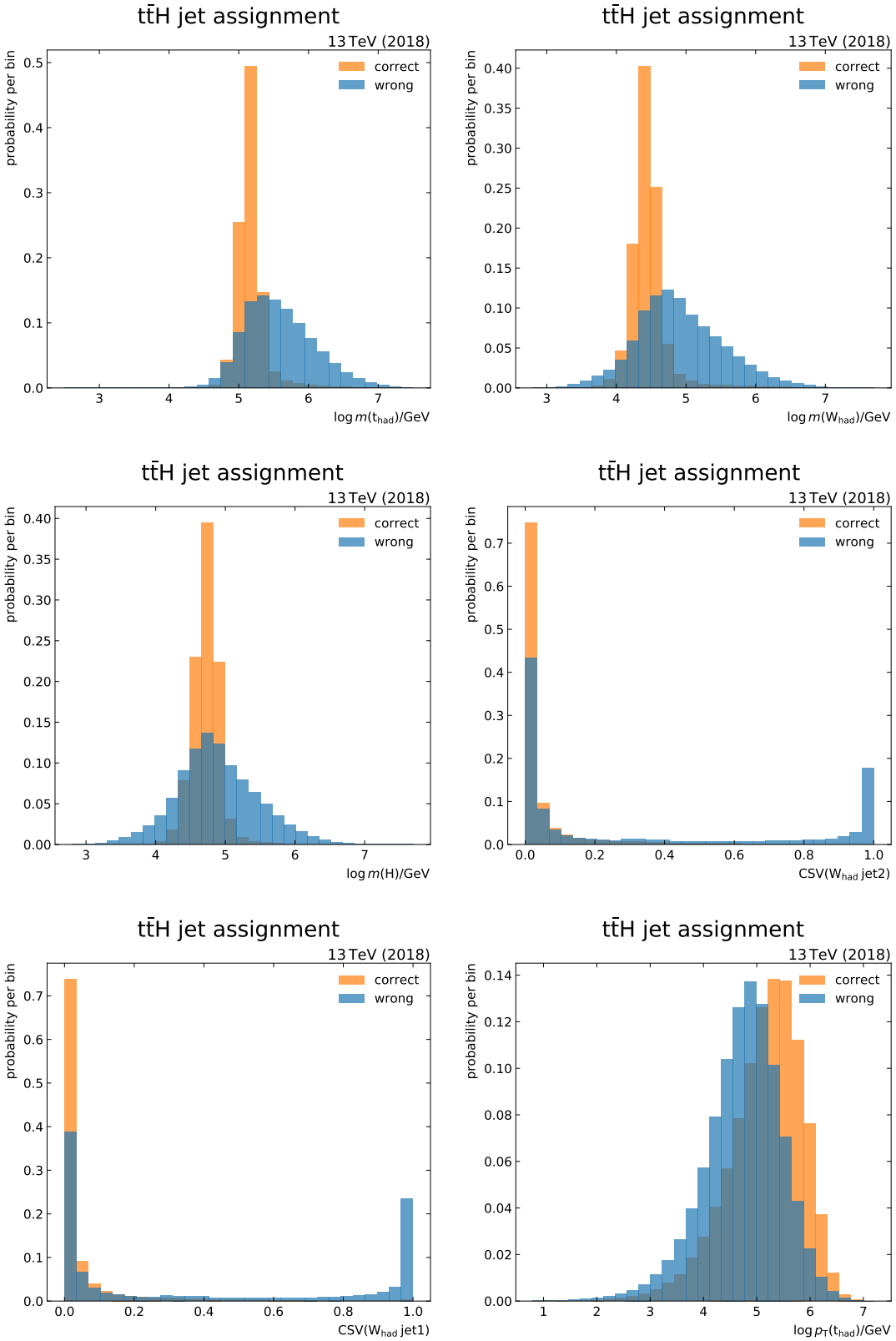

Figure B.7.: Distributions for the most important variables of the tt̄H jet assignment (2018): the distributions are shown for correct and wrong assignments for category A and B events combined.

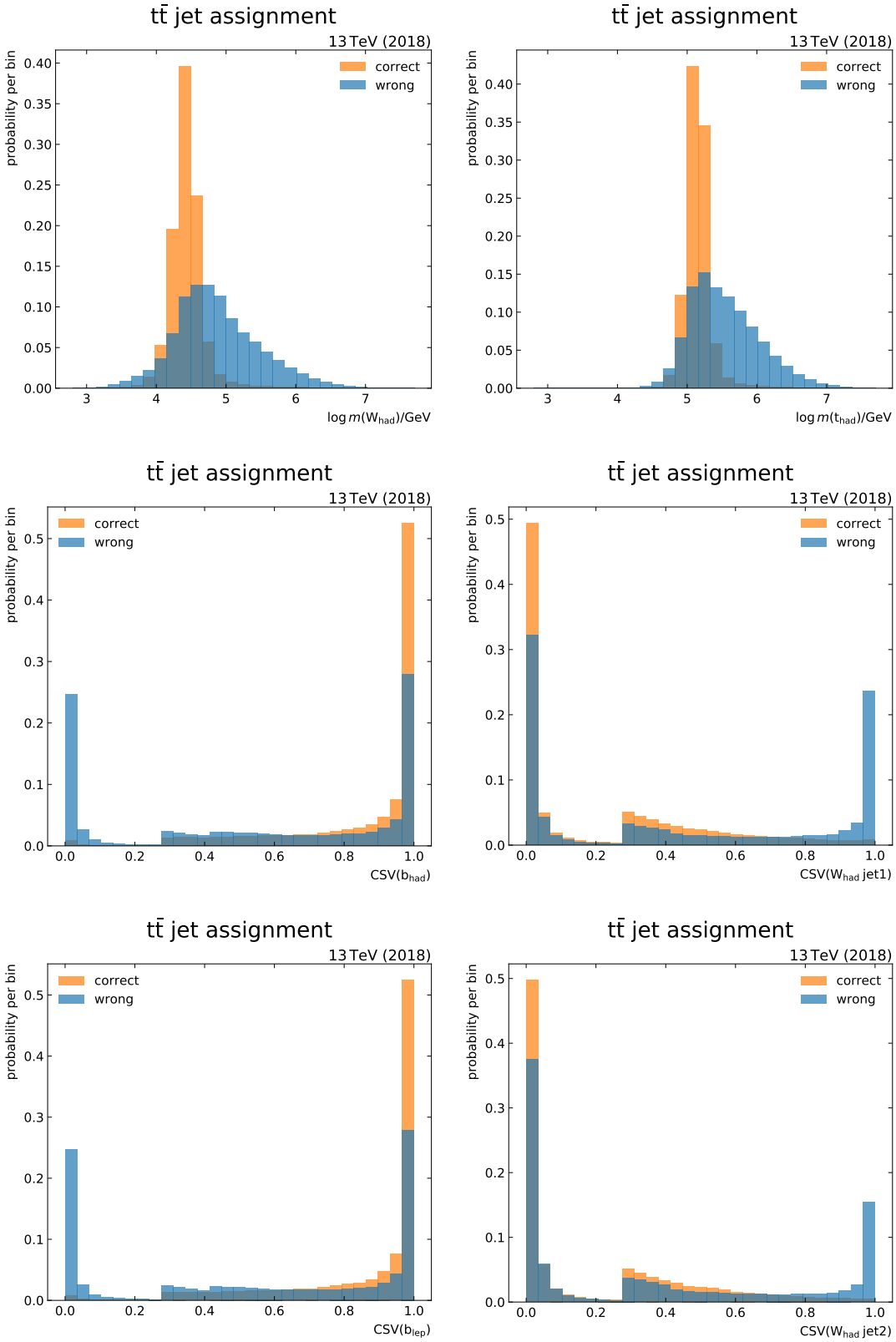

Figure B.8.: Distributions for the most important variables of the $t\bar{t}$ jet assignment (2018): the distributions are shown for correct and wrong assignments for category A and B events combined. The CSV output distributions for the jets assigned to the bottom quarks of the top quark decays are very similar. Nonetheless, both variables provide valuable input to the jet assignment.

Table B.1.: Input variables used in the tHq JA BDTs: these variables describe the properties of the final-state particles, i.e., the Higgs boson, the top quark, and the light quark. For variables with long tails, e.g., the reconstructed mass of the Higgs boson, the logarithm of the variable is used as narrower distributions are better handled by the BDT. The importance is given as the average over all three years and the training in categories A and B. The definition of the importance can be found in Section 3.5.

