Spin-correlation studies in a search for Higgs-boson production in association with top quarks at CMS

Master’s Thesis of

Michael Waßmer

at the Department of Physics
Institut für Experimentelle Kernphysik (IEKP)

Reviewer: Prof. Dr. Ulrich Husemann
Second reviewer: Prof. Dr. Thomas Müller

Accepted by the first referee of the master’s thesis.

Karlsruhe, 11.01.2017

(Prof. Dr. Ulrich Husemann)
I declare that I have developed and written the enclosed thesis completely by myself, and have not used sources or means without declaration in the text.

Karlsruhe, 11.01.2017

-----------------------------
(Michael Waßmer)
Contents

1. Introduction .......................................................... 1

2. The Standard Model of Particle Physics ......................... 3
   2.1. Fermions ....................................................... 3
   2.2. Bosons ....................................................... 4
   2.3. The Fundamental Interactions of Particle Physics .......... 5
   2.4. Feynman Diagrams, Cross Sections and Decay Rates ........ 14
   2.5. Hadron Collider Physics .................................... 16

3. The Compact Muon Solenoid (CMS) at the LHC .................. 19
   3.1. The LHC ..................................................... 19
   3.2. The Compact Muon Solenoid ................................ 21
   3.3. Object Reconstruction ....................................... 30

4. Search for the Associated Production of a Higgs Boson and a Top-Quark-Antiquark Pair ......................... 35
   4.1. Top Quark, Higgs Boson, and their Associated Production ...... 35
   4.2. Final State Topology ......................................... 39
   4.3. Backgrounds ................................................. 40

5. Event Simulation .................................................... 43
   5.1. Event Generation ............................................ 43
   5.2. Monte Carlo Datasets ........................................ 45
   5.3. Additional Jet-Flavour Identification in Top-Quark Pair Events .... 46
   5.4. Corrections to Simulated Data ................................ 46

6. Event Selection ...................................................... 49
   6.1. Physics Objects ............................................. 49
   6.2. Event Selection .............................................. 52

7. Statistical Model and Tools ....................................... 55
   7.1. Statistical Model ........................................... 55
   7.2. Systematic Uncertainties ................................... 58
   7.3. Boosted Decision Trees ..................................... 60

8. Theory of Spin Correlations ...................................... 63
   8.1. Introduction ................................................ 63
1. Introduction

In October 2013, Peter Higgs and François Englert were awarded the Nobel Prize in Physics “for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN’s Large Hadron Collider” [1–3]. The discovered particle is compatible with the properties of the Higgs boson of the Standard Model [4, 5]. In the Standard Model of Particle Physics, the masses of elementary particles are generated by the Higgs mechanism [6, 7] and their coupling to the Higgs field.

The top quark is the heaviest fundamental fermionic particle of the Standard Model. The coupling of fermions to the Higgs boson is described by a Yukawa interaction, which is proportional to the fermion mass after spontaneous symmetry breaking. Thus, compared to other fermions, the top quark coupling to the Higgs field is extraordinarily large. This makes the coupling of the top quark to the Higgs boson especially interesting. Furthermore, there are several theoretical questions depending on this coupling, for example the energy scale where new physics is expected to be found [8].

The coupling between the top quark and the Higgs boson can be measured in the associated production of a Higgs boson with a top-quark-antiquark pair (t¯tH). This process allows a model independent measurement of the coupling since its production rate depends exclusively on this coupling. If discrepancies between the predicted and the measured rate are observed, this could point towards physics beyond the Standard Model (BSM physics). However, the t¯tH process first has to be discovered before precise measurements of the coupling are possible. This thesis is dedicated to improving the search for the t¯tH process.

The Higgs boson and the top quark can decay in several different channels. In this thesis, the investigated final state is where the Higgs boson decays into two bottom-type quarks, from now on referred to as tH(bb), and the top-quark-antiquark pair decays in the semileptonic channel. This final state has a sizable branching fraction and still allows discrimination of background processes with the expected lepton and neutrino as well as the high number of jets originating from bottom quarks. Since the cross section of the tH process is already very low, approximately 0.5 pb at √s = 13 TeV [9], the expected number of semileptonic tH(bb) events is very small. Moreover, the irreducible tbb background with the same final state has a significantly higher cross section of O(1 pb) [10, 11]. Hence, multivariate techniques are used to improve the detection of signal events.
1. Introduction

These multivariate techniques detect differences and correlations in the distributions of observables, which can consequently be used to quantify of how signal- or background-like a specific event is. In this thesis, a set of angular variables, sensitive to spin correlations between the top quarks, is investigated. In theory, a qualitative and quantitative difference is expected between the signal process $t\bar{t}H(b\bar{b})$ and the irreducible background $t\bar{t}b\bar{b}$ regarding the spin correlations of the top quarks [12]. Therefore, angular variables of the decay products of top-quark-antiquark pairs which allow to exploit these physical differences within the semileptonic $t\bar{t}H(b\bar{b})$ analysis [13] have to be found and their impact on the analysis has to be tested. First, theoretically motivated angular variables were validated regarding their modelling in simulation. In addition, reference frames besides the lab frame which are more sensitive to spin correlations were investigated. After good compatibility of theoretical expectation and simulation was confirmed, the separation power of the angular variables at generator level was examined and consequently demonstrated. Finally, the step to reconstructed quantities was performed to investigate its impact on the angular variables. Despite several difficulties regarding the reconstruction of the $t\bar{t}$ system, an improvement of the expected upper limit of the $t\bar{t}H(b\bar{b})$ single-lepton analysis by 3% was achieved. In addition, several improvement and optimisation possibilities to further increase the benefit of the angular variables are presented.

In chapter 2, an overview of the current state of the Standard Model of Particle Physics is given. Afterwards, chapter 3 is focused on the CMS detector, which provides the measured data. In chapter 4, the CMS $t\bar{t}H(b\bar{b})$ analysis is reviewed in more detail. Chapter 5 is dedicated to an introduction to event generation with Monte Carlo event generators. Chapter 6 explains the selection of events. After that, chapter 7 introduces the statistical model and a multivariate classifier. Following this, an introduction to the theory of spin correlations is given in chapter 8 that builds the basis of the observables investigated in this thesis. Subsequently, the validation of the simulated data regarding their description of the spin correlations is performed in chapter 9. In chapter 10, the separation potential of the angular variables is evaluated at generator level. Chapter 11 is dedicated to the angular variables at reconstruction level, arising difficulties, and the impact on the existing analysis. In chapter 12, conclusions are drawn.
The properties of, and the interaction between, elementary particles are theoretically described by the Standard Model (SM) of particle physics. The SM contains a rich set of particles and describes the interactions between these particles with the help of quantum-mechanics [14–16], special relativity [17] and field theory. Those are combined to a quantum-field theory, which provides the foundation to derive and predict quantities to be compared with measurements. These quantities are called observables. The predictions of the SM were and still are constantly being tested by experiments to further investigate the predictions and precision of the model. The SM distinguishes between different kinds of particles and different interactions all with their own properties, which will be explained in the following sections. For orientation, two standard textbooks [18, 19] were used.

2. The Standard Model of Particle Physics

The properties of, and the interaction between, elementary particles are theoretically described by the Standard Model (SM) of particle physics. The SM contains a rich set of particles and describes the interactions between these particles with the help of quantum-mechanics [14–16], special relativity [17] and field theory. Those are combined to a quantum-field theory, which provides the foundation to derive and predict quantities to be compared with measurements. These quantities are called observables. The predictions of the SM were and still are constantly being tested by experiments to further investigate the predictions and precision of the model. The SM distinguishes between different kinds of particles and different interactions all with their own properties, which will be explained in the following sections. For orientation, two standard textbooks [18, 19] were used.

2.1. Fermions

Fermions are one of the two classes in which particles can be divided in the most general sense. Fermions are particles with half-integer spin in units of $\hbar$. An ensemble of fermions follows the Fermi-Dirac statistics [20]. Also the quantum mechanical wave function of a system of identical fermions has to be antisymmetric, if two particles are interchanged. For every fermion there is a antifermion with opposite charge and magnetic moment, but same mass and lifetime. The fermions of the Standard Model can further be split up into two categories: leptons and quarks, which mainly differ by the interactions they take part in.

2.1.1. Leptons

Leptons can only interact via the electroweak interaction. Therefore, they must carry electric charge and/or weak isospin, see 2.3.2 and 2.3.3. The electrons in the orbitals of an atom or the muons, which were discovered in cosmic radiation [21], are charged leptons. The muon has the same properties as the electron except its mass is around 200 times larger than the electron mass. In addition, there is another charged lepton, which was found in 1975 by Martin Perl et al. [22] and is named $\tau$-lepton. Its mass is around 3500 times larger than the electron mass. The heavier charged leptons can
decay into the lighter charged leptons in association with two additional neutrinos by the electroweak interaction. In addition, the $\tau$-lepton can decay hadronically due to its high mass. The already mentioned neutrinos are neutral leptons, which do not interact by the electromagnetic interaction but by the weak force. For every charged lepton, there is a corresponding neutrino. The three types of charged leptons can then be arranged into three generations with their corresponding neutrino. In the Standard Model, the neutrinos are predicted to be massless. However, this leads to conflicts with the observation of, for example, neutrino oscillations and the measurements of modern neutrino experiments, which require at least one neutrino to have mass.

2.1.2. Quarks

Quarks are the second kind of fermions in the Standard Model. They have spin 1/2 in units of $\hbar$, carry non-integer electric charge, and only interact by the electroweak and the strong interaction. Ordinary matter, i.e. protons and neutrons, is made up of the two lightest quarks, the up and the down quark. All together, there are six quarks, which can be ordered in three different families or generations analogous to the three different lepton flavours. In addition, every quark can be assigned one flavour and three colours due to the SU(3) colour symmetry of the strong force. The six quarks cover a wide mass range from $\mathcal{O}(\text{MeV})$ to $\mathcal{O}(100\ \text{GeV})$ and every quark is characterized by a set of quantum numbers, see “Quark Model and Masses” in [23]. The heavier quarks can decay into lighter quarks by the weak interaction, which can be parametrized by the CKM-matrix. Historically, the quark model was developed by Gell-Mann to explain the wide variety of particles that were discovered in the time before 1960 [24]. The quarks can form bound states, which are called mesons and baryons, but cannot exist as free particles because of a property of the strong force, called confinement. The mesons consist of a quark and an antiquark ($q\bar{q}'$), whereas the baryons are made up of three quarks ($qq'q''$) or three antiquarks. The quark model uses group theory to bring order into the many different possible combinations of quarks, which together can form bound states.

2.2. Bosons

In addition to fermions, there is a second kind of particles in the Standard Model, the bosons. In contrast to the fermions, they carry integer spin in units of $\hbar$ and an ensemble of bosons follows Bose-Einstein statistics [25]. The wave function of a system of identical bosons has to be completely symmetric if two particles are interchanged. In addition, there is a special kind of bosons, the gauge bosons, which are responsible for the fundamental interactions between particles in the Standard Model. In the SM, they result from symmetry demands, which will be explained in section 2.3 about the fundamental interactions of the Standard Model. These bosons are the mediators of the interactions and they are
The carriers of energy and momentum from an initial to a final state connected by an interaction.

### 2.3. The Fundamental Interactions of Particle Physics

To modern physics, there are four kinds of interactions in nature, the electromagnetic, weak, and strong interaction as well as the gravitational force. The first three will be explained in the following sections. The last one, gravitation, interacts between everything which can be assigned mass or therefore energy. The massless boson mediator is called the graviton and because of the tensor structure of general relativity, it has to carry spin 2 in units of $\hbar$. Gravitational interaction will not be considered further in this thesis because the relative strength of it compared to the other three interactions, see Tab. 2.1, is many magnitudes smaller and therefore of no practical interest. Moreover, gravitation is not part of the SM. The mediating bosons are so-called virtual particles, which do not have to be in accordance with Einsteins energy-momentum relation. This is only possible for a short amount of time, which is governed by the uncertainty relation $\Delta E \cdot \Delta t \approx \hbar$ of Heisenberg. This means that if the energy of this virtual particle is high, the time available for this state is very short leading to a short range ($\approx c \cdot \Delta t$) of the interaction. So if the mass and hence the energy of a mediator boson is very high, the lifetime of this particle is very short. This leads to the very short ranges of, for example, the weak interaction because of the high mass of its gauge bosons. In 1935, Yukawa derived the interaction potential for a model with a spinless mediator boson of mass $M$ \[26\],

$$V(r) \propto \frac{1}{r} e^{-\frac{r}{r_0}} \text{ with } r_0 = \frac{1}{M}, \quad (2.1)$$

where $\hbar = c = 1$ was set. This also explains the infinite range of the electromagnetic force and reproduces the Coulomb potential by setting $M \to 0$ corresponding to the rest mass of the photon. In addition, if this potential is used as a scattering potential, the scattering amplitude is calculated to be

$$F(q^2) = \frac{1}{-q^2 + M^2}, \quad (2.2)$$

where $q^2$ is the squared four-momentum transfer. This so-called propagator is also the general form of the amplitude for a spinless interaction. By taking the squared value of this and setting $M = 0$, it is also possible to reproduce Rutherford’s well-known scattering formula for a Coulomb potential.
2. The Standard Model of Particle Physics

2.3.1. Electromagnetic Interaction

The electromagnetic interaction is responsible for the coupling of charges to electric and magnetic fields. The theory of the electromagnetic interaction describes how these fields are created by their sources, which are electric charges and currents of moving charge, how they interact with other charges, and how they interact with each other. The first correct description of electromagnetism was given by Maxwell in the 1860s [27]. From these equations, it was possible to predict the phenomenon of electromagnetic waves, which are emitted by accelerated charges and were discovered by Heinrich Hertz in the 1880s [28]. Later, Planck and Einstein discovered the particle-like properties of light in their respective research regarding heat radiation [29] and the photoelectric effect [30] and it was possible to assign energy and momentum to these so-called photons. In addition to the formulation of special relativity, Einstein was able to show that the electric and magnetic field are part of one electromagnetic force and that these can transform from one to another depending on the reference frame of the observer. Later, the principles of quantum mechanics were used together with theory of electromagnetism to develop the theory of the hydrogen atom yet lacking the consideration of special relativity and spin. Dirac then developed his equation to solve the Lorentz-invariance problems of the Schrödinger theory. The Dirac equation [14] layed the foundation for the upcoming development of Quantum Electrodynamics (QED), the quantum field theory of the electromagnetic interaction. Feynman, Tomonaga and Schwinger [31–33] developed QED independently in the 1940s. With QED, it was possible to make predictions for many observables, which then were measured very precisely confirming the validity of the theory. In QED, the interaction between electric charges is mediated by the photon as a massless spin-1 gauge boson. The potential therefore has the form

\[ V_{em}(r) \propto -\frac{\alpha_{em}}{r}, \]  

(2.3)

with \( \alpha_{em} \) the electromagnetic coupling constant, defining how strong the photon couples to elementary electric charges. Contrary to its name, the electromagnetic coupling constant is not constant. It is a running coupling constant whose value changes dependent on the energy scale or distance of the interaction. The coupling is decreasing with increasing distance or decreasing energy.