Variable	Importance	Description
$\log\, m(\mathrm{H})/\mathrm{GeV}$	0.146	logarithm of the invariant mass of the Higgs boson
$\log\, m(\mathrm{t})/\mathrm{GeV}$	0.090	logarithm of invariant mass of the reconstructed top quark
$\Delta R(\mathrm{b_{top}}, \mathrm{lep})$	0.084	ΔR between the bottom quark from the top quark decay and the lepton
CSV(H jet1)	0.064	output of the CSV b-tagging algorithm for the first jet assigned to the Higgs boson
$\Delta R(\mathrm{H\,jet1}, \mathrm{H\,jet2})$	0.060	ΔR between the two jets from the Higgs boson decay
$\log\, p_{\mathrm{T}}(\mathrm{t})/\mathrm{GeV}$	0.060	logarithm of transverse momentum of the top quark
$\log\, p_{\mathrm{T}}(\mathrm{H})/\mathrm{GeV}$	0.058	logarithm of transverse momentum of the Higgs boson
CSV(btop)	0.056	output of the CSV b-tagging algorithm for the jet assigned to the bottom quark from the top quark decay
CSV(H jet2)	0.052	output of the CSV b-tagging algorithm for the second jet assigned to the Higgs boson
CSV(light jet)	0.051	output of the CSV b-tagging algorithm for the jet assigned to the light quark
$\lvert\eta(\text{light jet})\rvert$	0.048	absolute pseudorapidity of the light quark
$\Delta R(\mathrm{b_{top}}, \mathrm{lep})$	0.045	ΔR between the bottom quark from the top quark decay and the leptonic W boson
$\lvert\eta(\mathrm{b_{top}})\rvert$	0.041	absolute pseudorapidity of the bottom quark from the top quark decay
$\lvert\eta(\text{light jet}) - \eta(\mathrm{b_{top}})\rvert$	0.040	absolute value of the difference between the light quark and the bottom quark from the top quark decay
$\log\, p_{\mathrm{T}}(\text{light jet})/\mathrm{GeV}$	0.039	logarithm of transverse momentum of the light quark
$\frac{p_{\mathrm{T}}(\mathrm{t})+p_{\mathrm{T}}(\mathrm{H})+p_{\mathrm{T}}(\text{lightjet})}{H_{\mathrm{T}}+p_{\mathrm{T}}^{\mathrm{miss}}+p_{\mathrm{T}}(\mathrm{lep})}$	0.038	transverse momentum fraction assigned to event compared to sum of all transverse momenta
$\Delta R(\mathrm{t}, \mathrm{H})$	0.031	ΔR between the reconstructed top quark and the Higgs boson

B.2. Ranking of Input Variables to the Jet Assignment

The variables used for the jet assignments under the tHq, tHW, $t\bar{t}H$ and $t\bar{t}$ hypothesis are listed in Table B.1, B.2, B.3, and B.4, respectively.

Table B.2.: Input variables used in the tHW JA BDTs: these variables basically describe the properties of the final-state particles, i.e., the Higgs boson, the top quark, and the additional W boson. In order to figure out which of the W bosons decays leptonically, also properties of the hadronically and leptonically decaying W bosons are utilized. For variables with long tails, e.g., the reconstructed mass of the Higgs boson, the logarithm of the variable is used as narrower distributions are better handled by the BDT. The importance is the average importance for the trainings on category A and B. The definition of the importance can be found in Section 3.5.

Variable	Importance	Description
$\log\, m(\mathrm{W_{had}})/\mathrm{GeV}$	0.103	logarithm of invariant mass of the hadronically decaying W boson
$\log\, m(\mathrm{H})/\mathrm{GeV}$	0.087	logarithm of the invariant mass of the Higgs boson
$\mathrm{CSV}(\mathrm{W_{had}\,jet1})$	0.074	output of the CSV b-tagging algorithm for the first jet assigned to the hadronically decaying W boson
$\log\, m(\mathrm{t})/\mathrm{GeV}$	0.072	logarithm of invariant mass of the reconstructed top quark
$\mathrm{CSV}(\mathrm{H\,jet1})$	0.069	output of the CSV b-tagging algorithm for the first jet assigned to the Higgs boson
$\mathrm{CSV}(\mathrm{W_{had}\,jet2})$	0.065	output of the CSV b-tagging algorithm for the second jet assigned to the hadronically decaying W boson
$\Delta R(\mathrm{H\,jet1}, \mathrm{H\,jet2})$	0.064	ΔR between the two jets from the Higgs boson decay
$\mathrm{CSV}(\mathrm{b_{top}})$	0.063	output of the CSV b-tagging algorithm for the jet assigned to the bottom quark from the top quark decay
$\mathrm{CSV}(\mathrm{H\,jet2})$	0.062	output of the CSV b-tagging algorithm for the second jet assigned to the Higgs boson
$\log\, p_\mathrm{T}(\mathrm{W_{had}})/\mathrm{GeV}$	0.050	logarithm of transverse momentum of the hadronically decaying W boson
$\log\, p_\mathrm{T}(\mathrm{H})/\mathrm{GeV}$	0.050	logarithm of transverse momentum of the Higgs boson
$\cos\theta(\mathrm{btop}, \mathrm{lep})$	0.044	cosine of the angle between the bottom quark from the top quark decay and the lepton
$\Delta R(\mathrm{W_{had}\,jet1}, \mathrm{W_{had}\,jet2})$	0.042	ΔR between the two jets from the hadronically decaying W boson decay
leptonic top quark decay	0.041	type of top quark decay
$\frac{p_\mathrm{T}(\mathrm{t})+p_\mathrm{T}(\mathrm{H})+p_\mathrm{T}(\mathrm{W_b})}{H_\mathrm{T}+p_\mathrm{T}^\mathrm{miss}+p_\mathrm{T}(\mathrm{lep})}$	0.035	transverse momentum fraction assigned to event compared to sum of all transverse momenta
$(p_\mathrm{T}(\mathrm{W_{lep}}) - p_\mathrm{T}(\mathrm{W_{had}}))/\mathrm{GeV}$	0.030	difference between the transverse momentum of the leptonically and hadronically decaying W boson
$\Delta R(\mathrm{top}, \mathrm{H})$	0.027	ΔR between the reconstructed top quark and the reconstructed Higgs boson
$\log\, p_\mathrm{T}(\mathrm{t})/\mathrm{GeV}$	0.023	logarithm of transverse momentum of the top quark

Table B.3.: Input variables used in the $t\bar{t}H$ JA BDTs for the analysis presented in Chapter 6: these variables are ranked by their importance in the training. The importance is given by the average of the training on category A and category B. The definition of the importance can be found in Section 3.5.

Variable	Importance	Description
$\log m(\mathrm{W_{had}})/\mathrm{GeV}$	0.091	logarithm of invariant mass of the hadronically decaying W boson
$\log m(\mathrm{H})/\mathrm{GeV}$	0.087	logarithm of the invariant mass of the Higgs boson
$\log\left[m(\mathrm{t_{had}}) - m(\mathrm{W_{had}})\right]/\mathrm{GeV}$	0.082	logarithm of the difference between invariant mass of the hadronically decaying top quark and hadronically decaying W boson
$\log p_{\mathrm{T}}(\mathrm{t_{had}})/\mathrm{GeV}$	0.075	logarithm of transverse momentum of the hadronically decaying top quark
CSV($\mathrm{W_{had}}$ jet1)	0.075	output of the CSV b-tagging algorithm for the first jet assigned to the hadronically decaying W boson
CSV($\mathrm{W_{had}}$ jet2)	0.073	output of the CSV b-tagging algorithm for the second jet assigned to the hadronically decaying W boson
CSV($\mathrm{b_{lep}}$)	0.068	output of the CSV b-tagging algorithm for the jet assigned to the bottom quark from the leptonic top quark decay
CSV(H jet2)	0.065	output of the CSV b-tagging algorithm for the second jet assigned to the Higgs boson
CSV($\mathrm{b_{had}}$)	0.065	output of the CSV b-tagging algorithm for the jet assigned to the bottom quark from the hadronic top quark decay
CSV(H jet1)	0.062	output of the CSV b-tagging algorithm for the first jet assigned to the Higgs boson
$\log m(\mathrm{t_{lep}})/\mathrm{GeV}$	0.057	logarithm of invariant mass of the reconstructed leptonic top quark
$\frac{p_{\mathrm{T}}(\mathrm{t_{had}})+p_{\mathrm{T}}(\mathrm{t_{lep}})+p_{\mathrm{T}}(\mathrm{H})}{H_{\mathrm{T}}+p_{\mathrm{T}}^{\mathrm{miss}}+p_{\mathrm{T}}(\mathrm{lep})}$	0.056	transverse momentum fraction assigned to event compared to sum of all transverse momenta
$\Delta R(\mathrm{H\,jet1}, \mathrm{H\,jet2})$	0.039	ΔR between the two jets from the Higgs boson decay
$\log p_{\mathrm{T}}(\mathrm{t_{lep}})/\mathrm{GeV}$	0.039	logarithm of transverse momentum of the leptonic top quark
$\Delta R(\mathrm{W_{had}\,jet1}, \mathrm{W_{had}\,jet2})$	0.036	ΔR between the two jets from the hadronic W boson decay
$\log p_{\mathrm{T}}(\mathrm{H})/\mathrm{GeV}$	0.030	logarithm of transverse momentum of the Higgs boson

Table B.4.: Input variables used in the $t\bar{t}$ JA BDTs for the analysis presented in Chapter 6: these variables are ranked by their importance in the training. The importance is given by the average of the training on category A and category B. The definition of the importance can be found in Section 3.5.

Variable	Importance	Description
$\log\, m(\mathrm{W_{had}})/\mathrm{GeV}$	0.129	logarithm of invariant mass of the hadronically decaying W boson
$\log\, m(\mathrm{t_{had}})/\mathrm{GeV}$	0.126	logarithm of invariant mass of the hadronically decaying top quark
$\mathrm{CSV}(\mathrm{b_{had}})$	0.100	output of the CSV b-tagging algorithm for the jet assigned to the bottom quark from the hadronic top quark decay
$\mathrm{CSV}(\mathrm{b_{lep}})$	0.092	output of the CSV b-tagging algorithm for the jet assigned to the bottom quark from the leptonic top quark decay
$\mathrm{CSV}(\mathrm{W_{had}\ jet1})$	0.082	output of the CSV b-tagging algorithm for the first jet assigned to the hadronically decaying W boson
$\mathrm{CSV}(\mathrm{W_{had}\ jet2})$	0.075	output of the CSV b-tagging algorithm for the second jet assigned to the hadronically decaying W boson
$\Delta R(\mathrm{b_{lep}}, \mathrm{W_{lep}})$	0.074	ΔR between the bottom quark from the leptonically decaying top quark and the leptonically decaying W boson
$\log\, p_\mathrm{T}(\mathrm{t_{had}})/\mathrm{GeV}$	0.070	logarithm of transverse momentum of the hadronically decaying top quark
$\log\, m(\mathrm{t_{lep}})/\mathrm{GeV}$	0.068	logarithm of invariant mass of the reconstructed leptonically decaying top quark
$\Delta R(\mathrm{b_{had}}, \mathrm{W_{had}})$	0.061	ΔR between the bottom quark from the hadronically decaying top and the hadronically decaying W boson
$\Delta R(\mathrm{W_{had}\ jet1}, \mathrm{W_{had}\ jet2})$	0.051	ΔR between the two jets from the hadronic W boson decay
$\log\, p_\mathrm{T}(\mathrm{t_{lep}})/\mathrm{GeV}$	0.040	logarithm of transverse momentum of the leptonically decaying top quark
$\frac{p_\mathrm{T}(\mathrm{t_{had}})+p_\mathrm{T}(\mathrm{t_{lep}})}{H_\mathrm{T}+p_\mathrm{T}^\mathrm{miss}+p_\mathrm{T}(\mathrm{lep})}$	0.033	transverse momentum fraction assigned to event compared to sum of all transverse momenta

Table B.5.: Efficiencies of the tHq jet assignment: the efficiencies are listed for all three years for the BDT-based approach. The χ^2-based efficiencies are listed for comparison.

method	t	H	light jet	tH	full
2016					
χ^2	58.8%	60.2%	81.1%	48.9%	43.2%
BDT	72.0%	72.3%	85.9%	64.8%	57.4%
2017					
χ^2	57.7%	59.7%	81.0%	48.0%	42.1%
BDT	71.9%	72.4%	86.9%	65.0%	58.0%
2018					
χ^2	58.2%	59.5%	80.0%	48.2%	41.7%
BDT	72.0%	72.2%	86.5%	65.0%	58.0%

Table B.6.: Efficiencies of the tHW jet assignment: the efficiencies are listed for all three years for the BDT-based approach. The χ^2-based efficiencies are listed for comparison.

method	b	W_{had}	t	H	W_b	full
2016						
χ^2	46.0%	48.7%	36.0%	55.4%	59.6%	28.0%
BDT	75.5%	79.7%	58.2%	78.5%	68.4%	51.2%
2017						
χ^2	44.5%	46.6%	34.2%	54.0%	58.3%	26.4%
BDT	75.6%	80.0%	57.7%	78.6%	67.9%	50.8%
2018						
χ^2	44.5%	46.7%	34.4%	53.8%	58.5%	26.4%
BDT	75.9%	80.2%	58.3%	78.9%	68.7%	51.5%

B.3. Jet Assignment Efficiencies

The efficiencies of the jet assignment under the tHq hypothesis are listed in Table B.5, under the tHW assignment in Table B.6, under the $t\bar{t}H$ assignment in Table B.7, and under the $t\bar{t}$ assignment in Table B.8.

Table B.7.: Efficiencies of the $t\bar{t}H$ jet assignment: the efficiencies are listed for all three years for the BDT-based approach. The χ^2-based efficiencies are listed for comparison.

method	b_{had}	W_{had}	t_{had}	t_{lep}	H	full
2016						
χ^2	39.7%	40.1%	30.9%	47.8%	38.8%	20.5%
BDT	67.5%	75.0%	60.1%	63.6%	57.6%	45.1%
2017						
χ^2	37.1%	36.6%	27.9%	45.0%	35.8%	17.9%
BDT	66.8%	74.8%	59.0%	62.2%	56.3%	43.5%
2018						
χ^2	37.1%	36.6%	28.0%	45.1%	35.7%	17.9%
BDT	66.6%	74.7%	59.0%	61.8%	55.7%	43.2%

Table B.8.: Efficiencies of the $t\bar{t}$ jet assignment: the efficiencies are listed for all three years for the BDT-based approach. The χ^2-based efficiencies are listed for comparison.

method	b_{had}	W_{had}	t_{had}	t_{lep}	full
2016					
χ^2	53.1%	51.6%	43.5%	62.7%	35.5%
BDT	79.0%	78.0%	69.9%	81.2%	67.1%
2017					
χ^2	49.6%	47.7%	39.8%	59.7%	31.8%
BDT	78.1%	76.9%	68.4%	80.0%	65.5%
2018					
χ^2	49.0%	47.6%	39.6%	59.7%	31.9%
BDT	78.3%	77.1%	68.8%	80.5%	66.1%

B.4. Overtraining Check of JA BDTs

The distributions of the JA BDTs are shown in Fig. B.9 and B.10 for 2016 and in Fig. B.11 and B.12 for 2018.

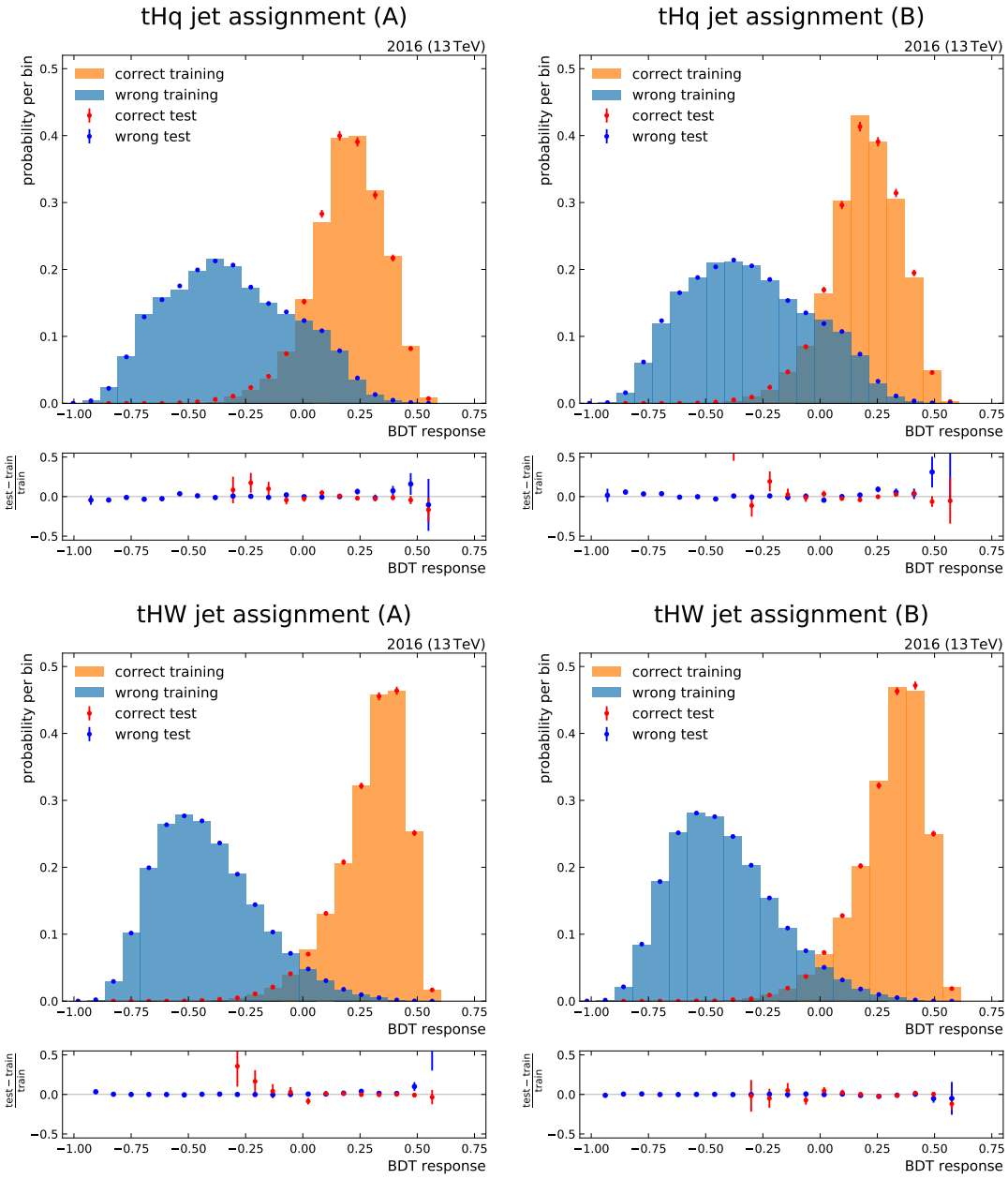

Figure B.9.: Overtraining check for tHq and tHW JA BDTs (2016): the distributions on the left show BDT response for the training on category A events and testing on category B. The right plots show the BDT response for the training on category B events and testing on category A. All distributions show good agreement between the outputs for the test and train data sets. No indication for overtraining is found. Since the wrong assignment contains the same events several times with different assignments, these events are not uncorrelated. Hence, a Kolmogorow-Smirnow test does not provide a good figure of merit, therefore it is not calculated here.