2.3.2. Strong Interaction

The strong interaction is responsible for the stability of nuclei. The repelling electromagnetic force between the protons has to be compensated to reach stable nuclei. Yukawa proposed a boson with a much higher mass than the electron but a significant lower mass than the proton itself to be the mediator of this force. Later, this particle was discovered as the pion, the lightest known meson. Today however, it is known that this meson exchange is just an effective model of the strong force. The strong force interacts between quarks
and its mediator boson is the gluon, which is massless and carries spin 1. The strong interaction only couples to particles which carry a degree of freedom named colour as an additional quantum number. It was shown by experiment that there are three different colour states. The three states are consequently called red, green, and blue and every quark exists in one of these three colour states. In contrast to the electromagnetic interaction, the mediator boson of the strong force also carries the specific charge of the interaction. A gluon is assigned with a colour and anticolour, so that at every vertex of the strong interaction, the colour is conserved. Altogether, there are eight possible gluon states. The quark-antiquark colour-potential can be written in the form

$$V_{\text{colour}}(r) = -\frac{4}{3} \alpha_s \frac{r}{r} + k \cdot r.$$  \hspace{1cm} (2.4)

The coefficient $\alpha_s$ contains the coupling of the quarks to the gluon. The first part of the formula shows an analogous form to the Coulomb potential because the gluon is massless as the photon. The second term takes into account a special property of the strong interaction, the confinement. Because the gluons carry colour charge, they couple to other gluons as well, causing the colour field lines to form a string. A linear increase in potential energy by separating quarks consequently results in the case that the production of a new quark-antiquark pair needs less energy than further separating the quarks. This leads to the result that free particles can only exit in a colour-neutral state. Consequently, baryons have to consist of all three colours and mesons of colour and anticolour to be colourless or "white". Moreover, there is another very interesting feature of the strong interaction. Its running coupling constant has a very different behavior in comparison to the electromagnetic one. The smaller the distance or the higher the energy of the mediating boson, the smaller the coupling constant becomes. This behaviour is called asymptotic freedom [34, 35]. So for high energy collisions, the quarks can be considered almost free of their interaction with the cloud of gluons constantly interacting with the quarks and other gluons.

### 2.3.3. Weak Interaction

The weak interaction was discovered indirectly during the investigation of the nuclear $\beta$-decay. Starting from an instable nucleus, its decay was examined. Out of the decay, an electron was emerging. After the decay, the nucleus was lighter and had a charge number increased by one unit. The particles emerging from the decay should have characteristic energies and momenta due to energy and momentum conservation. The result, however, showed a completely different behavior. Instead, a spectrum of electron energies was measured. By keeping energy and momentum conservation, the only explanation was given by Pauli, who proposed a third particle taking part in the decay, which today is called the neutrino. Later, Enrico Fermi published his theory of the nuclear $\beta$-decay [36] as an effective approximation of today’s weak interaction theory. In his theory, an interaction vertex with four participating particles occurred. In todays theory of the weak interaction, the force is mediated by the heavy vector bosons W and Z. They have high
masses of around 80 GeV and 90 GeV, respectively, that are responsible for the short range and weakness of the interaction. The weak force couples to a property of the fermions named weak isospin. The nuclear β-decay is mediated by the W boson. The scattering amplitude using the model of Yukawa then has the form

\[ F(q^2) \propto \frac{g_W^2}{-q^2 + m_W^2}, \]  

including the coupling constant \( g_W \) of the particles to the W boson. In nuclear β-decays the four-momentum transfer \( q^2 \) is low, so this quantity can be neglected in comparison with the high mass of the W boson. Thus, the scattering amplitude leads to an effective point-like coupling with the Fermi constant \( G_F = g_W^2 / m_W^2 \) as the effective coupling constant. In addition, there is not only interaction by the charged W boson, which couples to the charged current, but there is also a coupling to the neutral current by the neutral Z boson. Every fermion in the Standard Model, leptons and quarks, can participate in weak interactions. Nevertheless, there is some qualitative difference in the couplings of the quarks and leptons. The eigenstates of the weak interaction for the charged leptons are the \( f_l \)avour eigenstates, which are also mass eigenstates. In contrast to this, the interaction eigenstates for the quarks are linear combinations of the \( f_l \)avour eigenstates, that means they are no mass eigenstates. This behavior can be described with the CKM-matrix [37, 38], which specifies the \( f_l \)avour composition of the interaction eigenstates of the quarks. Finally, it is possible to unify the electromagnetic and weak force to the electroweak force. Interestingly, the coupling of the heavy bosons of the weak interaction and the photon of the electromagnetic interaction have the same basis. The weakness of the weak force in comparison to the electromagnetic force is only due to the high masses of the W and Z boson. If the energy is high or the distance small enough, the coupling is qualitatively the same because the mass in the denominator of the propagator can be neglected. This is known as the electroweak symmetry. However, as explained above, this symmetry is broken because of the different masses. The mechanism behind the broken electroweak symmetry will be addressed in more detail in section 2.3.4.4.

2.3.4. Gauge Theory

The fundamental interactions in particle physics are derived from gauge symmetries, which makes the Standard Model a gauge theory. Local symmetries are imposed on the Lagrangian of the system. To fulfill these symmetries, gauge fields have to be introduced, which are consequently the mediators of the interaction. The form of these gauge fields is not completely determined by the equations, there is some degree of freedom to gauge them as the name suggests. The following more theoretical sections have been created under the usage of [18]. Natural units (\( \hbar = c = 1 \)) are applied throughout the following sections.
2.3. The Fundamental Interactions of Particle Physics

2.3.4.1. Quantum Electrodynamics

The Lagrangian for a spin-1/2 particle with mass \( m \) is given by

\[
\mathcal{L} = i \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi ,
\]

where \( \psi \) is a spinor, \( \bar{\psi} \) the adjoint spinor, \( \gamma^\mu \) the Dirac matrices, and \( \partial_\mu \) the partial derivatives with regard to the space-time coordinates. This Lagrangian reproduces the Dirac equation if the Euler-Lagrange equations are applied. The symmetry being imposed on this Lagrangian, in the following, is motivated from quantum mechanics, where the state of a system is invariant under a global phase transformation, also called U(1) transformation. While being invariant under the transformation of a global phase factor (\( \psi \rightarrow e^{i \theta} \psi \)), this Lagrangian is not invariant if the phase factor is local, meaning it is dependent on the space-time coordinates (\( \theta = \theta(x) \)). However, this Lagrangian can be changed to be invariant under this kind of transformation. Therefore, a new vector field \( A_\mu \) has to be introduced and incorporated into the Lagrangian. This is usually done in gauge theory by replacing the partial derivative with the covariant derivative, \( \partial_\mu \rightarrow D_\mu \), which accommodates an additional term. In case of the U(1)-symmetry group of QED

\[
D_\mu = \partial_\mu + iqA_\mu.
\]

Using this procedure, the resulting Lagrangian becomes

\[
\mathcal{L} = i \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi - q \bar{\psi} \gamma^\mu \psi A_\mu ,
\]

where the vector field transforms as \( A_\mu \rightarrow A_\mu + \partial_\mu \lambda \) with \( \lambda(x) = -\frac{1}{q} \theta(x) \) for local phase transformations. This is the well-known gauge freedom from electrodynamics. In addition, another part has to be included for the Lagrangian to be complete. In analogy to classical mechanics, this part is the kinetic part of the newly introduced vector field \( A_\mu \). This term can be taken as the kinetic term from the Proca equation for vector fields. As a result, the final Lagrangian of Quantum Electrodynamics can be written as

\[
\mathcal{L} = i \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi - q \bar{\psi} \gamma^\mu \psi A_\mu - \frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} ,
\]

where \( F_{\mu\nu} \) can be interpreted as the field strength tensor of electrodynamics. In this form, the Lagrangian contains a part which describes a free particle of mass \( m \), the interaction of this particle with the vector field (electromagnetic field), and the free propagation of the electromagnetic vector field. The Proca equation could also supply a mass term for the gauge field but in this case the Lagrangian would not be gauge-invariant anymore. Hence, the photon field has to stay massless.

2.3.4.2. Electroweak Unification

From experimental facts, for example the Wu-experiment [39], it is known that the weak interaction only couples to left-chiral fermion states and right-chiral antifermion states.
Thus, it is natural to split up the Lagrangian into chiral states. Every spinor can be split up in a left- and right-chiral part with two projection operators
\[
\psi_L \equiv \frac{1}{2} (1 - \gamma_5) \psi \quad \text{and} \quad \psi_R \equiv \frac{1}{2} (1 + \gamma_5) \psi \quad \text{with} \quad \psi = \psi_L + \psi_R.
\] (2.10)

Consequently, the symmetry group is chosen as \(SU(2)_L \times U(1)_Y\), which means that only the left-chiral part of the fermions, carrying weak isospin, and particles with hypercharge are affected by this transformation. The Lagrangian before demanding the local gauge invariance for this symmetry group has the form
\[
\mathcal{L} = i \bar{\psi}_L \gamma^\mu \partial_\mu \psi_L + i \bar{\psi}_R \gamma^\mu \partial_\mu \psi_R
\] (2.11)

while for one generation of leptons
\[
\psi_L \equiv \begin{pmatrix} \nu_L \\ l_L \end{pmatrix} \quad \text{and} \quad \psi_R \equiv l_R.
\] (2.12)

There is no right-handed neutrino in the Standard Model because those have never been measured in experiments until now. Furthermore, there is no mass term in this Lagrangian because it would not be gauge invariant. This is because mass terms couple left- and right-chiral states. The mass terms for the fermions are later introduced ad-hoc together with the Higgs mechanism. The main difference between QED and the electroweak theory is the symmetry group, which is non-Abelian in case of SU(2). This leads to significant differences in the required transformation behaviour of the introduced gauge fields and the form of the field strength tensors, i.e. the part describing free propagation. The Lagrangian has to be made invariant under the combined \(SU(2)_L \times U(1)_Y\) transformation. This can be done the same way as it was done for QED, however, there have to be four gauge fields. Three of those fields \((W^\mu_1, W^\mu_2, W^\mu_3)\) have their origin in the SU(2) transformation and the fourth one \((B^\mu)\) is to compensate for the hypercharge phase transformation. For those gauge fields, the free propagation term from the Proca equation also has to be included in the final Lagrangian.
\[
\mathcal{L} = i \bar{\psi}_L \gamma^\mu D_\mu \psi_L + i \bar{\psi}_R Y^\mu D_\mu' \psi_R - \frac{1}{16\pi} B_\mu \nabla_{\mu} B_{\nu} - \frac{1}{16\pi} W^i_\mu W_i^{\mu \nu},
\] (2.13)

with the two covariant derivatives
\[
D_\mu \equiv \partial_\mu + ig \frac{\tau^i}{2} W^i_\mu + ig' \frac{Y}{2} B_\mu \quad \text{and} \quad D'_\mu \equiv \partial_\mu + ig' \frac{Y}{2} B_\mu.
\] (2.14)

Here, the Pauli-matrices \(\tau^i\), which are the generators of the SU(2) group, emerge. Because the U(1)-hypercharge transformation is Abelian, the field-strength tensor \(B_{\mu \nu}\) is defined as
\[
B_{\mu \nu} = \partial_\mu B_\nu - \partial_\nu B_\mu,
\] (2.15)

whereas the field-strength tensors of the other three gauge fields
\[
W^i_\mu = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - ge^{ijk} W^j_\mu W^k_\nu
\] (2.16)
show the influence of the non-Abelian symmetry group SU(2). Two different coupling constants \( g \) and \( g' \) are associated with the two different gauge symmetries and their mediating gauge fields. The gauge fields are massless to retain the gauge symmetry. In addition, the introduced gauge fields are not the physical fields but those can be constructed as linear combinations of the gauge fields. Their masses will later be generated with the Higgs mechanism.