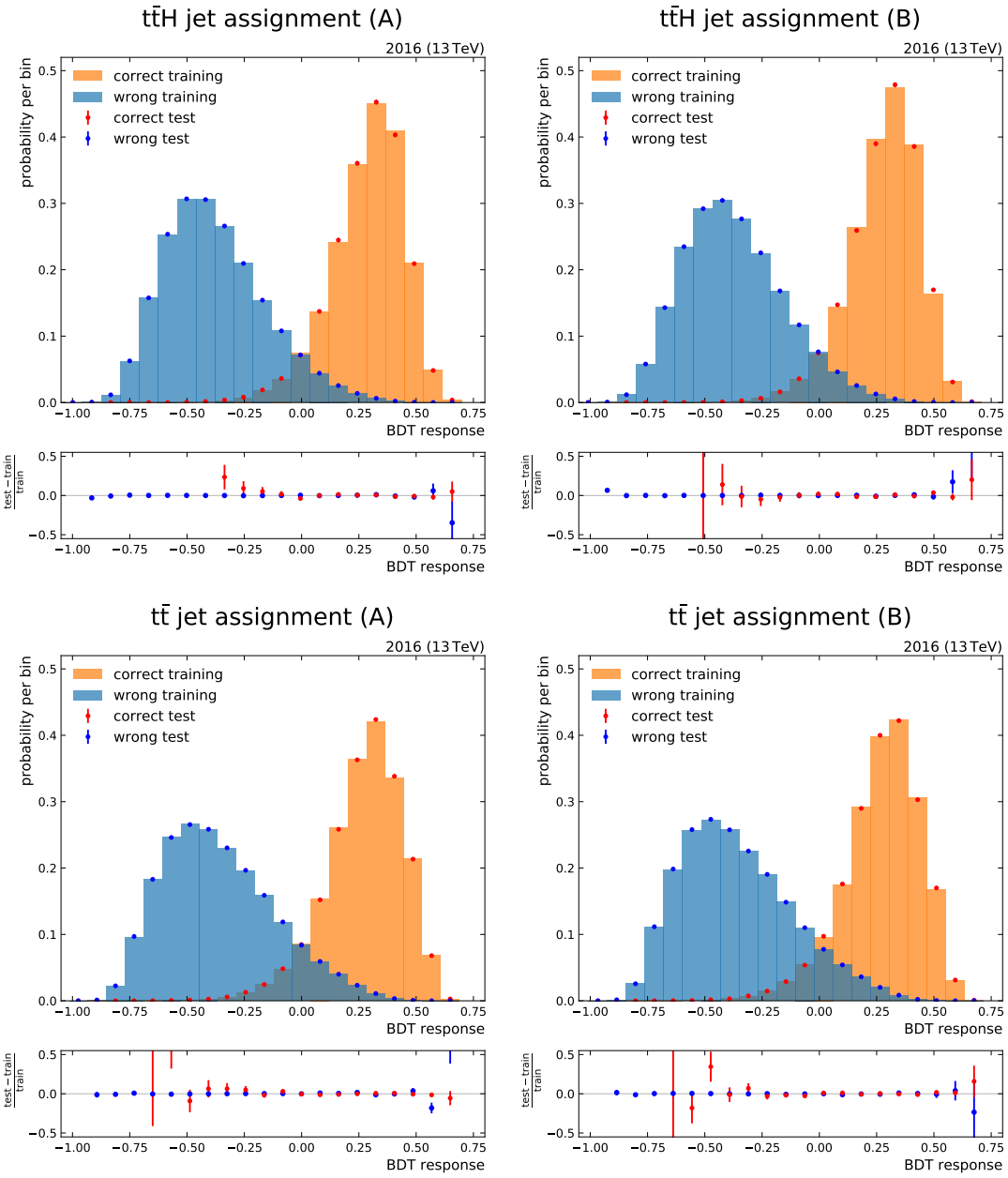

Figure B.10.: Overtraining check for $t\bar{t}H$ and $t\bar{t}$ JA BDTs (2016): the distributions on the left show the BDT response for the training on category A events and testing on category B. The right plots show the BDT response for the training on category B events and testing on category A. All distributions show good agreement between the outputs for the test and train data sets. No indication for overtraining is found. Since the wrong assignment contains the same events several times with different assignments, these events are not uncorrelated. Hence, a Kolmogorow-Smirnow test does not provide a good figure of merit, therefore it is not calculated here.

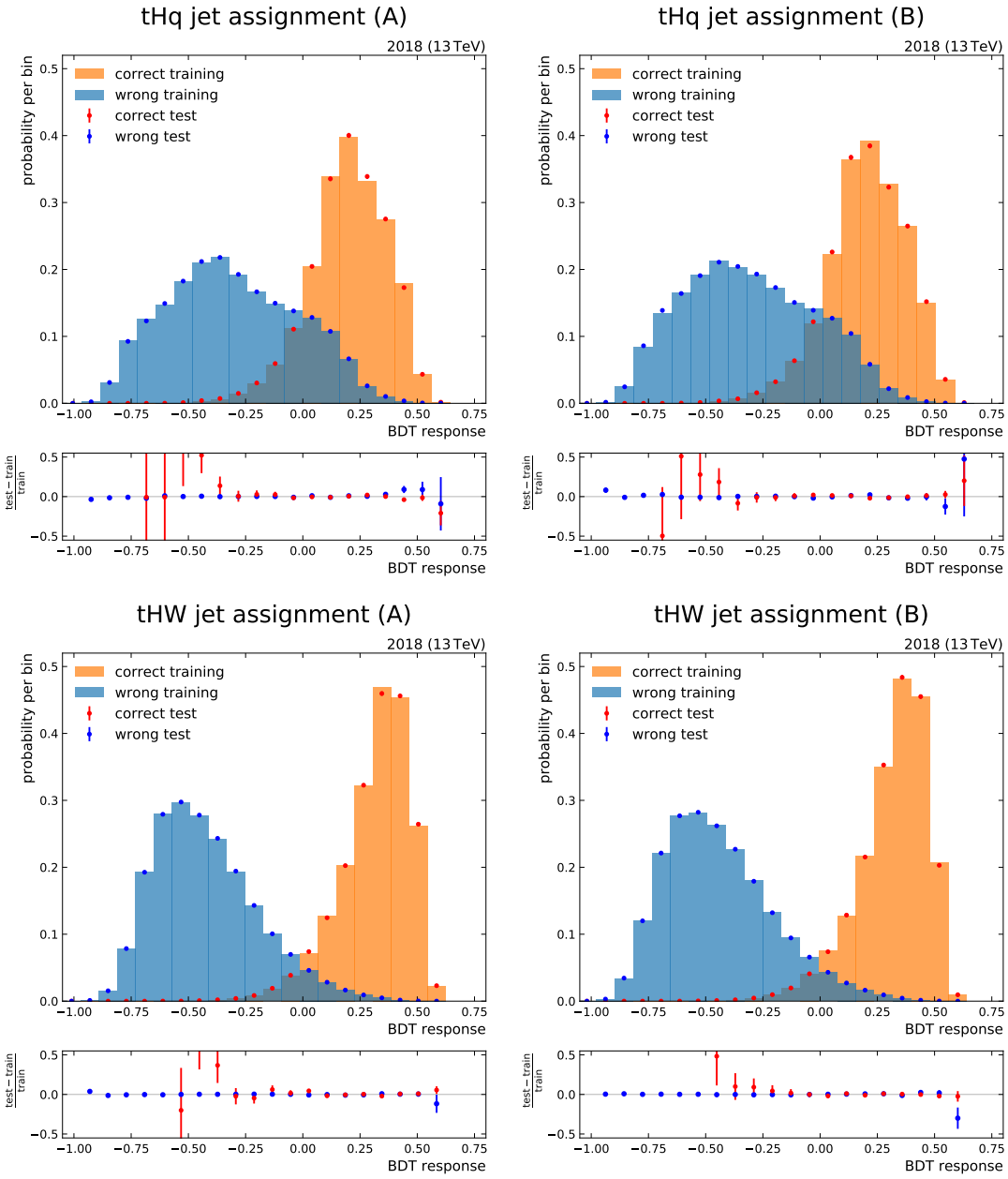

Figure B.11.: Overtraining check for tHq and tHW JA BDTs (2018): the distributions on the left show the BDT response for the training on category A events and testing on category B. The right plots show the BDT response for the training on category B events and testing on category A. All distributions show good agreement between the outputs for the test and train data sets. No indication for overtraining is found. Since the wrong assignment contains the same events several times with different assignments, these events are not uncorrelated. Hence, a Kolmogorow-Smirnow test does not provide a good figure of merit, therefore it is not calculated here.