### 2.3.4.3. Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the underlying theory of the strong interaction. After the discovery of the colour degree of freedom, every quark flavour was assigned three possible colour states. The simplest way to accommodate for this change is to extend the already known Lagrangian for every colour state.

\[
\mathcal{L} = i \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi ,
\]

where

\[
\psi \equiv \left( \begin{array}{c} \psi_r \\ \psi_b \\ \psi_g \end{array} \right) \quad \text{and} \quad \bar{\psi} \equiv (\bar{\psi}_r \bar{\psi}_b \bar{\psi}_g) .
\]

In analogy to the Dirac Lagrangian for which U(1) symmetry was demanded, U(3) symmetry is imposed \((\psi \to U\psi)\), where \(U\) is an arbitrary unitary \(3 \times 3\) matrix. This transformation matrix can be split up into a phase transformation factor and another matrix, which belongs to the group SU(3). The only difference is that the determinant of this SU(3) matrix has to be 1. As in the previous case, this invariance must also hold in case of local transformations. The procedure for the local phase transformation has already been shown in section 2.3.4.1 on quantum electrodynamics leading to the interaction with the photon. Thus, only the local SU(3) transformation behaviour \((\psi \to S\psi)\) is of importance here. The SU(3) matrix \(S\) can also be written in terms of an exponential \(S = \exp (i \sum_{i=1}^{8} \lambda_i \Phi_i(x))\) where \(\lambda_i\) are the Gell-Mann matrices, which are the generators of the group of SU(3) matrices. Analogous to the additional vector field in the Dirac Lagrangian, the invariance of the Lagrangian requires the introduction of eight additional vector fields \(A^i_\mu\) within the covariant derivatives \(D_\mu = \partial_\mu + i g_s \lambda^i A^i_\mu\). For those gauge fields, the transformation rule is more complex than in the previous cases and for an infinitesimal transformation has the form

\[
A^i_\mu \to A^i_\mu - \frac{1}{g_s} \partial_\mu \Phi_i - \sum_{j,k=1}^{8} f_{ijk} \Phi_j A^k_\mu ,
\]

with the structure constants \(f_{ijk}\) of SU(3). Furthermore, the kinetic part of the gauge fields has to be added as well. The final Lagrangian for quantum chromodynamics then is

\[
\mathcal{L} = i \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi - g_s \bar{\psi} \gamma^\mu \lambda^i \psi A^i_\mu - \frac{1}{16 \pi} F^{\mu\nu} F^i_{\mu\nu} .
\]
This describes one quark flavour with its three colour states all having the same mass, their interaction with the eight gluons, and the free kinetic part of the gluons. In contrast to QED, the field strength tensor for QCD has an additional part

\[ F_{\mu\nu} = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu} - g \sum_{j,k=1}^{8} f_{ijk} A_{\mu}^{j} A_{\nu}^{k}, \]  

which describes the interaction between the gluon fields themselves.

### 2.3.4.4. Higgs mechanism

The Higgs mechanism was introduced to account for the missing masses in the electroweak theory while retaining the corresponding gauge symmetries. Essentially, the underlying principle is the introduction of a new field, whose ground state is not the zero-field configuration. This new minimum is determined spontaneously, which breaks the symmetry of the original Lagrangian. Usually, the Lagrangian is formulated in terms of the fields as small fluctuations around the ground state. If this is not the case, the ground state and the Lagrangian do not have to share a common symmetry. This effect is called spontaneous symmetry-breaking. After reformulating the Lagrangian in terms of deviations from the ground state, the reformulated Lagrangian does not have the same symmetry properties as the original one. For continuous global symmetries that are broken spontaneously, so-called Goldstone bosons appear as the consequence of the broken symmetries. These bosons are massless and carry spin 0. A typical Lagrangian, chosen to illustrate this behavior is

\[ L = \frac{1}{2} \partial_{\mu} \Phi \partial^{\mu} \Phi + \frac{1}{2} \partial_{\mu} \Phi_{1} \partial^{\mu} \Phi_{1} + \frac{1}{2} \partial_{\mu} \Phi_{2} \partial^{\mu} \Phi_{2} + \frac{1}{2} \mu^{2} (\Phi_{1}^{2} + \Phi_{2}^{2}) - \frac{1}{4} \lambda^{2} (\Phi_{1}^{2} + \Phi_{2}^{2})^{2}, \]  

where \( \Phi_{1}, \Phi_{2} \) are real scalar fields. This Lagrangian is invariant under rotations in the space of \( \Phi_{1}, \Phi_{2} \). The minimum-energy state can be found by minimizing the potential energy and is given by \( \Phi_{1,\text{min}}^{2} + \Phi_{2,\text{min}}^{2} = \mu^{2}/\lambda^{2} \). The Lagrangian can then be reformulated in terms of the deviations of the minimum state. Therefore, one state has to be picked from set of minima. After choosing \( \Phi_{1,\text{min}} = \mu/\lambda \) and \( \Phi_{2,\text{min}} = 0 \), the resulting Lagrangian is

\[ L = \frac{1}{2} \partial_{\mu} \eta \partial^{\mu} \eta + \frac{1}{2} \partial_{\mu} \xi \partial^{\mu} \xi + \frac{1}{2} \mu^{2} \eta^{2} + \frac{1}{2} \lambda^{2} \xi^{2} + \text{h.c.} \]  

where h.c. means higher corrections, \( \eta \equiv \Phi_{1} - \mu/\lambda \), and \( \xi \equiv \Phi_{2} \). This Lagrangian describes a Klein-Gordon field \( \eta \) with mass \( m_{\eta} = \sqrt{2}\mu \) and a free scalar field \( \xi \) without mass, which is the already mentioned Goldstone boson. This result has to be combined with local gauge invariance to generate the masses of the heavy vector bosons. To obtain the result, the two original fields \( \Phi_{1} \) and \( \Phi_{2} \) are written as one complex scalar field

\[ \Phi = \Phi_{1} + i\Phi_{2}. \]
The Lagrangian from Eq. 2.22 can then be reformulated as
\[ \mathcal{L} = \frac{1}{2} (\partial_{\mu} \Phi)^* (\partial^\mu \Phi) + \frac{1}{2} \mu^2 (\Phi^* \Phi) - \frac{1}{4} \lambda^2 (\Phi^* \Phi)^2. \] (2.25)

As in the previous sections, local gauge-invariance is demanded for the broken symmetry, which was a rotational symmetry SO(2), but can be regarded a U(1) symmetry in this form of the Lagrangian. The procedure of the introduction of the covariant derivative can be applied in this case to create invariance under local U(1) transformations as well. After doing this, the fields are rewritten in terms of fluctuations around one ground state. The process leads to a Lagrangian which again contains a massive scalar particle, the Higgs boson, a massless Goldstone boson, and a vector field (the gauge field) with a mass. In addition, it is possible to remove the gauge boson from the theory by using a specific gauge without changing the physics of the theory. The final Lagrangian then becomes
\[ \mathcal{L} = \frac{1}{2} \partial_{\mu} \eta \partial^\mu \eta - \mu^2 \eta^2 - \frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2} \left( q_{\mu}^2 \right)^2 A_\mu A^\mu + \text{h.c.}, \] (2.26)

containing the spin 0 Higgs field \( \eta \) and a massive gauge field \( A_\mu \). This result was obtained by starting with two additional degrees of freedom (\( \Phi_1, \Phi_2 \)), which were converted into one degree of freedom for the massive Higgs boson and one for the massive gauge field. From experiment, three massive and one massless gauge boson are known. Therefore, by starting with four degrees of freedom, three massive electroweak gauge fields and one massive Higgs boson could be generated. There is no degree of freedom left for the photon, which hence stays massless. Technically, this can be done by using two complex scalar fields, altogether describing four degrees of freedom, and combining them into a weak isospin-doublet
\[ \Phi \equiv \left( \begin{array}{c} \Phi^+ \\ \Phi^0 \end{array} \right), \] (2.27)

which is also affected by the electroweak gauge symmetry \( \text{SU}(2)_L \times \text{U}(1)_Y \). After introducing the covariant derivatives for the electroweak symmetry-group and by expanding around the vacuum expectation value \( \langle 0 | v \rangle \), the masses of the gauge bosons can be extracted from the Lagrangian. To do this, the physical states of the photon and the Z boson have to be recognized as linear combinations of \( B_\mu \) and \( W^0_\mu \). This linear combination can be parametrized with a rotation matrix in two dimensions with the Weingberg angle
\[ \cos \theta_W = \frac{g}{\sqrt{g^2 + g^2}}. \] (2.28)

The mass values of the heavy vector bosons are dependent on the electroweak coupling constants \( g, g' \) and the vacuum expectation value of the Higgs field \( v \):
\[ m_{W}^2 = \frac{g^2 v^2}{4} \quad \text{and} \quad m_{Z}^2 = \frac{v^2}{4} (g^2 + g'^2) \] (2.29)

The usual mass terms for the fermions which couple left- and right-chiral states are not invariant under the electroweak gauge symmetry. The fermion masses can, however, be
accommodated for within the Higgs mechanism assuming a Yukawa coupling term of the form $\bar{\psi}_L \Phi \psi_R$ and $\bar{\psi}_R \Phi^{\dagger} \psi_L$. The Yukawa-Lagrangian is

$$L_{\text{Yukawa}} = -f_l \left( \bar{\psi}_L \Phi \psi_R + \bar{\psi}_R \Phi^{\dagger} \psi_L \right).$$

(2.30)

After spontaneous symmetry-breaking and expansion around the vacuum expectation value, this becomes

$$L_{\text{Yukawa}} = -m_l (\bar{l}_L l_R + \bar{l}_R l_L) \left( 1 + \frac{H}{v} \right)$$

(2.31)

for a lepton $l$. This means that the leptons acquired mass $m_l = f_l v$ with $f_l$ being the Yukawa coupling for the respective lepton and $H$ the Higgs field. The second terms proportional to the Higgs field describe the interaction vertex of the lepton with the Higgs boson. The same can be done for the quarks but the already mentioned linear combination of the flavour/mass eigenstates to weak eigenstates, parametrized by the CKM-matrix, has to be taken into account.

### 2.4. Feynman Diagrams, Cross Sections and Decay Rates

The theory from the previous sections shows the mathematical and physical structure of the interactions between elementary particles. However, it does not provide the tools to directly calculate observables, which can be compared to experiment. An observable which can in general be measured and is commonly used in particle physics is the reaction rate. The reaction rate can be observed most generally in two kinds of reactions. Almost all particles decay unless there are conservation laws preventing the decay. For an ensemble of particles decaying, the decay rate is the interesting physical observable. The decay rate $\Gamma$ is related to the mean lifetime $\tau$ of a particle by

$$\Gamma = \frac{1}{\tau}.$$  

(2.32)

The decay of a particle is a quantum mechanical statistical process and it follows an exponential decay distribution. The probability density distribution for the decay times $\Delta t$ of a particle is

$$P(\Delta t) = \frac{1}{\tau} \exp \left( -\frac{1}{\tau} \Delta t \right).$$

(2.33)

If a particle can decay in $n$ different ways, the decay rate is the sum of the single decay rates $\Gamma = \Gamma_1 + \Gamma_2 + \ldots + \Gamma_n$. The fraction $B_i = \Gamma_i/\Gamma$ is then called the branching fraction for the decay $i$.

The second observable rate is the interaction rate for scattering processes. Essential for this rate is the cross section $\sigma$ of a process, defining an effective interaction area for a flux of incoming particles. The resulting interaction rate is given by

$$\frac{dN}{dt} = \sigma \cdot L.$$  

(2.34)
2.4. Feynman Diagrams, Cross Sections and Decay Rates

Figure 2.1.: The leading order Feynman diagram for the process $e^+e^- \rightarrow \mu^+\mu^-$ is shown. In addition, the terms corresponding to the lines of the Feynman diagram are given, which can be put together to evaluate the matrix element. The particles are assigned a particle spinor, the antiparticles are assigned a antiparticle spinor, the vertices contribute the electromagnetic coupling constant $e$, and the photon contributes its propagator. [41]

where $L$ is called the instantaneous luminosity. The luminosity is the effective flux of the accelerated particles normalized to unit area and unit time. Thus, the interaction rate is the product of the effective interaction area and the flux. The luminosity is an accelerator dependent quantity.

Both, the decay rate and the cross section, are proportional to the quantum mechanical transition amplitude $|M|^2 = |\langle \psi_f | \hat{U} | \psi_i \rangle|^2$, basically representing the probability for a transition from the initial state $|\psi_i \rangle$ to the final state $|\psi_f \rangle$ caused by an interaction potential $\hat{U}$. In addition, the decay rate and cross section are proportional to the phase space available for the process. The phase space is a measure of the number of possible final state configurations. Hereby, the final state configurations have to respect energy and momentum conservation. This is also known as Fermi’s Golden Rule, but it was first derived by Dirac in 1927 [40].

To calculate the matrix elements, a procedure was developed by Feynman using perturbation theory. The so-called Feynman rules are derived from the Lagrangian density of the theory. For every term in the expansion, a Feynman diagram can be drawn. These diagrams are a graphical representation of the corresponding process. Moreover, starting from the diagrams, a recipe can be used to write down the mathematical expression for the corresponding terms, which can then be calculated. A simple Feynman diagram for the process of two electrons annihilating into a photon, which subsequently creates a pair of muons, is shown in Fig. 2.1. Each Feynman diagram consists of outer lines, representing the initial and final state physical particles, and inner lines, representing virtual particles. For each line, a factor is added to the expression for the matrix element. In addition, there are vertices where different lines cross. The vertices describe the fundamental interaction between particles within a theory. The interactions are given by the terms in the Lagrangian density containing three or more fields. Depending on what kind of interaction
2. The Standard Model of Particle Physics

takes place, the corresponding coupling constant has to be taken into account. The number of vertices in a diagram determines the order or perturbation. If the coupling constant is small, higher orders can usually be neglected and the perturbative series converges. However, because of the running nature of the coupling constants, the coupling changes for different energy/distance scales. This has an important impact on QCD calculations. Because the strong coupling increases for low energies, the coupling constant is no longer small. This has the result that a perturbative expansion is no longer possible and can consequently not be used. This has a major impact, for example, in the description of parton showers or hadronisation, see chapter 5.