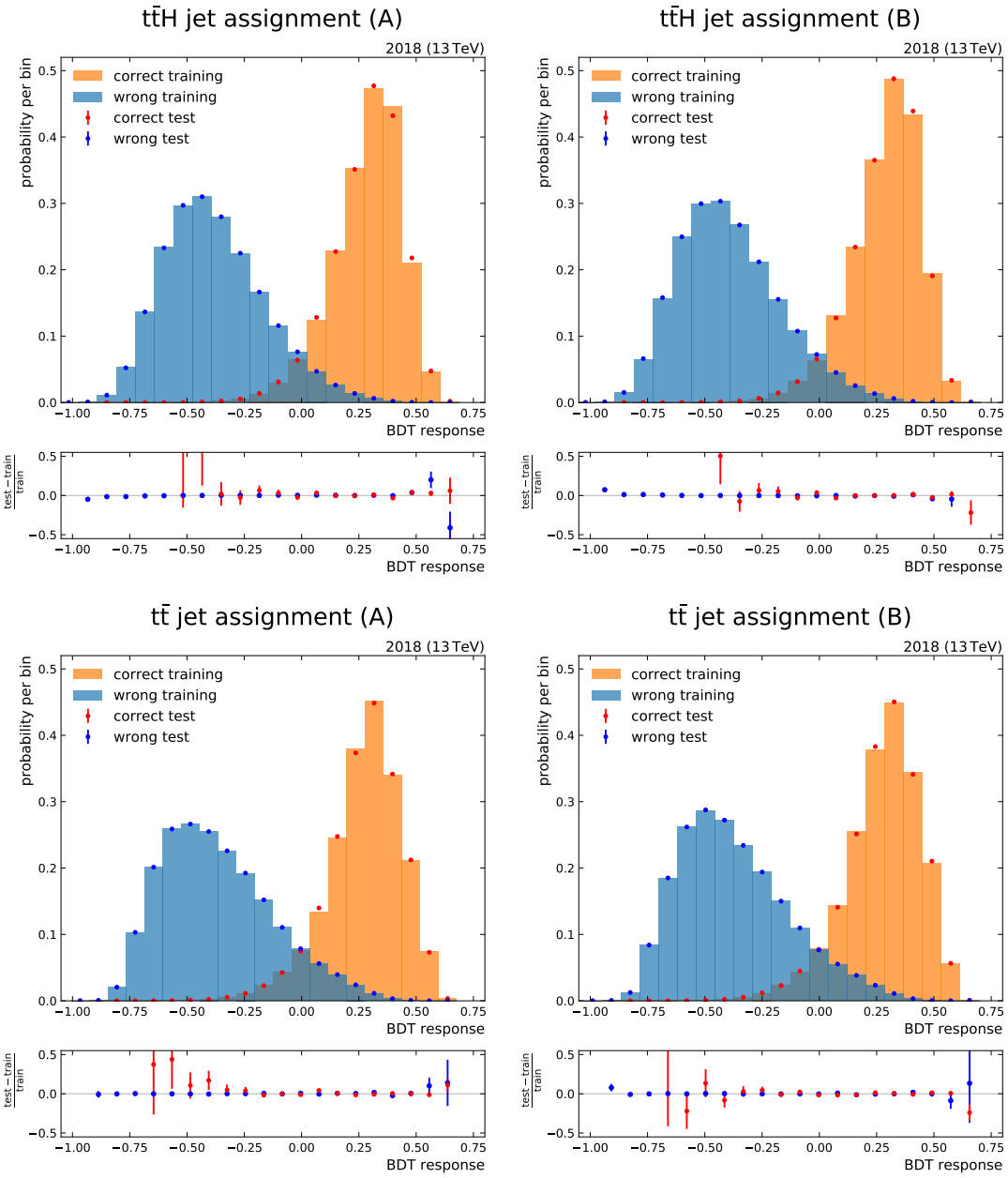

Figure B.12.: Overtraining check for $t\bar{t}H$ and $t\bar{t}$ JA BDTs (2018): the distributions on the left show the BDT response for the training on category A events and testing on category B. The right plots show the BDT response for the training on category B events and testing on category A. All distributions show good agreement between the outputs for the test and train data sets. No indication for overtraining is found. Since the wrong assignment contains the same events several times with different assignments, these events are not uncorrelated. Hence, a Kolmogorow-Smirnow test does not provide a good figure of merit, therefore it is not calculated here.

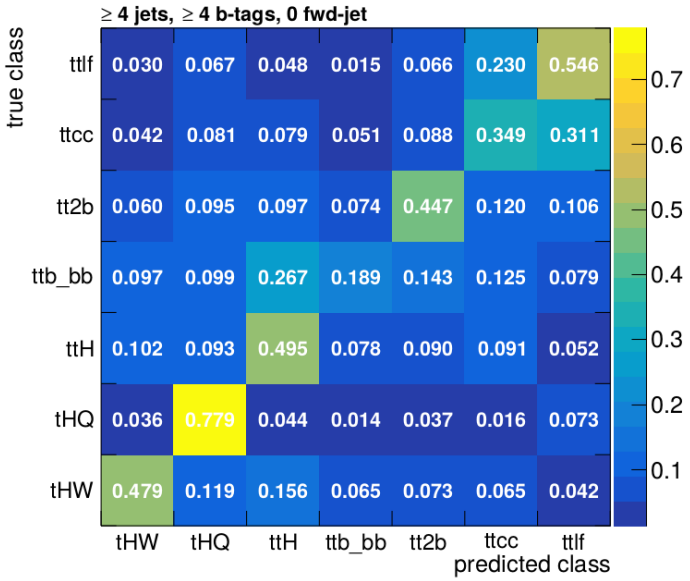

Figure B.13.: Confusion matrix of the classification DNN for the $\geq$ 4j $\geq$ 4t 0f region: the matrix shows the predicted class for a given true class.

B.5. DNN Training

The confusion matrices for the trainings in the $\geq$ 4j $\geq$ 4t 0f and $\geq$ 3j $\geq$ 3t $\geq$ 1f region are shown in Fig. B.13 and Fig. B.14, respectively.

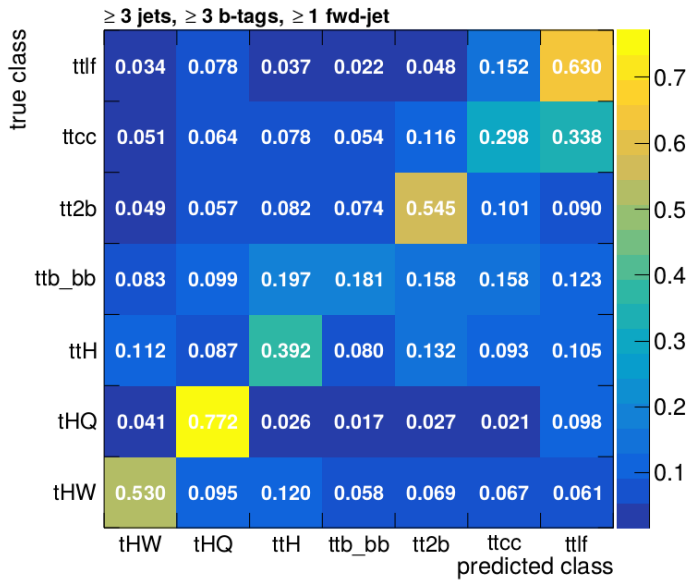

Figure B.14.: Confusion matrix of the classification DNN for the $\geq$ 3j $\geq$ 3t $\geq$ 1f region: the matrix shows the predicted class for a given true class.

B.6. Systematic Impacts

The impacts of the most important uncertainties to the fit on an Asimov dataset are shown in Fig. B.15.

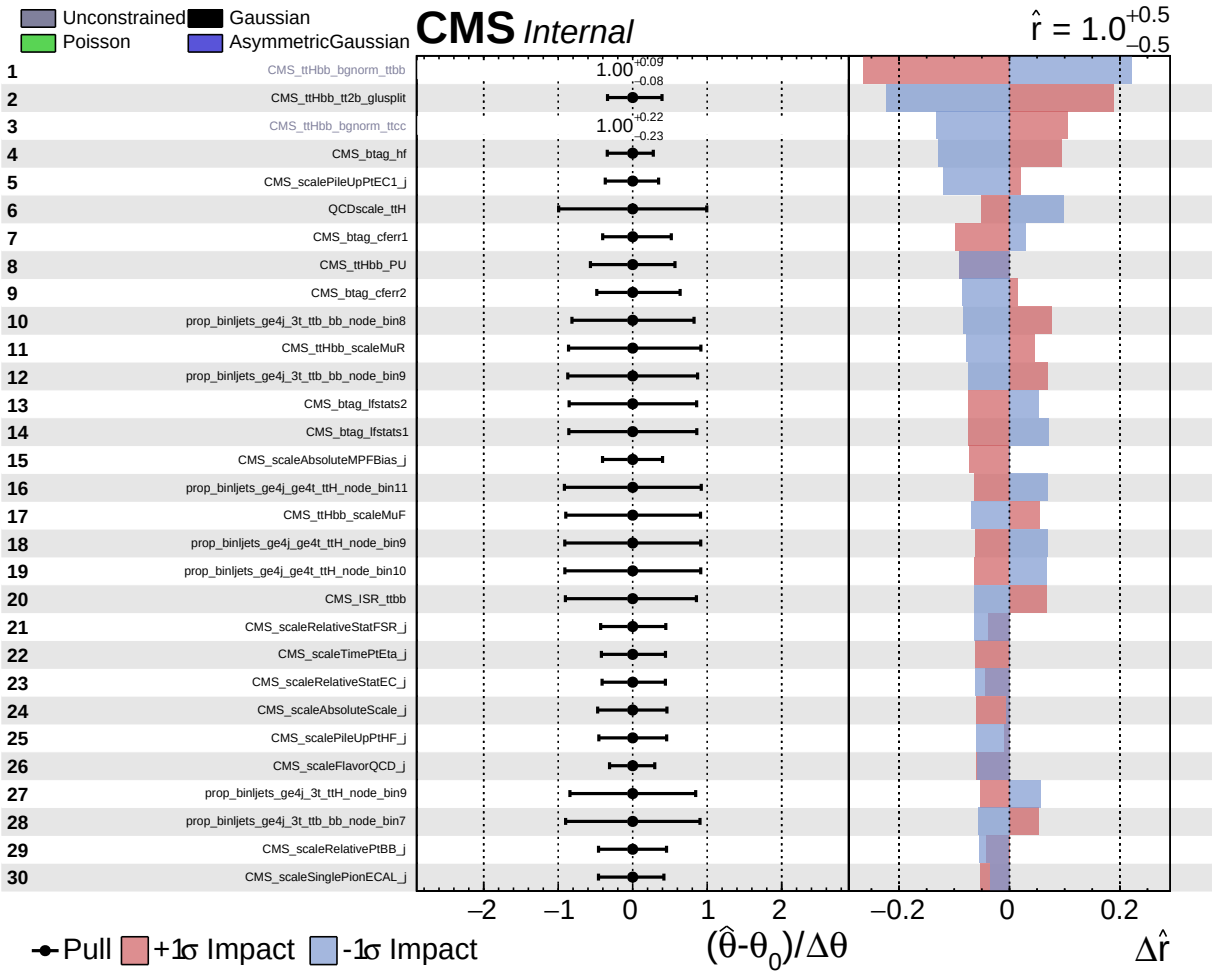

Figure B.15.: Impacts of the different systematic uncertainty sources: the uncertainties are ranked by their impact on the signal strength modifier.

Bibliography

[1] The 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-PAS-HIG-17-016, CERN, Geneva, 2018.

[2] The CMS Collaboration, "Search for associated production of a Higgs boson and a single top quark in proton-proton collisions at $\sqrt{s} = 13$ TeV," *Phys. Rev. D*, vol. 99, no. 9, p. 092005, 2019.

[3] M. Born, W. Heisenberg, and P. Jordan, "Zur Quantenmechanik. II.," *Z. Phys.*, vol. 35, no. 8-9, pp. 557–615, 1926.

[4] J. J. Aubert *et al.*, "Experimental observation of a heavy particle *j*," *Phys. Rev. Lett.*, vol. 33, pp. 1404–1406, 1974.

[5] J. E. Augustin *et al.*, "Discovery of a narrow resonance in e^+e^- annihilation," *Phys. Rev. Lett.*, vol. 33, pp. 1406–1408, 1974.

[6] 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.

[7] The D0 Collaboration, "Observation of the top quark," *Phys. Rev. Lett.*, vol. 74, pp. 2632–2637, 1995.

[8] The CDF Collaboration, "Observation of top quark production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab," *Phys. Rev. Lett.*, vol. 74, pp. 2626–2631, 1995.

[9] K. Kodama *et al.*, "Observation of tau neutrino interactions," *Phys. Lett. B*, vol. 504, pp. 218–224, 2001.

[10] The CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys. Lett. B*, vol. 716, no. 1, pp. 30 – 61, 2012.

[11] The ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett. B*, vol. 716, no. 1, pp. 1 – 29, 2012.

[12] The CMS Collaboration, "Observation of t$\bar{\text{t}}$H production," *Phys. Rev. Lett.*, vol. 120, no. 23, p. 231801, 2018.

[13] The CMS Collaboration, "Observation of Higgs Boson Decay to Bottom Quarks," *Phys. Rev. Lett.*, vol. 121, p. 121801, 2018.

[14] The ATLAS Collaboration, "Observation of H$\rightarrow$b$\bar{\text{b}}$ decays and VH production with the ATLAS detector," *Phys. Lett. B*, vol. 786, pp. 59 – 86, 2018.

[15] 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.

[16] The CMS Collaboration, "CMS Supersymmetry Physics Results," `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS`, May 2019.

[17] A. Einstein, "Die Grundlage der allgemeinen Relativitätstheorie," *Annalen der Phys.*, vol. 49, pp. 769–822, 1916.

[18] M. Tanabashi *et al.*, "Review of particle physics," *Phys. Rev. D*, vol. 98, p. 030001, 2018.

[19] Y. Fukuda *et al.*, "Evidence for oscillation of atmospheric neutrinos," *Phys. Rev. Lett.*, vol. 81, pp. 1562–1567, 1998.

[20] R. P. Feynman, "Space - time approach to quantum electrodynamics," *Phys. Rev. Lett.*, vol. 76, pp. 769–789, 1949.

[21] S. L. Glashow, "The renormalizability of vector meson interactions," *Nucl. Phys.*, vol. 10, pp. 107–117, 1959.

[22] S. Weinberg, "A model of leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 1967.

[23] A. Salam and J. C. Ward, "Weak and electromagnetic interactions," *Il Nuovo Cimento (1955-1965)*, vol. 11, no. 4, pp. 568–577, 1959.

[24] A. S. Kronfeld and C. Quigg, "Resource Letter: Quantum Chromodynamics," *Am. J. Phys.*, vol. 78, pp. 1081–1116, 2010.

[25] M. Faraday, "Experimental researches in electricity," 1839.

[26] J. C. Maxwell, "A dynamical theory of the electromagnetic field," *Philosophical Transactions of the Royal Society of London*, vol. 155, pp. 459–513, 1865.

[27] E. Fermi, "An attempt of a theory of beta radiation. 1.," *Z. Phys.*, vol. 88, pp. 161–177, 1934.

[28] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction," *Progress of Theoretical Physics*, vol. 49, pp. 652–657, 02 1973.

[29] L. Álvarez Gaumé and J. Ellis, "Eyes on a prize particle," *Nat. Phys.*, vol. 7, no. 1, pp. 2–3, 2011.

[30] R. D. Ball *et al.*, "Parton distributions from high-precision collider data," *Eur. Phys. J. C*, vol. 77, no. 10, p. 663, 2017.

[31] Y. Dokshitzer, "Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics," *Zh. Eksp. Teor. Fiz*, vol. 73, p. 1216, 1977.

[32] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl. Phys. B*, vol. 126, p. 298, 1977.

[33] V. Gribov and L. Lipatov, "Deep inelastic e p scattering in perturbation theory," *Sov. J. Nucl. Phys.*, vol. 15, pp. 438–450, 1972.

[34] The CMS Collaboration, "Summaries of CMS cross section measurements," `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined`, Jan. 2019.

[35] The D0 Collaboration, "Observation of Single Top Quark Production," *Phys. Rev. Lett.*, vol. 103, p. 092001, 2009.

[36] The CDF Collaboration, "First Observation of Electroweak Single Top Quark Production," *Phys. Rev. Lett.*, vol. 103, p. 092002, 2009.

[37] The CDF and D0 Collaborations, "Observation of Single Top Quark Production at the Tevatron Collider," *Mod. Phys. Lett. A*, vol. 25, pp. 309–339, 2010.

[38] The CMS Collaboration, "Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s}$ =8 TeV," *Phys. Rev. Lett.*, vol. 112, no. 23, p. 231802, 2014.

[39] The CDF and D0 Collaborations, "Observation of s-channel production of single top quarks at the Tevatron," *Phys. Rev. Lett.*, vol. 112, p. 231803, 2014.

[40] D. de Florian *et al.*, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector," 2016.

[41] The CMS Collaboration, "Combined measurements of Higgs boson couplings in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$," *Eur. Phys. J. C*, vol. 79, no. 5, p. 421, 2019.

[42] The ATLAS and CMS Collaborations, "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.

[43] The CMS Collaboration, "Observation of $t\bar{t}H$ Production," *Phys. Rev. Lett.*, vol. 120, p. 231801, 2018.

[44] M. Farina *et al.*, "Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson," *JHEP*, vol. 05, p. 022, 2013.

[45] F. Demartin *et al.*, "Higgs production in association with a single top quark at the LHC," *Eur. Phys. J. C*, vol. 75, no. 6, p. 267, 2015.

[46] The CMS Collaboration, "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.

[47] The ATLAS Collaboration, "Observation and measurement of Higgs boson decays to WW* with the ATLAS detector," *Phys. Rev. D*, vol. 92, no. 1, p. 012006, 2015.

[48] The CMS Collaboration, "Measurement of the properties of a Higgs boson in the four-lepton final state," *Phys. Rev. D*, vol. 89, no. 9, p. 092007, 2014.

[49] L. R. Evans and P. Bryant, "LHC Machine," *JINST*, vol. 3, p. S08001. 164 p, 2008. This report is an abridged version of the LHC Design Report (CERN-2004-003).

[50] The CMS Collaboration, "The CMS experiment at the CERN LHC," *JINST*, vol. 3, p. S08004, 2008.

[51] E. Mobs, "The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018," 2018.

[52] C. Wyss, "LEP design report, v.3: LEP2," CERN, Geneva, 1996. Vol. 1-2 publ. in 1983-84.

[53] European Organization for Nuclear Research, "Linear Accelerator 2," `https://home.cern/science/accelerators/linear-accelerator-2`, June 2019.

[54] European Organization for Nuclear Research, "The Proton Synchrotron Booster," `https://home.cern/science/accelerators/proton-synchrotron-booster`, June 2019.

[55] European Organization for Nuclear Research, "The Proton Synchrotron," `https://home.cern/science/accelerators/proton-synchrotron`, June 2019.

[56] The ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST*, vol. 3, p. S08003, 2008.

[57] The ALICE Collaboration, "The ALICE experiment at the CERN LHC," *JINST*, vol. 3, p. S08002, 2008.

[58] The LHCb Collaboration, "The LHCb Detector at the LHC," *JINST*, vol. 3, p. S08005, 2008.

[59] The TeVI Group, "Design Report Tevatron 1 project," FERMILAB-DESIGN-1984-01, 1984.

[60] V. Papadimitriou, "Luminosity determination at the Tevatron," in *Proceedings, LHC Lumi Days, LHC Workshop on LHC Luminosity Calibration: Geneva, Switzerland, 13-14 Jan*, 2011.

[61] The CMS Collaboration, "Public CMS Luminosity Information," `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#Online_Luminosity`, June 2019.

[62] T. Sakuma and T. McCauley, "Detector and event visualization with SketchUp at the CMS experiment," *Journal of Physics: Conference Series*, vol. 513, no. 2, p. 022032, 2014.