2.5. Hadron Collider Physics

In the last section, the idea of the cross section was explained. It was based on a clean initial state which transitions into another final state with a specific probability within an allowed phase space. However, at hadron colliders, the initial state is not known a-priori. At the LHC, protons are brought to collision. Protons are composite objects consisting of three valence quarks. Two of the valence quarks are up quarks and one is a down quark. The valence quarks interact with each other by emitting and absorbing gluons. The gluons can also interact by creating two virtual quarks or by separating into two gluons. This leads to a huge number of partons constantly interacting by the strong force inside the proton. The four-momentum transfer of such interactions is low, hence, the strong coupling constant is large and a perturbative calculation is impossible. Therefore, a calculation does not seem possible.

The processes with interesting physics properties are the ones with a high energy and momentum transfer, thus happening on a much shorter time and length scale than low energy processes. This leads to the fact that the additional interaction of the partons within the proton can be considered frozen from a time perspective. Seen from a length perspective, the resolution is high enough that the partons interact with each other and not with the soft QCD background in the proton. This separation leads to a factorization of the hard interaction process of the two colliding partons and the soft processes happening inside of the proton. This means that these two processes can be treated independently \[42\]. The energy scale defining this border is called the factorization scale \( \mu_F^2 \) and has to be chosen. The hard partonic cross section can be calculated as was explained in the previous section as a function of the momenta of the two partons. The influence of the soft interactions cannot be calculated with perturbation theory. Instead, one uses parton-density-functions (PDFs) to summarize (and parametrize) all soft interactions and effectively describe the state of the proton without actually calculating all the interactions. The resulting parton-density-functions then give the probability to find a specific parton \( i \) inside proton \( k \) (having momentum \( p_k \)) with a momentum \( p_i = x_i p_k \). The variable \( x_i \) is the Bjorken variable and describes the momentum fraction of parton \( i \) compared to the total proton momentum. The PDFs \( f_i(x_i, \mu_F^2) \) cannot be calculated from first principles and have to be
measured. This can be done, for example, with deep inelastic scattering [43]. In Fig. 2.2, two PDFs are shown. The total cross section for the process $pp \rightarrow X$ is then given by

$$\sigma_{pp \rightarrow X}(p_1, p_2) = \sum_{i,j} \int_0^1 dx_i \int_0^1 dx_j f_i(x_i, \mu_F^2) f_j(x_j, \mu_F^2) \hat{\sigma}(\hat{p}_i, \hat{p}_j \rightarrow X).$$  (2.35)

Here, $p_1, p_2$ represent the momenta of proton 1 and proton 2, the indices $i, j$ represent parton 1 and parton 2 with their respective momentum $\hat{p}_i, \hat{p}_j$ and $\hat{\sigma}$ is the partonic cross section.

Another challenge at hadron colliders like the LHC is the underlying event and pile-up. The underlying event describes the possible interactions from the additional partons in the protons. These interactions usually happen at low energy scales. The pile-up is caused by the fact that the protons are arranged in packages within the LHC. These packages contain of $O(10^{11})$ protons. When these bunches collide, not only one proton of each bunch collides, but several interactions happen at the same bunch crossing. Both, pile-up and the underlying event, cause a large number of additional objects in the detector.

The breakdown of the perturbative expansion at low energy scales also has an impact on the description of final state particles which interact by the strong force. Low energy gluon radiation or gluon splitting into quarks therefore have to be described in a non-perturbative
way. The same has to be done for hadronisation. Again, a border can be defined which separates the perturbative treatment from the regime where this is not possible. Hence, another factorization scale can be defined.
3. The Compact Muon Solenoid (CMS) at the LHC

This chapter is focused on giving a brief overview about the Large Hadron Collider (LHC) at CERN and the CMS detector following [45] and [46], respectively.

3.1. The LHC

The LHC is a hadron collider containing a ring with two beam pipes, which are equipped with superconducting magnets. Its circumference is around 27 km and it is situated between 45 m and 170 m below the surface. The tunnel was already used for the LEP collider [47] in the years before. Two transfer tunnels connect the LHC with the preacceleration facilities at CERN, see Fig. 3.1. The LHC is a particle-particle collider (protons or lead ions) causing the need for two beam pipes with one beam each in opposite directions. The LHC was designed for center-of-mass energies of up to 14 TeV. The protons and lead ions are planned to be accelerated to an energy of up to 7 TeV and 2.8 TeV, respectively. In addition, the protons are arranged in packages, which are called bunches. The energy of the protons is limited by the Lorentz force, which keeps the bunches on their trajectory. For a particle with momentum $p$ and charge $q$ on a trajectory with radius $R$ within a magnetic field $B$ the relation

$$p \propto qBR$$

(3.1)

holds. This means that the center-of-mass energy is limited by the strength of the magnets because of the given dimensions of the already existing tunnel.

Four major experiments are stationed at four different spots, where the beams can be brought to collision. The ALICE experiment [49] investigates physics with heavy ions and the properties of the strong interaction at very high densities and temperatures. The LHCb experiment [50] is focused on flavour physics of B hadrons. By measuring flavour oscillations, CP violation can be investigated. The aim of the two other so-called multi-purpose experiments, CMS [46] and ATLAS [51], is to search for new physics, the confirmation of the electroweak symmetry breaking in form of the Higgs mechanism, and precision measurements of already known quantities. There are two distinct ways to search for new physics, precision measurements and discovery measurements. Typically, precision measurements are done with lepton colliders because of the clean initial state and
3. The Compact Muon Solenoid (CMS) at the LHC

Figure 3.1.: The acceleration facilities at CERN with the LHC and the conducted experiments [48]. Starting from LINAC2 accelerator, the protons traverse the BOOSTER and are injected into the Proton Synchrotron (PS). From there, they are injected into the Super Proton Synchrotron (SPS). After, the protons finally enter the LHC.
the non-existing underlying event. Experiments with the aim of discovering new particles or reaching very high energies, in general use hadron colliders because of technical reasons. Still, the LHC tries to bring those two types together in the best possible way. The high energies are needed to create hypothetically new particles with masses too high to have ever been created with previous colliders. Furthermore, high luminosities are needed because the particles at the LHC are hadrons and, therefore, composite particles. This has the result that most of the collisions are low energy collisions giving rise to a large QCD background, which is not of much physical interest. Thus, a large number of collisions have to happen to get a sizable amount of high energy collisions. A measure of the number of collisions happening per second is the instantaneous luminosity

$$L = \frac{f n_b N_b^2}{4\pi \sigma_x \sigma_y}.$$  \hspace{1cm} (3.2)

Here, $f$ is the rotation frequency of the bunches, $N_b$ the number of proton bunches and $n_b$ the number of protons in one bunch. In addition, $\sigma_x, \sigma_y$ is the width of the beam. The LHC was designed to reach a luminosity of $L = 10^{34}\text{cm}^{-2}/\text{s}$ with $N_b = 2808$ and $n_b = 1.1 \cdot 10^{11}$. The LHC itself is only the last accelerator in a chain of multiple steps, accelerating and modifying the beam. First, a linear accelerator (LINAC2) accelerates the protons to 50 MeV. The next accelerating step by the Proton Synchrotron (PS) leads to 25 GeV protons. After that, the protons reach 450 GeV within the Super Proton Synchrotron (SPS). The final step is the acceleration by the LHC, which uses superconducting 400 MHz radio frequency cavities [52] for acceleration and compensation of the energy loss due to synchrotron radiation. The energy loss of an accelerated particle with mass $m$ and energy $E$ in one turn is given by

$$\Delta E \propto \frac{E^4}{m^4}.$$ \hspace{1cm} (3.3)

Therefore, an electron collider would have to deal with a immensely larger energy loss, which would have to be compensated in every turn. In the end, the proton bunches are brought to collision at the four experiment spots using quadrupole magnets.

### 3.2. The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is a multi-purpose detector operating at the LHC. It is situated approximately 100 m below the surface near the village Cessy in France between Lake Geneva and the Jura mountains. The detector is 21.6 m long with a diameter of 14.6 m and weighs around 14000 t. Among others, its purpose is to explore the TeV energy scale and search for new physics. Several problems and challenges had to be solved. First, the huge data rate of around $10^9$ events per second at design luminosity has to be reduced to approximately 100 events/s to be able to store and process the interesting data. This is done by the online event selection with triggers. In addition, the bunch crossing time of 25 ns leads to several requirements for the read-out electronics, detectors and the trigger
system to accommodate for multiple interactions in one bunch crossing or electric signals from earlier bunch crossings. This can be done with high-granularity detectors and a large number of read-out channels, where the single detector elements need a good time resolution. Moreover, it is expected that around 1000 charged particles result from one collision every 25 ns. This implies a very high flux of ionizing particles on the detectors and electronics. Therefore, they have to be hardened against radiation. To be able to achieve the set goals, several needs have to be fulfilled: The most important are the identification and reconstruction of muons, electrons, photons, and, in general, charged particles with the precise measurement of their momenta and energies. Furthermore, it has to be possible to measure the decays of heavy particles and possible missing transverse energy resulting out of them. The CMS detector and its components are shown in Fig. 3.2.

3.2.1. Coordinate System and Kinematic Quantities

In this section, the basic geometric and kinematic quantities used in collider physics are presented. The z-axis of the detector coordinate system is placed along the beamline and points towards the Jura mountains. The x-axis points towards the center of the LHC and the y-axis towards the surface. Since the CMS detector has a cylindrical form, it is more natural to use the polar angle $\theta$ with respect to the z-axis and the azimuthal angle $\phi$ in the $x$-$y$ plane. The colliding partons in general have different momenta along the beam axis resulting in a movement of the center-of-mass system along the beamline. Therefore, the quantities measured in the laboratory/detector coordinate system cannot be used as the center-of-mass system quantities. However, the Lorentz boost of the center-of-mass
system is mainly along the beamline. Thus, the transverse momentum

\[ p_T = \sqrt{p_x^2 + p_y^2} \]

is invariant regarding the boost from the laboratory system to the center-of-mass system. In addition, the pseudorapidity \( \eta \) is defined as

\[ \eta = -\ln \tan \frac{\theta}{2}. \]

With increasing pseudorapidity, the movement of a particle is increasingly more parallel to the beamline. Another measure of the geometrical separation of two particles is

\[ \Delta R_{12} = \sqrt{(\Delta \phi)^2 + (\Delta \eta)^2}, \]

where \( \Delta \phi = \phi_1 - \phi_2 \) and \( \Delta \eta = \eta_1 - \eta_2 \). With this quantity, a geometrical cone can be defined around the movement direction of a particle.

### 3.2.2. The Superconducting Solenoid

One of the main goals of the CMS detector is to precisely measure the momentum of highly energetic particles. This can be done by using again Eq. 3.1. The magnetic field and the momentum of the charged particles determine the radius of the trajectory. Hence, a very strong magnetic field is needed for highly energetic particles to gain enough bending in their trajectory that it can be measured with the tracking system. The solenoid within CMS reaches a magnetic field strength of 3.8 T, with a current of 20 kA due to the usage of superconducting NbTi cable, cooled to a temperature of 4.6 K. This leads to 2.6 GJ of stored energy in form of the magnetic field. The diameter of the solenoid is around 6.3 m and it has 12.5 m of length. To return the magnetic field lines, an iron yoke with a mass of 12500 t is built around the solenoid. The tracker, the electromagnetic and the hadronic calorimeter are situated within the solenoid.

### 3.2.3. The Inner Tracker

The tracker [53] is the innermost part of CMS. With only 4.4 cm distance to the beam pipe, it is closest to the nominal collision point. It has a length of 5.8 m and a diameter of 2.5 m and covers the \( |\eta| < 2.5 \) regime.

The tracker is supposed to enable the identification of particle tracks after the collision of two protons. As already mentioned, by reconstructing the track and its curvature, it is possible to measure the momenta of charged particles. Furthermore, by tracking the flight path backwards, the interaction vertex can be reconstructed. This is the point where the
two protons collided. Another purpose is the identification of so-called secondary vertices. Those vertices are created when a particle decays into lighter ones at a certain point in time and space. From this point the decay particles emerge and their tracks can be used to identify the secondary vertex. The high level trigger system also relies significantly on tracking information.

High granularity and fast response result in a large number of read-out channels with high power consumption of the responsible electronics. Thus, a powerful cooling procedure is required which of course needs cooling material. On the other hand, the aim for the tracker was to use as little material as possible, because effects like multiple scattering, bremsstrahlung, photon conversion, and nuclear interactions should be kept under control.

The influence of the radiation has already been mentioned in the introduction.

To fulfill all these requirements, the chosen detector design is based on silicon. Effectively, a silicon based detector works comparable to a diode in reverse direction. It uses the charge-depletion zone created between the two differently doped areas. When a charged particle traverses this zone it creates electron-hole pairs, which are separated by the reverse voltage leading to a measurable current.

The inner tracker has two major parts: Closest to the nominal interaction point is the pixel detector which has three barrel layers at radii between 4.4 cm and 10.2 cm and two disks as endcaps. The mounted pixels have a size of $100 \times 150 \mu m^2$ in $r - \Phi$ and $z$. The pixel detector reaches a spatial resolution of around $20 \mu m$ by taking into account the charge-spreading between different pixels due to the Lorentz drift of the electrons in the magnetic field of the solenoid.

Then follow the micro-strip detectors with 10 barrel layers and an outermost radius of 1.1 m starting with a size of $10 cm \times 80 \mu m$ and growing in size with increasing distance or decreasing particle flux, respectively. The strip detector also has several disks as endcaps, see Fig. 3.3. In the end, around $200 m^2$ of active silicon were implemented within the CMS tracker consisting of 1440 pixel and 15148 strip detector modules.