[63] S. R. Davis, "Interactive Slice of the CMS detector," 2016. CMS-OUTREACH-2016-027.

[64] M. Schott and M. Dunford, "Review of single vector boson production in pp collisions at $\sqrt{s} = 7$ TeV," *Eur. Phys. J. C*, vol. 74, p. 2916, 2014.

[65] D. Sprenger *et al.*, "Validation of Kalman Filter alignment algorithm with cosmic-ray data using a CMS silicon strip tracker endcap," *JINST*, vol. 5, p. P06007, 2010.

[66] A. Dominguez *et al.*, "CMS Technical Design Report for the Pixel Detector Upgrade," CERN-LHCC-2012-016. CMS-TDR-11, 2012.

[67] A. Benaglia, "The CMS ECAL performance with examples," *JINST*, vol. 9, p. C02008, 2014.

[68] F. De Guio *et al.*, "Performance of the CMS electromagnetic calorimeter and its role in the hunt for the higgs boson in the two-photon channel," *Journal of Physics: Conference Series*, vol. 455, p. 012028, 2013.

[69] G. Abbiendi, "The CMS muon system in Run2: preparation, status and first results," *PoS*, vol. EPS-HEP2015, p. 237, 2015.

[70] S. Cittolin, A. Rácz, and P. Sphicas, "CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project," Technical Design Report CMS, CERN, Geneva, 2002.

[71] The CMS Collaboration, "The CMS trigger system," *JINST*, vol. 12, no. 01, p. P01020, 2017.

[72] K. Bos *et al.*, "LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)," Technical Design Report LCG, CERN, Geneva, 2005.

[73] I. Bird *et al.*, "Update of the Computing Models of the WLCG and the LHC Experiments," CERN-LHCC-2014-014. LCG-TDR-002, 2014.

[74] The CMS Collaboration, "Documentation of the RooStats-based statistics tools for Higgs PAG," 2019.

[75] The ATLAS and CMS Collaborations, The LHC Higgs Combination Group, "Procedure for the LHC Higgs boson search combination in Summer 2011," CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, 2011.

[76] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling," *eConf*, vol. C0303241, 2003.

[77] V. Blobel and E. Lohrmann, "Statistische und numerische Methoden der Datenanalyse," Teubner Studienbücher Physik, Teubner, Stuttgart, 1998.

[78] 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.

[79] M. Baak *et al.*, "Interpolation between multi-dimensional histograms using a new non-linear moment morphing method," *Nuclear Instruments and Methods in Physics*, vol. 771, pp. 39–48, 2015.

[80] R. J. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples," *Comput. Phys. Commun.*, vol. 77, pp. 219–228, 1993.

[81] G. Cowan *et al.*, "Asymptotic formulae for likelihood-based tests of new physics," *Eur. Phys. J. C*, vol. 71, p. 1554, 2011.

[82] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," 1992.

[83] 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.

[84] A. L. Read, "Modified frequentist analysis of search results (the CL_s method)," 2000. CERN-OPEN-2000-205.

[85] 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.

[86] S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *Ann. Math. Statist.*, vol. 9, pp. 60–62, 03 1938.

[87] A. Hoecker *et al.*, "TMVA-Toolkit for multivariate data analysis," *arXiv preprint physics/0703039*, 2007.

[88] I. Antcheva *et al.*, "ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization," *Computer Physics Communications*, vol. 182, no. 6, pp. 1384–1385, 2011.

[89] C. Gini, "Variabilità e mutabilità," 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] R. E. Schapire and Y. Singer, "Improved boosting algorithms using confidence-rated predictions," *Machine Learning*, vol. 37, no. 3, pp. 297–336, 1999.

[92] L. Breiman *et al.*, "Classification and regression trees," The Wadsworth statistics/probability series, Wadsworth and Brooks/Cole Advanced Books and Software, Monterey, CA, 1984.

[93] Y. LeCun *et al.*, "Deep learning," *Nature*, vol. 521, pp. 436–444, 5 2015.

[94] D. E. Rumelhart *et al.*, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 1986.

[95] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014.

[96] G. E. Hinton *et al.*, "Improving neural networks by preventing co-adaptation of feature detectors," *CoRR*, vol. abs/1207.0580, 2012.

[97] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *CoRR*, vol. abs/1502.03167, 2015.

[98] S. Fink, "Probing the Top-Yukawa Coupling by Searching for Associated Higgs Boson Production with a Single Top Quark at the CMS Experiment," PhD thesis, KIT, 2016.

[99] G. Marsaglia *et al.*, "Evaluating kolmogorov's distribution," *Journal of Statistical Software, Articles*, vol. 8, no. 18, pp. 1–4, 2003.

[100] N. Smirnov, "Table for estimating the goodness of fit of empirical distributions," *Ann. Math. Statist.*, vol. 19, no. 2, pp. 279–281, 1948.

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

[102] R. D. Ball *et al.*, "Reweighting and Unweighting of Parton Distributions and the LHC W lepton asymmetry data," *Nucl. Phys. B*, vol. 855, pp. 608–638, 2012.

[103] R. D. Ball *et al.*, "Parton distributions for the LHC Run II," *JHEP*, vol. 04, p. 040, 2015.

[104] 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)].

[105] R. Frederix and S. Frixione, "Merging meets matching in MC@NLO," *JHEP*, vol. 12, p. 061, 2012.

[106] B. Andersson *et al.*, "Parton Fragmentation and String Dynamics," *Phys. Rept.*, vol. 97, pp. 31–145, 1983.

[107] A. Buckley *et al.*, "General-purpose event generators for LHC physics," *Phys. Rept.*, vol. 504, pp. 145–233, 2011.

[108] J. Alwall *et al.*, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP*, vol. 2014, no. 7, p. 1, 2014.

[109] J. Alwall *et al.*, "MadGraph 5 : Going Beyond," *JHEP*, vol. 1106, p. 128, 2011.

[110] J. Alwall *et al.*, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP*, vol. 1407, p. 079, 2014.

[111] S. Frixione *et al.*, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," *JHEP*, vol. 0711, p. 070, 2007.

[112] S. Alioli *et al.*, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP*, vol. 1006, p. 043, 2010.

[113] T. Sjöstrand *et al.*, "PYTHIA 6.4 Physics and Manual," *JHEP*, vol. 0605, p. 026, 2006.

[114] T. Sjöstrand *et al.*, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.*, vol. 178, pp. 852–867, 2008.

[115] J. Alwall *et al.*, "A Standard format for Les Houches event files," *Comput. Phys. Commun.*, vol. 176, pp. 300–304, 2007.

[116] P. Skands *et al.*, "Tuning PYTHIA 8.1: the Monash 2013 Tune," *Eur. Phys. J. C*, vol. 74, no. 8, p. 3024, 2014.

[117] The CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements," *Eur. Phys. J. C*, vol. 76, no. 3, p. 155, 2016.

[118] The CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements," 2019.

[119] S. Agostinelli *et al.*, "GEANT4: A Simulation toolkit," *Nuclear Instruments and Methods in Physics*, vol. A506, pp. 250–303, 2003.

[120] J. Allison *et al.*, "Geant4 developments and applications," *IEEE Transactions on Nuclear Science*, vol. 53, p. 270, 2006.

[121] The CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector," *JINST*, vol. 12, no. 10, p. P10003, 2017.

[122] The CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET," CMS-PAS-PFT-09-001, CERN, Geneva, 2009.

[123] The CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector," CMS-PAS-PFT-10-001, CERN, Geneva, 2010.

[124] The CMS Collaboration, "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV," CMS-PAS-PFT-10-002, CERN, Geneva, 2010.

[125] The CMS Collaboration, "Particle-flow commissioning with muons and electrons from J/Psi and W events at 7 TeV," CMS-PAS-PFT-10-003, CERN, Geneva, 2010.

[126] The CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker," *JINST*, vol. 9, no. 10, p. P10009, 2014.

[127] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting," *Nucl. Instrum. Meth.*, vol. A262, pp. 444–450, 1987.

[128] P. Billoir, "Progressive track recognition with a Kalman like fitting procedure," *Comput. Phys. Commun.*, vol. 57, pp. 390–394, 1989.

[129] P. Billoir and S. Qian, "Simultaneous pattern recognition and track fitting by the Kalman filtering method," *Nucl. Instrum. Meth.*, vol. A294, pp. 219–228, 1990.

[130] P. Billoir and S. Qian, "Further test for the simultaneous pattern recognition and track fitting by the Kalman filtering method," *Nucl. Instrum. Meth.*, vol. A295, pp. 492–500, 1990.

[131] T. Speer *et al.*, "Vertex fitting in the CMS tracker," 2006.

[132] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems," *IEEE Proc.*, vol. 86, no. 11, pp. 2210–2239, 1998.

[133] R. Frühwirth *et al.*, "Adaptive vertex fitting," *J. Phys. G*, vol. 34, p. N343, 2007.

[134] The CMS Collaboration, "Pileup Removal Algorithms," CMS-PAS-JME-14-001, CERN, Geneva, 2014.

[135] The CMS Collaboration, "Search for neutral higgs bosons decaying to τ pairs in pp collisions at $\sqrt{s} = 7$ TeV," *Phys. Lett. B*, vol. 713, p. 68, 2012.

[136] The CMS Collaboration, "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.

[137] The CMS Collaboration, "Performance of Jet Algorithms in CMS," CMS-PAS-JME-07-003, CERN, Geneva, 2007.