### 3.2.4. The Electromagnetic Calorimeter (ECAL)

The purpose of the electromagnetic calorimeter (ECAL) [54] is to measure the energy of electromagnetically interacting particles. The ECAL should cover a wide range of solid angle and have good electromagnetic energy resolution. In addition, a good diphoton resolution is required to measure the process $H \rightarrow \gamma \gamma$. This process has a very low branching fraction, but a very unique signature.

For the ECAL, two processes of electrons and photons within a material are of interest. Photons traversing a material can produce electron-positron pairs if they are in the vicinity of nuclei. Such a nucleus is necessary because the photons need another particle
3.2. The Compact Muon Solenoid

Figure 3.3.: The tracker system of the CMS detector with its modules in a schematic view [46] around the nominal collision point. The lines represent detector modules. In addition, the pseudo-rapidity, the distance in beam direction (z) and the radial distance is shown. The different sections of the silicon micro-strip detector are also visualized.

for conservation of energy and momentum. If electrons traverse a material, they emit photons in the electromagnetic field of a nucleus because they are accelerated (positively or negatively). This process is called bremsstrahlung. The cross section of both processes is proportional to $Z^2$, where $Z$ represents the atomic number of the nuclei. The result of these two processes is an electromagnetic shower. The electron-positron pairs created through pair creation emit photons and those again create more pairs. This leads to a cascade of electromagnetic particles, a so-called shower. The shower can be characterized by several quantities, for example the length of a shower, which is usually given in radiation lengths $X_0$. The radiation length is a quantity representing the typical energy loss of an electron due to bremsstrahlung in a specific material. Another quantity characterizing the shower is the Molière radius. It is a measure of the transverse expansion of the electromagnetic shower.

The electrons also deposit energy in the material by ionisation effects. These effects are used in the ECAL, for example, in form of a scintillating material. The energy which is absorbed by the material is converted into light by scintillation. The scintillation light and its energy can be detected with photosensitive detectors and from their signal, the energy deposited into the material, which is in relation with (almost proportional to) the original energy of the particle, can be inferred. The electric signal gets amplified before it is read out.

The ECAL of the CMS detector consists of 61200 lead tungstate ($\text{PbWO}_4$) crystals which are mounted in a hermetic way around the central barrel part. Lead tungstate is radiation tolerant, has a small radiation length and Molière radius of $X_0 = 0.89 \text{ cm}$ and $r_M = 2.19 \text{ cm}$, respectively, and 99% of the emitted light can be collected within 100 ns. As the inner tracker, the ECAL also has two endcaps containing 7324 crystals each, see Fig. 3.4. The contained oxygen in the crystals turns the material transparent and makes it a scintillator. Because of the metal proportion, the density of the material is around $8.3 \text{ g/cm}^3$. To detect
The Compact Muon Solenoid (CMS) at the LHC

3. The Compact Muon Solenoid (CMS) at the LHC

Figure 3.4.: The electromagnetic calorimeter is shown in a side view. The numbers represent the pseudo-rapidity $\eta$. Taken from [55].

the light which is emitted by the scintillating crystals, Avalanche photodiodes are used in the barrel. In front of the endcap crystals a more granular preshower detector is mounted. This is done to resolve boosted pion ($\pi^0$) decays to two almost collinear photons, which could otherwise be reconstructed as only one high energy photon. In the endcaps, vacuum phototriodes are used to detect the scintillation light.

Summarized, the ECAL was designed to have fast response, a granular structure, and to be hardened against radiation.

3.2.5. The Hadronic Calorimeter

The hadronic calorimeter (HCAL) [56] is supposed to measure the energy of particles mainly interacting by the strong force, for example, pions and kaons. These particles are typically found in jets emerging from quarks and gluons which result of an interaction.

The HCAL was built in alternating layers of absorber and detector. The detector is realized as a plastic scintillator. Hadrons can traverse a material much easier than the other particles. Consequently, the layers of absorber material are made of a high density material to trigger hadronic showering and also limit the length of the emerging hadronic shower to avoid leakage and thus lost energy. The energy of the particles in the shower are again converted to scintillation light within the detector layers. The light is then collected and wavelength-shifted by fibres and sent to photodetectors.

Analogous to the ECAL, the HCAL is also structured into barrel and endcap, which are composed of brass. Furthermore, there are two forward calorimeters, which are made of iron to account for the very high flux of hadrons in the region of high pseudorapidities. They are supposed to detect particles flying at very small angles relative to the beam line. Moreover, there is an outer part (after the solenoid) to account for particles that interact...
3.2. The Compact Muon Solenoid

Figure 3.5.: A quarter of the hadronic calorimeter is shown in a side view. The segmentation into a barrel (HB), endcap (HE), and forward region (HF) can be observed. Taken from [57].

the. The hadronic interaction length is about one order of magnitude greater than the radiation length.

It is important that the HCAL is almost hermetically closed, because this builds the basis to recognize missing transverse energy or momentum in the event. Only if every particle coming out of the collision is detected, it is possible to look at the sum of all measured transverse energy and momentum and conclude if the sum does not add up. This is especially important in searches for new physics, where BSM particles are expected to escape the detector without any interaction besides leaving traces of missing energy and momentum. The geometrical details can be seen in Fig. 3.5.

3.2.6. The Muon Chambers

The precise measurement of muons at the CMS detector is a very important task. Muons are often used to distinguish interesting processes from the overwhelming QCD background at the LHC. Furthermore, the measurement of the Higgs boson in the clear decay channel to four muons is one important goal of the CMS experiment because of its good mass resolution.

The information of the muon system [58] is used in three different tasks. These tasks are the identification of potential muons, the measurement of their momenta, and providing trigger information.

To measure the momenta precisely, a significant bending radius is needed. This is achieved by the high-field solenoidal magnet. The hits in the four so-called muon stations are
Figure 3.6.: The muon system is shown in a side view of a quarter of the CMS detector. Taken from [59]. It is possible to see the barrel and the endcaps as well as the different types of muon chambers.

used together with information of the inner tracker (within a curve fitting algorithm) to determine the flight path of the muons.

Muons penetrate material more easy than most of the other particles created at the LHC. Consequently, the detectors are placed at the most outside part of the experiment between the layers of the iron return yoke, which other particles are not expected to reach. This information can be used efficiently within the triggering system. The reason for the small energy loss of the muons in material is essentially their around 200 times higher mass in comparison to the electrons. This leads to smaller ionization losses for muons (in the interesting momentum regime above 1 GeV) described by the Bethe formula (see Passage of particles through matter in [23]). Also radiative losses are small because $\sigma_{\text{brems}} \propto 1/m^4$.

The muon system, see Fig. 3.6, consists of 1400 muon chambers of three different types: There are 250 drift tubes (DTs), 540 cathode strip chambers (CSCs), and 610 resistive plate chambers (RPCs). The DTs and a part of the RPCs are mounted in the barrel region. The CSCs and the other part of the RPCs are situated in the endcaps.

### 3.2.7. The Trigger System

As was already mentioned in the previous sections, the bunch crossing frequency of 40 MHz and the pile-up of about 25 simultaneous collisions leads to a huge amount of data (around $10^9$ proton-proton collisions per second). Not only is it impossible to store and process such a large amount of data, also the major part of these collisions is less interesting and dominated by QCD background. Hence, the amount of data has to be reduced significantly. This is the task of the trigger system [60]. The trigger system is designed in two parts: The first part is the Level-1 (L1) trigger, see Fig. 3.7. The L1 trigger is built with custom-designed programmable electronics, mainly FPGAs, reducing the
output rate below 100 kHz with very fast response and good time resolution. The L1 trigger is separated into local, regional and global components. The local triggers look for energy deposits in calorimeters or hits in the track segments of the muon system. The subsequent regional triggers combine the information of the local triggers and construct trigger objects, for example electron or muon candidates. The trigger objects are ranked dependent on energy or momentum and a quality measure. The last steps of the L1 trigger are the Global Calorimeter and Muon Triggers which choose the highest-rank objects and pass them to the L1 Global Trigger. The final decision to reject or accept the event has to be made within a maximum time of $3.2 \mu s$. The complete detector information is stored for this short amount of time in a pipeline. The second part is the High-Level trigger (HLT). The HLT is implemented in software running on a large processor farm. If the L1 Trigger accepted an event, the detector information is made available for the HLT. The HLT performs more sophisticated calculations than the L1 trigger comparable to the offline analysis. This leads to a final event rate of $O(100\text{-}1000$ events/s).

3.2.8. The Computing Grid

Although the trigger system reduces the amount of data significantly, the size of the remaining datasets is still huge. Therefore, a computing and storage system [61] named Worldwide LHC Computing Grid (WLCG) has been developed. The WLCG connects several computer and storage clusters around the world and provides real data as well as Monte Carlo (MC) data to scientists and students all over the world. The WLCG is structured in a hierarchical way where the different layers are called tiers. At CERN and in Budapest, the Tier 0 centre is situated. Its task is to store the raw detector data of all four LHC experiments and do a first reconstruction of the collisions. The next step are the thirteen Tier 1 centres, which are located around the world. There, at least one copy of the events is kept. In addition, data can be reconstructed again with, for example, new calibrations. The Tier 1 centres then distribute the events needed for specific analysis.
3. The Compact Muon Solenoid (CMS) at the LHC

tasks to the 160 Tier 2 centres, where the data can be examined by scientists and students within their specific analysis.

3.3. Object Reconstruction

Different physics objects are reconstructed from the signals of the single detector parts. The deposits of energy in the calorimeter system and the hits in the tracking system are combined by the CMS Particle Flow (PF) algorithm [62, 63] to reconstruct each particle created by the proton-proton collision. With the reconstructed particles, other more complex objects like jets and missing transverse energy are reconstructed.

3.3.1. Track and Vertex Reconstruction

To reconstruct the trajectories of the particles in the detector, an algorithm is applied which combines the hits recorded by the tracking system [64] to find the corresponding tracks. The first step is called local reconstruction. In this step the electric signals are converted into hits defining the position in each pixel or strip sensor. The next step is the track reconstruction. It uses the information of the reconstructed hits to estimate the momentum and the position of the charged particles. The hits have to be converted from their local sensor coordinate system to the global track coordinate system considering real positions and deformations of detector elements described by the alignment procedure [65]. The reconstruction of the trajectories needs high computing power and is performed by the Combinatorial Track Finder (CTF). This software is based on the Kalman filter [66]. First of all, a collection of reconstructed tracks is created by iterative tracking. This procedure first searches for tracks which are the least challenging to find. Then, the hits corresponding to those tracks are removed. This reduces the difficulty to reconstruct the remaining more challenging tracks. Each iteration of the software has four steps.

1. An initial collection of track candidates is created by considering only a low number of hits. The parameters of the track candidates are called seeds. This step is called seed generation.

2. The Kalman filter is used for track finding. The trajectories corresponding to the different seeds are extrapolated and additional hits compatible to this track are searched for.

3. The Kalman filter is used again for track-fitting. This is done to obtain the best estimates for the parameters of the track candidate.
4. A track selection is performed. Only tracks fulfilling the defined criteria are further considered.

Each of the iterations differs in the configuration of the track seed generation and in the track selection.

The tracks are then used to reconstruct the primary and secondary vertices in the event. Primary vertices are the interaction points of the hard parton interactions of the proton-proton collisions. Secondary vertices are created when particles decay and the decay products emerge from the nominal decay point. The vertices are found by looking for tracks which can be traced back to a common interaction point. Primary vertices usually have smaller impact parameters compared to secondary vertices. The vertices are found by a deterministic annealing algorithm [67]. To obtain the point of the vertex, a fit is performed using the assigned tracks. This is done by the adaptive vertex fitter method [68]. Secondary vertices use other track collections but are reconstructed in an analogous way.

3.3.2. Electron Reconstruction

Because electrons loose a significant amount of energy in the tracker material due to bremsstrahlung, the reconstruction of the hits and the estimation of the track parameters would not yield precise results. Therefore, the electrons are handled with a specific algorithm [69]. In the ECAL, the bremsstrahlung photons are mainly split over an azimuthal range but have the same pseudorapidity as the electron itself. There, patterns of such energy deposits are used to construct electron track seeds. Together with track seeds from the inner tracker which have low quality and are close to ECAL energy depositions, an electron track collection is created. However, the Gaussian-Sum filter method [70] is applied instead of the Kalman filter.

3.3.3. Muon Reconstruction

If a track has a momentum higher than a specific threshold, the search for additional hits is continued within the muon system. If a hit is found which is compatible to the track, this track is defined as a tracker muon [71]. In addition, seeds can be generated in the muon system and corresponding tracks are reconstructed with the Kalman filter. If muon system tracks and tracks from the inner tracking system can be matched, they are called global muon tracks.
3. The Compact Muon Solenoid (CMS) at the LHC

3.3.4. Calorimeter Clustering

An algorithm is used to cluster energy deposits in the calorimeters. The algorithm is optimized to assign the energy deposits to the respective particle and to distinguish different showers which are close to one another. Starting from seeds of maximum local calorimeter energy, topological clusters are built around them until no further calorimeter cells can be added. These objects are then called PF clusters. The energy of the cluster is then calculated by using the energy deposited in the cluster cells weighted by their distance to the original seed.

3.3.5. The Particle Flow Algorithm

The Particle Flow algorithm combines information from all subdetectors to identify and reconstruct all particles from the collision. It uses the large magnetic field, the high granularity of the calorimeters and the efficient tracking. The reconstruction of tracks and the clustering of calorimeter showers is linked together to blocks to connect every energy deposit of a particle. After a particle has been identified, its assigned track and calorimeter clusters are removed to simplify the event. The PF algorithm works in a specific order.