[138] M. Cacciari *et al.*, "The Anti-k(t) jet clustering algorithm," *JHEP*, vol. 04, p. 063, 2008.

[139] Y. L. Dokshitzer *et al.*, "Better jet clustering algorithms," *JHEP*, vol. 08, p. 001, 1997.

[140] S. Catani *et al.*, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions," *Nucl. Phys. B*, vol. 406, pp. 187–224, 1993.

[141] The CMS Collaboration, "Jet algorithms performance in 13 TeV data," CMS-PAS-JME-16-003, CERN, Geneva, 2017.

[142] The CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV," *JINST*, vol. 12, no. 02, p. P02014, 2017.

[143] The CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS," *JINST*, vol. 6, p. 11002, 2011.

[144] The CMS Collaboration, "CMS JEC Run I legacy performance plots," 2015. CMS-DP-2015-044.

[145] The CMS Collaboration, "Introduction to Jet Energy Corrections at CMS," `https://twiki.cern.ch/twiki/bin/viewauth/CMS/IntroToJEC`, Sept. 2019.

[146] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas," *Phys. Lett. B*, vol. 659, pp. 119–126, 2008.

[147] The CMS Collaboration, "Identification of b-quark jets with the CMS experiment," *JINST*, vol. 8, p. P04013, 2013.

[148] The CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2," CMS-PAS-BTV-15-001, CERN, Geneva, 2016.

[149] The CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV," *JINST*, vol. 13, no. 05, p. P05011, 2018.

[150] The CMS Collaboration, "Performance of the CMS missing transverse momentum reconstruction in pp data at $\sqrt{s}$ = 8 TeV," *JINST*, vol. 10, no. 02, p. P02006, 2015.

[151] The CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector," *JINST*, vol. 14, no. 07, p. P07004, 2019.

[152] The CMS Collaboration, "MET Analysis," `https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookMetAnalysis`, Oct. 2019.

[153] 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, KIT, CERN-THESIS-2010-255, 2010.

[154] L. Stockmeier, "Study of the Impact of Higgs Boson Variables in the Jet-to-Quark Assignment for the Investigation of Higgs Boson Couplings to Top Quarks and Vector Bosons with the CMS Experiment," Bachelor thesis, KIT, 2019.

[155] J. Rauser, "Search for Higgs boson production in association with a single top quark with the CMS experiment," Master thesis, KIT, 2017.

[156] K. Flöh, "Search for single top quarks in association with a Higgs boson, $H\rightarrow b\bar{b}$, at a center-of-mass energy of 13 TeV with the CMS detector," Master thesis, KIT, 2017.

[157] The CMS Collaboration, "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.

[158] The CMS Collaboration, "Search for $H\rightarrow b\bar{b}$ in association with a single top quark as a test of Higgs boson couplings at $\sqrt{s} = 13$ TeV," CMS-PAS-HIG-16-019, CERN, Geneva, 2016.

[159] N. Faltermann, "Single top quark production at 13 tev with the cms experiment: from rediscovery to search for rare channels and determination of higgs boson couplings,". PhD thesis, Karlsruher Institut für Technologie (KIT), 2018.

[160] The 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-PAS-HIG-17-005, CERN, Geneva, 2017.

[161] The CMS Collaboration, "Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV," *JHEP*, vol. 11, p. 185, 2018.

[162] The CMS Collaboration, "A deep neural network for simultaneous estimation of b jet energy and resolution," *arXiv preprint 1912.06046*, 2019.

[163] D. Bühler, "Improvement of the b-Jet Energy Reconstruction via Multivariate Regression for the Investigation of Higgs Boson Couplings to Top Quarks and Vector Bosons," Master thesis, KIT, 2019.

[164] D. Bühler, "Datenbasierte QCD-Untergrundmodellierung für die Untersuchung von Higgs-Boson-Produktion in Assoziation mit einzelnen Top-Quarks im Elektron- Kanal," Bachelor thesis, KIT, 2015.

[165] 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 thesis, KIT, 2017.

[166] The CMS Collaboration, "Performance of CMS Muon Reconstruction in pp Collision Events at $\sqrt{s} = 7$ TeV," *JINST*, vol. 7, p. P10002, 2012.

[167] The CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV," *JINST*, vol. 13, no. 06, p. P06015, 2018.

[168] The CMS Collaboration, "Electron Tag-and-Probe," `https://twiki.cern.ch/twiki/bin/view/CMSPublic/ElectronTagAndProbe`, Oct. 2019.

[169] The CMS Collaboration, "Utilities for Accessing Pileup Information for Data," `https://twiki.cern.ch/twiki/bin/view/CMS/PileupJSONFileforData`, Oct. 2019.

[170] The CMS Collaboration, "CMS Luminosity Measurements for the 2016 Data Taking Period," CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

[171] The CMS Collaboration, "Jet Energy Resolution," `https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution`, Oct. 2019.

[172] The CMS Collaboration, "Jet energy scale uncertainty sources," `https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources`, Oct. 2019.

[173] The CMS Collaboration, "pt(top-quark) based reweighting of ttbar MC," `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopPtReweighting`, Oct. 2019.

[174] The CMS Collaboration, "Measurement of the $\mathrm{t\bar{t}}$ production cross section in the all-jets final state in pp collisions at $\sqrt{s} = 8$ TeV," *Eur. Phys. J. C*, vol. 76, no. 3, p. 128, 2016.

[175] The CMS Collaboration, "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.

[176] The CMS Collaboration, "Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV," *Phys. Rev. D*, vol. 95, no. 9, p. 092001, 2017.

[177] The CMS Collaboration, "Measurement of the differential cross section for $\mathrm{t\bar{t}}$ production in the dilepton final state at $\sqrt{s} = 13$ TeV," CMS-PAS-TOP-16-011, CERN, Geneva, 2016.

[178] A. Kalogeropoulos and J. Alwall, "The SysCalc code: A tool to derive theoretical systematic uncertainties," 2018.

[179] V. D. Barger, J. Ohnemus, and R. J. N. Phillips, "Event shape criteria for single lepton top signals," *Phys. Rev.*, vol. D48, pp. R3953–R3956, 1993.

[180] E. Sarauer, "Study of the Jet Assignment BDT Output Shape in the tH Analysis," Bachelor thesis, KIT, 2019.

[181] T. Friedrich, "Event Classification based on a Deep Neural Network for the Investigation of Higgs Boson Couplings to Top Quarks and Vector Bosons with the CMS Experiment," Bachelor thesis, KIT, 2019.

[182] The CMS Collaboration, "Measurement of $\mathrm{t\bar{t}H}$ production in the $\mathrm{H} \to \mathrm{b\bar{b}}$ decay channel in $41.5\,\mathrm{fb}^{-1}$ of proton-proton collision data at $\sqrt{s} = 13\,\mathrm{TeV}$," CMS-PAS-HIG-18-030, CERN, Geneva, 2019.

[183] The 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`, Oct. 2019.

Danksagungen

Ich danke Prof. Dr. Thomas Müller für die Betreuung dieser Arbeit. Außerdem bedanke ich mich für die Konferenzen und Sommerschulen, die ich besuchen durfte und die wertvollen Erfahrungen, die ich dadurch sammeln konnte. Auch für den halbjährigen CERN-Aufenthalt bin ich sehr dankbar, dadurch wurde mir die Kombination meiner Analyse mit den Analysen der anderen Zerfallskanäle erheblich erleichtert.

Des Weiteren danke ich Prof. Dr. Günter Quast für die Übernahme des Korreferats und die Kommentare zu dieser Arbeit. Auch für die Aufnahme in die Computing-Gruppe bedanke ich mich.

Dr. Thorsten Chwalek und Dr. Nils Faltermann danke ich für die hervorragende Betreuung und die Beantwortung meiner Fragen.

Weiterhin danke ich Darius Bühler, Dr. Nils Faltermann, Dr. Daniela Schäfer und PD Dr. Roger Wolf für das Korrekturlesen dieser Arbeit und die vielen wertvollen Anmerkungen. Für die exzellente Zusammenarbeit innerhalb der Analysegruppe danke ich Darius Bühler, Dr. Nils Faltermann und Marco Link. Zudem danke ich meinen Studenten und Praktikanten Marc Buckmakowski, Bastian Czech, Tabea Friedrich, Ellen Sarauer, David Schilling, Lea Stockmeier und Komal Tauqeer. Es hat mir immer große Freude bereitet euch zu betreuen. Der $t\bar{t}H$ Gruppe des Campus Nord danke ich für die Zusammenarbeit in der kombinierten tH+$t\bar{t}H$ Analyse. Auch den restlichen Mitgliedern der Arbeitsgruppe danke ich für die Unterstützung, die ich in den letzten Jahren erfahren habe.

Matthias Schnepf danke ich für die Hilfestellungen bei der Nutzung der Computing Resourcen. Die Nutzung des bwForCluster NEMO an der Universität Freiburg war für die Produktion der zusätzlichen Signalereignisse besonders hilfreich. Den Admins des Instituts danke ich für die Bereitstellung der IT-Infrastruktur. Frau Bräunling danke ich für die Hilfe bei der Bewältigung sämtlicher bürokratischer Hürden. Allen weiteren Kollegen am ETP danke ich für die ausgezeichnete Arbeitsathmosphäre.

Meiner Familie und meinen Freunden danke ich für die Unterstützung, die ich die ganze Zeit über erfahren habe. Insbesondere danke ich Helena Lamprecht für die vielen Stunden, die sie aufgebracht hat um mich für die Prüfung abzufragen.