1. PF muons are identified with global muons if their momentum is in accordance with the inner track momentum measurement within three standard deviations.

2. The electron reconstruction procedure from section 3.3.2 is applied to obtain electron candidates. If those candidates fulfill certain criteria, they are identified as PF electrons.

3. The remaining tracks and calorimeter clusters which can be linked by the PF algorithm are tagged as charged hadrons given their momentum uncertainty is smaller than the expected calorimetric energy resolution.

4. The remaining ECAL and HCAL clusters are assigned with PF photons and PF neutral hadrons, respectively.

3.3.6. Jets

After the hard interaction took place, the partons in general radiate gluons or split into two other partons leading to a cascade of strongly interacting partons. Because of the running coupling constant of the strong interaction, quarks and gluons are confined. Consequently, the partons hadronize resulting in a vast number of strongly interacting particles moving along the initial parton direction. This group of particles moving roughly in the same
direction is called a jet. To really work with events in which partons hadronize, a jet definition has to be created which is consistent throughout several stages of the event. This enables to use jets on parton level for theoretical calculations, on particle level for example with PF candidates, and on detector level with calorimeter clusters. Furthermore, the definition has to be infrared safe, meaning that soft radiation or collinear splitting does not influence the clustering. In general, jets are defined by the procedure to construct them. Although there are several algorithms [72], all of them somehow cluster the objects corresponding to the applied event stage. The algorithms can be grouped in cone and sequential recombination algorithms. At the LHC, sequential recombination algorithms are commonly used because they are more robust. The jets which are used throughout this thesis were all clustered with the anti-\(k_T\) algorithm, which belongs to the class of sequential recombination algorithms. Within this algorithm, two distance measures are defined as

\[ d_{ij} = \min(p_{T,i}^{-2}, p_{T,j}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \]  

\[ d_{iB} = p_{T,i}^{-2}, \]

where \(p_{T,i}\) is the transverse momentum of particle \(i\) and \(R\) is a specified radius parameter corresponding to the approximate maximum cone radius of the jet in the \(\eta-\phi\) plane. Within this thesis, the radius parameter was chosen to be \(R = 0.4\). The distance \(d_{ij}\) corresponds to the distance between two particles and \(d_{iB}\) is a measure of the distance to the beamline. First of all, the minimum of all possible \(d_{ij}\) and \(d_{iB}\) is found. If the minimum is one of the \(d_{ij}\), the corresponding two particles are combined and their four-momenta are added. However, if one of the \(d_{iB}\) is the minimum, particle \(i\) is defined as a jet and removed from the particle collection under consideration. The distances then have to be recalculated and the procedure can be started again. This process is repeated until no particles are left. The algorithm is designed to first cluster particles with high transverse momentum with their neighbour particles. This ensures that the jets have an approximately round shape.

### 3.3.7. Missing Transverse Energy

The missing transverse energy is the missing momentum within the event to reach a vanishing total transverse momentum. The transverse momentum of the colliding protons is negligible compared to the momentum along the beam axis. Thus, significant deviations from a \(p_T\) balance are signs of particles escaping the detector without detection. Known particles which show this behaviour are neutrinos due to their almost vanishing interaction. The missing transverse energy can be calculated as

\[ \vec{E}_T = - \sum_{\text{visible}} p_T = \sum_{\text{invisible}} p_T. \]  

Since a lot of theories beyond the Standard Model (BSM) predict particles only interacting weakly and which consequently would escape the detector, the measurement of the transverse momentum imbalance is a very important part of the corresponding searches.
3. The Compact Muon Solenoid (CMS) at the LHC

3.3.8. b-tagging

The final state of $\bar{t}tH(b\bar{b})$ contains four jets originating from bottom-type quarks. Naturally, it is crucial to detect those jets. To do this, a special property of particles containing b-quarks is exploited. The b-quarks can only decay to lighter quarks through decay channels which are suppressed by off-diagonal CKM matrix elements. This suppression leads to longer lifetimes of b-hadrons compared to typical weak decays. The longer lifetimes and the consequently longer flight distances enable a displacement measurement of the decay vertex from the primary vertex which is of $O(\text{mm})$.

To perform the task of recognizing jets from b-quarks, the combined secondary vertex algorithm (CSV) is used [73]. This algorithm uses the reconstructed secondary vertices, its properties and associated tracks to evaluate a multivariate method. The output of the method, called CSV value, is a value between zero and one. The closer the value is to one, the more probable the jet originates from a b-quark. A cut is defined to classify the considered jet as a b-jet if its CSV value is higher than the defined cut. In addition, three typical working points are defined, which are called loose, medium, and tight. These working points usually correspond to mistag-rates of 10%, 1%, and 0.1%. The mistag rate is the fraction of jets which is incorrectly tagged as a b-jet at the considered WP. Throughout this thesis, the medium working point is used to classify b-jets.
4. Search for the Associated Production of a Higgs Boson and a Top-Quark-Antiquark Pair

In this chapter, a short summary about the production of a Higgs boson in association with a top-quark-antiquark pair and the corresponding search is given. The process and its constituents are explained, the topology of the final state is defined, and the major background processes are introduced.

4.1. Top Quark, Higgs Boson, and their Associated Production

The Higgs boson can be produced in association with a pair of top quarks if the energy of the collision is high enough to create the masses of the three heavy particles. This specific process is of great interest because the coupling of the Higgs boson to the top quark can be examined in a model-independent way. As was shown in section 2.3.4.4 (Higgs mechanism), the mass of a fermion is proportional to the coupling of the Higgs boson to the fermion. Because of the high mass of the top quark, the coupling is expected to be of order $O(1)$, distinguishing it significantly from the other fermions. The Yukawa coupling of the top quark has major influence on the Higgs scalar self coupling and therefore has an influence on the scale at which new physics is expected to be necessary [8].

4.1.1. The Top Quark

The top quark was discovered by the CDF and DØ collaboration at the Tevatron in 1995 [74, 75]. A combination of mass measurements results in

$$m_{\text{top}} = 173.21 \pm 0.51 \pm 0.71 \text{ GeV},$$

see [23]. The top quark carries an electric charge of $q = 2/3e$, weak isospin $T_3 = 1/2$ and colour charge. Therefore, the top quark can interact by any kind of interaction. Top-quark pairs are produced at the LHC by the strong interaction (mainly by gluon-gluon fusion) whereas single top quarks are produced by the weak interaction. The top quark
decays almost exclusively into a bottom quark and a W boson because the corresponding CKM matrix element is practically one. Its short lifetime is of great importance for the properties investigated in this thesis. The decay of a top quark is classified according to the decay of the created W boson, which can decay leptonically or hadronically. The W boson decays leptonically in around $\frac{1}{3}$ and hadronically in around $\frac{2}{3}$ of all cases. Top-quark pairs can thus decay full hadronically in approximately $\frac{4}{9}$ of all cases. The semileptonic (one hadronically decaying top, one leptonically top) decay happens with a probability of about $\frac{4}{9}$ and the dileptonic decay (two leptonically decaying tops) with a probability of $\frac{1}{9}$. Top-quark pair production has an important role within the analysis of the associated Higgs production with a top-quark pair because it is by far the largest background process.

4.1.2. The Higgs Boson

The Higgs boson is a massive scalar particle predicted by the Higgs mechanism. It couples to massive elementary particles. Therefore, it couples favourably to heavy particles like the top quark or the heavy electroweak vector bosons. In addition, the Higgs boson couples to itself. In 2012, CMS and ATLAS discovered a particle \cite{atlas_higgs, cms_higgs} compatible with the properties of the long-sought Higgs boson \cite{bourjaily1979possibility, gildener1981tree}. Today, its mass has been combined \cite{higgs_mass_combination} to

$$m_{\text{Higgs}} = 125.09 \pm 0.24 \text{ GeV}. \quad (4.2)$$

There are several decay channels for the Higgs boson in the Standard Model. In Fig. 4.1, the branching fractions are shown dependent on the Higgs-boson mass. The coupling and hence the decay rates differ for fermions and vector bosons. The decay rate for the decay
into two fermions is given \[76\] by

$$\Gamma_f = N_c \frac{G_F m_f^2}{4 \sqrt{2} \pi} m_H \beta_f^3 \quad (4.3)$$

at leading order, where \(N_c\) is the number of available colours (three for quarks and one for leptons) and \(\beta_f\) is the relativistic velocity of the fermions. Because the top quark is more massive than the Higgs boson, the decay into two bottom quarks is the most probable. The decay into leptons is additionally suppressed by the colour factor of one.

The decay rate into vector bosons is given \[76\] by

$$\Gamma_V = N_V \frac{G_F m_H^3}{16 \sqrt{2} \pi} (1 - 4x + 12x^2) \beta_V \quad (4.4)$$

The factor \(N_V\) is 2 for the decay into two \(W\) bosons and 1 for the decay into two \(Z\) bosons, \(x = m_V^2 / m_H^2\), and \(\beta_V\) is the relativistic velocity of the vector bosons. Since the mass of two vector bosons is higher than the Higgs-boson mass, one of the vector bosons is a virtual particle. The decay into two \(Z\) bosons is a very clean channel because of its final state containing 4 charged leptons. Moreover, leptons can be measured with a high resolution, thus enabling a precise mass and spin measurement. This channel is one of the channels in which the Higgs boson was discovered by detecting its mass peak within the spectrum of the invariant mass of the four leptons. Besides the direct coupling of the Higgs boson to its decay products, there are also loop-induced decays in which the Higgs boson couples to a loop of heavy particles. This is the case, for example, in the Higgs-boson decay into two photons. The particles in the loop are the top quark and \(W\) boson, which are both charged particles. This means that the photon can interact with them. This is also one of the channels in which the Higgs boson was discovered. The signal of two photons can be distinguished from the background and the electromagnetic calorimeters of the LHC experiment are designed to provide good diphoton mass resolution. Consequently, it was possible to measure the mass peak in this spectrum as well.

The production of the Higgs boson at the LHC is possible in a lot of different ways. Because the partons initiating the hard collision have a low mass or no mass at all, the coupling of the Higgs to them is weak. Hence, the Higgs boson is produced only indirectly. In Fig. 4.2, the leading order Feynman diagrams for the dominant production processes are shown and will be summarized in the following:

**Gluon-Gluon Fusion:** This production process dominates the production cross section because of the properties of the proton PDFs, cf. section 2.5. The gluons couple to a top-quark loop by the strong interaction. Because of the high top quark mass, the Yukawa coupling of the top quark dominates the other known particles in the loop. Unknown beyond the Standard Model particles could also have a contribution in this loop creating a model dependence.

**Vector Boson Fusion:** Two initial state quarks emit one electroweak vector boson each, which produce the Higgs boson. In the final state, two quarks and the Higgs boson
4. Search for the Associated Production of a Higgs Boson and a Top-Quark-Antiquark Pair

[Diagram: Higgs–strahlung and Vector boson fusion]

Figure 4.2.: Example leading order Feynman diagrams of the four major production processes of the Higgs boson are shown [77].

[Diagram: Gluon–gluon fusion and Associated production with $Q\bar{Q}$]

Figure 4.3.: The four leading order Feynman diagrams for $t\bar{t}H$ production [78].

emerge. This process depends on the coupling strength of the Higgs boson to the electroweak bosons.

**Higgs Strahlung:** Two quarks create a highly virtual W or Z boson. This virtual vector bosons then radiates a real Higgs boson and is converted to a real final state vector boson.

**Associated Higgs Production:** In this process, the Higgs boson is produced together with two additional quarks. Relevant for this thesis is the production in association with a top-quark-antiquark pair. Because of the additional high-mass final state particle in $t\bar{t}H$, this process needs a more energetic hard interaction compared to $t\bar{t}$ production. The partonic center-of-mass energy has to be at least around 500 GeV. The partons can be two gluons or a quark and antiquark, see Fig. 4.3 for LO diagrams. The top quark again couples directly to the Higgs boson by the Yukawa coupling. This means that at leading order the cross section is proportional to the top Yukawa coupling squared opening the possibility for its measurement. For experiments probing the top Yukawa coupling, this process has an advantage compared to the gluon-gluon fusion. Because there are no loops involved where BSM particles could have an
4.2. Final State Topology

In the final state considered in this thesis, the Higgs boson decays into a $b\bar{b}$ pair and the top quark-antiquark pair decays in the semileptonic channel. This channel combines the largest Higgs-boson branching fraction into two bottom quarks with the sizable semileptonic branching fraction of the top-quark pair. In addition, the lepton can be used to discriminate against QCD multijet events by requiring exactly one reconstructed lepton. Moreover, missing transverse energy from the neutrino in the final state is detected. The Higgs boson and the decaying top quarks contribute four bottom quarks in addition to the two quarks coming from the $W$ boson decay. A Feynman diagram showing the production of this final state can be found in Fig. 4.5. This final state is from now on referred to as $t\bar{t}H(b\bar{b})$.
4. Search for the Associated Production of a Higgs Boson and a Top-Quark-Antiquark Pair

![Feynman diagrams showing Higgs boson and top-quark production](image)

Figure 4.5.: One of the Feynman diagrams for $t\bar{t}H$ production (left) where the Higgs boson decays into a pair of bottom quarks and the top-quark pair system decays semileptonically [13]. In addition, a Feynman diagram of the irreducible $t\bar{t}b\bar{b}$ background (right) is shown [79].

4.3. Backgrounds

Although the event selection, cf. chapter 6, is designed to discriminate the $t\bar{t}H(b\bar{b})$ signal from several background processes, background events with final states similar or comparable to the signal will usually be selected as well. In the following sections, three classes of background events will be explained. These events are top-quark-antiquark events with additional heavy-flavour and light-flavour jets as well as additional background events.

4.3.1. Top-Quark Pair Production with two Additional Heavy-flavour Jets

This process, consisting of a top–quark pair with two additional bottom quarks, has exactly the same final state as the signal process, cf. Fig. 4.5 on the right side, and is therefore an irreducible background. NLO calculations [10, 11] predict the cross section of $t\bar{t}b\bar{b}$ to be approximately eight times larger than $t\bar{t}H(b\bar{b})$ [80]. This (irreducible) background can only be discriminated by exploiting physical differences showing themselves in distributions of observables. This thesis is focused on such a physical difference to separate $t\bar{t}b\bar{b}$ from $t\bar{t}H(b\bar{b})$. Spin correlations of the top-pair system are investigated regarding their use in the separation of $t\bar{t}b\bar{b}$ and $t\bar{t}H(b\bar{b})$.

4.3.2. Top-Quark Pair Production with Additional Light-flavour Jets

In general, $t\bar{t}$ events with additional light flavour or c-jets can be discriminated by requiring more than two b-tags. Since the inclusive cross section of around 830 pb for the $t\bar{t}$ process [81] is $O(1000)$ times larger than the $t\bar{t}H$ cross section, even a small misidentification rate of non b-jets, identified as b-jets, leads to a significant amount of background.
4.3. Backgrounds

4.3.3. Additional Backgrounds

Finally, a few other backgrounds contribute to the events selected in the tH(bb) analysis.

The production of single top quarks can result in two bottom quarks, a lepton, missing transverse energy, and additional jets. Single top quarks can be produced in several channels. The t-channel production has the largest contribution to the tH(bb) background.

Another small contribution is due to t-channel production in association with a vector boson. Although processes like ttZ or ttW have comparable production cross sections to tH [82] and can have comparable final states by decays like Z → b¯b or W → bq′, the corresponding branching fractions are smaller or CKM suppressed. Thus, this background is neglected in this thesis.

Furthermore, the production of a vector boson decaying leptonically in association with additional jets is a minor background. A W boson decaying into an electron or muon and its associated neutrino, together with four additional jets, will pass the selection. However, the cross section of such processes decreases strongly with increasing number of jets [83].

In addition, a small contribution can be attributed to the pair production of vector bosons. If a Z boson decays into two bottom quarks or light jets are misidentified as b-jets in association with one lepton, the final state looks comparable to the signal. However, the cross section is much smaller than the tt + jets cross section [84]. This background is neglected as well in this thesis.

The remaining QCD multijet background can be severely reduced by requiring an isolated lepton and several b-tags. For this thesis, this background is neglected.
5. Event Simulation

The data recorded by the CMS detector is supposed to be compared to theoretical expectations given by the Standard Model. To compare detector data with expectations on parton level, several effects have to be taken into account. The pure parton level physics is influenced by the parton-density-functions describing the probability to have a specific set of partons in the initial state. Besides the matrix element of the hard parton interaction, the showering and hadronisation of strongly interacting partons in the initial and final state has to be considered. Finally, the stable particles pass through the detector and interact with its material. Here, detector resolution as well as efficiency and acceptance effects impact the recorded data. Summarized, the relation between the initial-state partons and the recorded detector data is described by a complex convolution of several effects and influences. To avoid a deconvolution of the recorded data to obtain the parton level results, a Monte Carlo (MC) simulation is performed considering all the effects from the initial partons to the data recorded by the CMS detector.

5.1. Event Generation

The generation of proton-proton collisions at the LHC using MC simulation starts by exploiting the QCD factorization theorem which allows to separate the processes happening inside the protons from the hard partonic interaction, cf. section 2.5. For a specific process to be generated, the matrix element (ME) has to be computed. This is usually done by dedicated ME generators like “Powheg” [85, 86] or “MG5_aMC@NLO” [87]. Those two generators work at NLO perturbation theory. There are leading order generators as well. The LO part of “MG5_aMC@NLO” will be denoted by “MadGraph” from now on. The initial parton state is sampled with an MC method from the PDFs provided, for example, by the “MMHT” [88], “NNPDF” [89], and “CTEQ” [90] groups. To obtain the probability distribution of the final states, an integration over the final-state phase space is performed. The integrals are usually computed with MC integration techniques because of their high dimensionality. From the obtained probability distribution, the final states are sampled. The next step is the parton shower (PS) algorithm. This algorithm simulates gluon radiation and quark-antiquark pair production of initial- and final-state partons leading to showers of strongly interacting partons. Often used parton shower programs are “Pythia” [91] and “Herwig” [92, 93]. Difficulties can arise by interfacing the ME generators with partons shower programs. A $t\bar{t} + 1$ jet event can be generated by a NLO matrix element generator...
Table 5.1.: A summary of the simulated MC datasets used in this thesis. The generated processes and their theoretical cross sections $\sigma_{\text{theo}}$, the number of events in the datasets, and the generators are listed. Furthermore, additional information is given in the “Comment” column. The “$W + \text{jets}, W \rightarrow l\nu$” datasets are divided by $H_T$, which is the scalar sum of the transverse momenta of all jets in the event.

<table>
<thead>
<tr>
<th>Process</th>
<th>Comment</th>
<th>Generator</th>
<th># Events</th>
<th>$\sigma_{\text{theo}}$ [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}H, H \rightarrow b\bar{b}$</td>
<td>$m_H = 125$ GeV</td>
<td>Powheg</td>
<td>3993304</td>
<td>0.29</td>
</tr>
<tr>
<td>$t\bar{t}H, H \rightarrow \text{non } b\bar{b}$</td>
<td>$m_H = 125$ GeV</td>
<td>Powheg</td>
<td>3975854</td>
<td>0.21</td>
</tr>
<tr>
<td>$t\bar{t}H, H \rightarrow b\bar{b}$</td>
<td>$m_H = 125$ GeV</td>
<td>MG5_aMC@NLO</td>
<td>9827285</td>
<td>0.29</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>semileptonic</td>
<td>Powheg</td>
<td>157387260</td>
<td>364.31</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>dileptonic</td>
<td>Powheg</td>
<td>75383000</td>
<td>87.33</td>
</tr>
<tr>
<td>$t\bar{t}b\bar{b}$</td>
<td>inclusive</td>
<td>MG5_aMC@NLO</td>
<td>5953262</td>
<td>1</td>
</tr>
<tr>
<td>$t$</td>
<td>t-channel</td>
<td>Powheg</td>
<td>3279200</td>
<td>45.34</td>
</tr>
<tr>
<td>$\bar{t}$</td>
<td>t-channel</td>
<td>Powheg</td>
<td>1682400</td>
<td>27.98</td>
</tr>
<tr>
<td>$t$</td>
<td>tW-channel</td>
<td>Powheg</td>
<td>998400</td>
<td>35.60</td>
</tr>
<tr>
<td>$\bar{t}$</td>
<td>tW-channel</td>
<td>Powheg</td>
<td>985000</td>
<td>35.60</td>
</tr>
<tr>
<td>$W + \text{jets}, W \rightarrow l\nu$</td>
<td>$100$ GeV $&lt; H_T \leq 200$ GeV</td>
<td>MadGraph</td>
<td>37778906</td>
<td>1345</td>
</tr>
<tr>
<td>$W + \text{jets}, W \rightarrow l\nu$</td>
<td>$200$ GeV $&lt; H_T \leq 400$ GeV</td>
<td>MadGraph</td>
<td>19851624</td>
<td>359.70</td>
</tr>
<tr>
<td>$W + \text{jets}, W \rightarrow l\nu$</td>
<td>$400$ GeV $&lt; H_T \leq 600$ GeV</td>
<td>MadGraph</td>
<td>7432746</td>
<td>48.91</td>
</tr>
<tr>
<td>$W + \text{jets}, W \rightarrow l\nu$</td>
<td>$600$ GeV $&lt; H_T \leq 800$ GeV</td>
<td>MadGraph</td>
<td>18133257</td>
<td>12.05</td>
</tr>
<tr>
<td>$W + \text{jets}, W \rightarrow l\nu$</td>
<td>$800$ GeV $&lt; H_T \leq 1200$ GeV</td>
<td>MadGraph</td>
<td>7854734</td>
<td>5.50</td>
</tr>
<tr>
<td>$W + \text{jets}, W \rightarrow l\nu$</td>
<td>$1200$ GeV $&lt; H_T \leq 2500$ GeV</td>
<td>MadGraph</td>
<td>7063909</td>
<td>1.32</td>
</tr>
<tr>
<td>$W + \text{jets}, W \rightarrow l\nu$</td>
<td>$2500$ GeV $&lt; H_T$</td>
<td>MadGraph</td>
<td>2507809</td>
<td>0.03</td>
</tr>
</tbody>
</table>
or by the combination of a $t\bar{t}$ event with the parton shower radiating another parton. Therefore, merging algorithms were developed to use the optimal description for every part of the event. Hard emissions allowing a perturbative treatment are described better by the ME generators whereas soft or collinear emission have a better description with PS generators. A comparison between several merging algorithms can be found in [94]. After the parton shower reached a cut-off energy scale, the hadronisation is performed by phenomenological models since a description by perturbative theory is not possible anymore. The Lund string model [95], used in Pythia, and the cluster hadronisation model [96], used in Herwig, are two well-known models. Finally, the underlying event and pile-up interactions also have to be simulated. The underlying event is taken care of by the parton shower program as well. The influence due to pile-up is simulated by adding additional proton-proton collisions to the event. After the hadronisation, the interaction of the final state particles with the detector is simulated with the “GEANT4” framework [97]. The simulation includes the interaction of the particles with the material, the influence of the magnetic field, and the electronics. The results of all the explained steps are then stored in the same way as real data allowing an analogous analysis of simulated and measured data.

5.2. Monte Carlo Datasets

For this thesis, simulated datasets at a center-of-mass energy of 13 TeV, centrally produced by the CMS collaboration, were used. The datasets consist of the signal process $t\bar{t}Hb\bar{b}$ and the major backgrounds, see Tab. 5.1. The number of events generated for the different processes is much higher than the number of expected events in the data. This is done to reduce the statistical fluctuations and uncertainties as well as to provide enough events for training and testing of multivariate classifiers. To compare the simulated distributions to data and use the classifier trained with MC events on data, the MC samples have to be reweighted to the number of events expected in data. The expected number of events is calculated with the measured integrated luminosity $L_{int}$ and the theoretical cross section at least to NLO. Hence, the MC events in a sample containing $N_{gen}$ events for a process with a cross section of $\sigma_{\text{theo}}$ have to be reweighted with

$$w_{\sigma} = \frac{\sigma_{\text{theo}} \cdot L_{\text{int}}}{N_{\text{gen}}}.$$  (5.1)

If the considered dataset is generated in a specific exclusive decay channel, the cross section has to be replaced with the product of the cross section and the branching fraction of the considered decay channel. Some generators like “MG5_aMC@NLO” use negative event weights. For those MC samples, the number $N_{\text{gen}}$ has to be replaced with the difference of events with a positive and negative weight, $N_{\text{pos}} - N_{\text{neg}}$.

A major part of the background datasets were generated with the matrix element generator “Powheg” (v.2). For the $t\bar{t}H$ samples, the Higgs-boson mass was set to 125 GeV. The mass
5. Event Simulation

of the top quark was set to 172.5 GeV for all processes containing of top quarks. The events containing a W boson with additional jets were generated with “MadGraph”. An additional $t\bar{t}H(bb)$ and $t\bar{t}b\bar{b}$ sample was generated with “MG5_aMC@NLO”. The structure of the proton was described by the PDF set “NNPDF3.0”. In addition, “Pythia” (v.8.2) with the tune “CUETP8M1” [98, 99] was used for the parton shower and the hadronisation. The tune of the parton shower is used to describe the underlying event.

5.3. Additional Jet-Flavour Identification in Top-Quark Pair Events

The $t\bar{t}$ events usually contain additional jets with different flavours. With a dedicated tool [100], it is possible to further classify the main background ($t\bar{t}$+jets) dependent on the additional jet flavours. This tool uses generator information to search for jets containing b-hadrons or c-hadrons. Subsequently, it is checked if the corresponding b- or c-quarks originate from the top decay. Only if this is not the case, the jets are used as additional jets to classify the event further in one of the following categories.

$t\bar{t} + b\bar{b}$:
Two additional b-jets, each with at least one b-hadron, were found. These are the background events addressed in this thesis. This category will be denoted by “$t\bar{t}b\bar{b}$” from now on.

$t\bar{t} + b$:
One additional b-jet with exactly one b-hadron was found.

$t\bar{t} + 2b$:
One additional b-jet with at least two b-hadrons was found.

$t\bar{t} + c\bar{c}$:
Any number of additional c-jets with at least one c-hadron and no additional b-jets were found.

$t\bar{t} + l\bar{l}$:
No additional b-jets or c-jets were found.

5.4. Corrections to Simulated Data

Since the MC samples are generated with several approximations and assumptions, it is expected that the simulated data cannot describe the real data without shortcomings.
Thus, the simulated data have to be corrected dependent on the observations of the real data. This is usually done with scale factors which can be applied on physics objects or events.

### 5.4.1. Jet Energy Resolution

It has been observed that differences occur considering the jet energy resolution in data and MC simulation. The jet energy resolution in MC events was observed to be smaller than in events from real data. Consequently, dijet events and the transverse momentum asymmetry

\[
\mathcal{A} = \frac{p_{T,j1} - p_{T,j2}}{p_{T,j1} + p_{T,j2}}
\]

are used to derive scale factors dependent on \( \eta \) for the transverse momentum of the simulated jets [101]. This can be done because the width of the distribution of \( \mathcal{A} \) is directly related to the jet energy resolution. The scale factors \( f_{\text{JER}} \) are then applied according to the following formula:

\[
p_{T,\text{corr}} = p_{T}^{\text{reco}} + f_{\text{JER}}(p_{T}^{\text{reco}} - p_{T}^{\text{gen}})
\]

The corrected transverse momentum \( p_{T,\text{corr}} \) is obtained by shifting the originally reconstructed jet \( p_{T} \) with a fraction of the difference between reconstructed jet \( p_{T} \) and the \( p_{T} \) of the particle level jet obtained from generator information.

### 5.4.2. CSV Reweighting

The output of the CSV algorithm also indicates differences between data and simulation. Since the \( t\bar{t}H(\bar{b}b) \) final state consists of four b-quarks and the CSV value determines if a jet is considered as a b-jet, it is important to work with well modelled CSV distributions. Scale factors are derived by using a tag-and-probe method in a specific control region [102]. Different scale factors are derived for different jet flavours. Jets from up, down, and strange quarks as well as gluon jets are considered as light-flavour (LF) jets whereas bottom jets are considered as heavy-flavour (HF) jets. Jets from charm quarks are treated differently as discussed below. The control region consists of exactly two leptons and two jets. In addition, cuts on the dilepton mass and on the missing transverse energy are applied to suppress Z boson events. One of the two jets, the tag jet, has to pass the medium working point of the CSV algorithm. The remaining jet then is the probe jet. These selections are applied to data as well as simulation. The collection of probe jets from all the events passing the criteria form a high purity heavy flavour jet sample. Thus, it is possible to estimate the light flavour jet contamination in a heavy flavour jet enriched sample by using the generator information of the simulated datasets. This can also be done
5. Event Simulation

binned in jet $p_T$ and the CSV value of the jet. Thus, the scale factors can be calculated as

$$SF_{HF}(CSV, p_T) = \frac{\text{Data}(CSV, p_T) - MC_{LF}(CSV, p_T)}{MC_{HF}(CSV, p_T)}.$$  \hspace{1cm} (5.4)

For the LF scale factors, a light-flavour enriched sample has to be created. This can be done by inverting the Z boson cuts and requiring the tag jet to fail the loose working point of the CSV algorithm. The LF scale factors can then be extracted in an analogous way. However, the LF scale factors are additionally binned in $\eta$ resulting in

$$SF_{LF}(CSV, p_T, \eta) = \frac{\text{Data}(CSV, p_T, \eta) - MC_{HF}(CSV, p_T, \eta)}{MC_{LF}(CSV, p_T, \eta)}.$$  \hspace{1cm} (5.5)

Since the scale factors change the distribution of the CSV values on which the calculation of the scale factors depends, an iterative procedure is used. The scale factors are applied and recalculated several times until stable values are reached. After the calculation of the scale factors, a polynomial fit is performed on the LF scale factors to obtain continuous scale factors and reduce statistical uncertainties. For the HF scale factors, an interpolation between CSV bins is performed. Finally, the event is reweighted by the product of the scale factors for all jets contained in the considered event:

$$w_{CSV} = \prod_i SF_i.$$ \hspace{1cm} (5.6)

Since it is difficult to create a c-jets enriched sample, the scale factors for c-jets are set to unity but with a doubled uncertainty of the HF scale factors. Important systematic uncertainties influencing the CSV scale factors are the uncertainty on the jet energy scale, uncertainties on the sample purity, the uncertainty on the c-jets scale factors, and statistical fluctuations.

5.4.3. Pile-up Reweighting

The number of additional inelastic proton-proton collisions in one bunch crossing depends on the instantaneous luminosity. Since the instantaneous luminosity cannot be known at the time when the MC samples are generated, a rough estimate is used. Quantities like the muon isolation, used to set quality criteria on muon candidates, depend on the pile-up density. Therefore, it is useful to correct the pile-up scenario of the MC simulation to match the data. This is achieved with a reweighting procedure, adding an additional weight for every event. The event weight for an MC event with $n$ primary vertices can be calculated as the ratio of the expected number of events containing $n$ primary vertices calculated from data and the expected number of events with $n$ primary vertices determined by the MC production parameters. The number of primary vertices in data depends on the instantaneous luminosity and the minimum-bias cross section, for which a value of 69.4 mb was found to reproduce the observed primary vertex distribution well.
6. Event Selection

The events which are used in the \( t\bar{t}H(b\bar{b}) \) single-lepton analysis have to fulfill certain selection criteria. Those selection criteria are essentially based on the topology of the final state as explained in the previous chapter. The selection criteria are formulated using specifically defined physics objects, which will be explained in this chapter.

6.1. Physics Objects

After the physics objects are returned by the particle flow algorithm, additional criteria defined by the corresponding “CMS Physics Object Groups” are applied to the objects. A summary of reference selections and recommendations is provided by the “CMS Top Physics Analysis Group” [103].

6.1.1. Electrons

The criteria which PF electron candidates have to pass are defined by the “\( e-\gamma \) Physics Object Group” (POG) [104].

Electron Isolation:

To distinguish prompt electrons from secondary electrons and fake electrons, the relative electron isolation \( I_e \) is used. Prompt electrons are created in the hard interaction whereas secondary electrons are decay products from weakly decaying hadrons. Since hadrons are usually a part of jets, such secondary electrons are often close to a jet. This enables to define a measure of jet activity within a cone radius of \( \Delta R = 0.3 \) around the electron candidate. This measure is called relative isolation:

\[
I_e = \frac{\sum \text{charged hadrons } p_T + \max \left( 0, \sum \text{neutral hadrons } p_T + \sum \text{photons } - p_{A\text{eff}} \right)}{p_{T,e}} \tag{6.1}
\]

Basically, the transverse momentum of all hadrons and photons within the cone radius is summed and divided by the \( p_T \) of the electron candidate. However, the
charged hadrons coming from a secondary vertex are discarded by the charged hadron subtraction (CHS) algorithm. In addition, neutral hadrons and photons can contribute to the energy deposit in the cone due to pile-up collisions. To handle this, an average energy contribution from such pile-up collisions $\rho A_{\text{eff}}$ is subtracted where the average energy density of the event is $\rho$ and $A_{\text{eff}}$ is an effective $\eta$ dependent jet area defined in [105].

Multivariate Electron ID:

In addition to the relative isolation, the electron candidate is evaluated by a multivariate classifier and is consequently assigned an electron ID [104]. A boosted decision tree (BDT), see section 7.3, combines several observables from the calorimeters and the tracker into the multivariate electron ID. In this thesis, the MVA Electron ID is used to define PF electron candidates as electrons which can be used in the $t\bar{t}H(b\bar{b})$ analysis.

Conversion Rejection:

If photons which pass through detector layers are highly energetic, electron-positron pairs are created. The electrons and positrons can then be misinterpreted as prompt electrons. However, these conversion electrons show a track only after their production. This means that tracker layers close to the nominal interaction point are usually missed. Therefore, electron candidates are discarded if their track does not have hits in the first tracker layer. Moreover, since conversion electrons or positrons are produced in pairs, another track with inverse curvature is expected close to the original PF candidate. If this is the case, the candidate is discarded as well.

Loose and Tight Electrons:

All electron candidates need to have a pseudorapidity which does not lie between the barrel and endcap region. This results in an allowed region of $1.444 < |\eta| < 1.566$. The relative isolation needs to be smaller than a maximum value of 0.15. Usually, these criteria are implemented in the multivariate classifier assigning the electron ID. Because of this, it is possible to only work with cuts on $p_T$, $\eta$, and the value of the electron ID itself. If electrons have $p_T \geq 30$ GeV and $|\eta| < 2.1$, they are categorized as tight electrons. If the electrons only fulfill $p_T \geq 15$ GeV and $|\eta| < 2.4$, they belong to the loose electron collection.

6.1.2. Muons

The “CMS Muon POG” provides the recommendations for the muon selection [106].

Muon Isolation:
6.1. Physics Objects

Analogous to the relative electron isolation, an isolation variable is defined for the muons as well. The isolation cone size is chosen to be $\Delta R = 0.4$. This results in a difference concerning the subtraction of the neutral pile-up contribution described by the $\Delta \beta$ correction [107].

Loose and Tight Muons:

Only PF muon candidates having hits in the inner tracker and the muon system are considered further. The muon system has to recognize at least two hit segments. This leads to a suppression of hadrons which were able to traverse the calorimeters completely and reach the muon system. In addition, this requirement improves the $p_T$ measurement. Moreover, the muon isolation needs to be larger than a threshold of 0.15. The transverse and longitudinal impact parameters $d_{\text{trans}}$ and $d_{\text{long}}$ of the track with respect to the primary vertex have to fulfill $d_{\text{trans}} < 0.2 \text{ cm}$ and $d_{\text{long}} < 0.5 \text{ cm}$. Loose muons are then defined by $p_T \geq 15 \text{ GeV}$ and $|\eta| < 2.4$ whereas tight muons fulfill $p_T \geq 25 \text{ GeV}$ and $|\eta| < 2.1$.

6.1.3. Jets

In this thesis, only anti-$k_T$ jets with a cone radius of $\Delta R = 0.4$ are considered. These jets are obtained by clustering particle flow candidates with the anti-$k_T$ algorithm. The already mentioned charged hadron subtraction is applied before the start of the clustering to reduce the influence of pile-up. In addition, the clustering excludes PF muons and electrons. Other criteria are necessary to improve the purity of the jet collection because jets could also be clustered from misidentified particles or calorimeter noise. To avoid jet clustering of misidentified leptons, the jets are required to have more than one constituent from which at least one should be charged. Furthermore, the energy contributed by charged hadrons divided by the total jet energy (charged hadronic energy fraction) should be greater than zero. The neutral hadronic, charged electromagnetic and neutral electromagnetic energy fractions are required to be below 0.99. Finally, it is checked if the jet is within a $\Delta R = 0.4$ cone of any selected lepton. If this is the case, the jet is discarded.

After the creation of the jet collection, its constituents have to be corrected regarding their energies. Since the response of the calorimeter is non-linear as well as $\eta$ and $p_T$ dependent, this is a necessary procedure. The jet energy corrections essentially have three steps at CMS and also depend on the detector state at the time of data-taking. They are derived by the “Jet/$E_T$ POG” [108]. First, the jets are corrected with the L1 pile-up correction. This step removes energy from pile-up events and creates a luminosity independent sample. The next step is the L2L3 MC-truth correction. The two corrections relate the particle level $p_T$ to the reconstructed $p_T$. In addition, the jet response is made uniform in $\eta$ and $p_T$. The last step is the L2L3 Residual correction which is only applied to data to accommodate for small differences between data and MC simulation. After the jet energy corrections, cuts on the transverse momentum and the pseudorapidity can be applied. For this analysis,
jets are only selected if they have $p_T > 30$ GeV and $|\eta| < 2.4$. This suppresses influences from pile-up and background processes and ensures good reconstruction quality in the central detector region.

### 6.1.4. Missing Transverse Energy

The procedure to calculate the missing transverse energy was explained in section 3.3.7. The energy corrections on the jets are consequently propagated to $\vec{E}_T$ as well as the correction regarding the jet energy resolution, cf. section 5.4.1.

### 6.2. Event Selection

Several selection criteria are applied to suppress background processes contained in the data. The data which was used in this thesis corresponds to proton-proton collisions which were recorded with the CMS detector from April 22$^{\text{th}}$ 2016 to July 15$^{\text{th}}$ 2016. The amount of data corresponds to an integrated luminosity of $12.9 \, \text{fb}^{-1}$. The following selection is designed for semileptonic ttH(bb) events.

#### 6.2.1. Trigger Selection

All considered MC events have to be accepted by a high level trigger as it would be the case with real events. For electron events, the “HLT_Ele27_eta2p1_WPTight_Gsf” trigger path has to be passed. This trigger is passed by events in which the electron fulfills $p_T > 27$ GeV and $|\eta| < 2.1$. For muon events, two trigger paths, “HLT_IsoMu22” and “HLT_IsoTkMu22”, are available. These paths require muons with $p_T > 22$ GeV.

#### 6.2.2. Vertex Selection

Another requirement for an event is the reconstruction of a good quality primary vertex. Its position has to lie between $\pm 24$ cm in beam direction and $\pm 2$ cm in radial direction from the nominal beam spot. In addition, the vertex fit needs to have a number of degrees of freedom greater than four.
6.2. Event Selection

6.2.3. Lepton Selection

Events are only kept if they contain exactly one tight lepton and no additional loose leptons with the above definitions. Combined with the trigger selection, this requirement reduces QCD multijet, Drell-Yan and boson pair processes significantly. Furthermore, this selection separates the top-quark pair dileptonic final state from the semileptonic one.

6.2.4. Jet Selection

Each event needs to have at least four jets to pass this selection. This allows to keep $t\bar{t}H(bb)$ events in which jets are outside of the acceptance, jets were missed, or events with overlapping jets. In general, background processes like vector-boson pair production, vector bosons in association with jets or $t\bar{t}$ events with fewer than four jets can be suppressed with this requirement.

6.2.5. b-tag Selection

The last step is the requirement of at least two jets which are tagged as b-jets. The medium working point of the CSV algorithm is chosen. After this selection, the background consists mainly of $t\bar{t} +$ jets events with small contributions of single top production as well as $W+$jets events.
7. Statistical Model and Tools

The aim of the tH(bb) analysis is to establish the corresponding process in experimental data. Because of uncertainties in the experimental measurements as well as in the theoretical calculations, and the finite amount of data, it is only possible to make statistical statements. To discover a new process, a hypothesis testing procedure is used, which checks the compatibility of the data with the new process and the respective incompatibility with the known processes. The following explanations are based on [109].

7.1. Statistical Model

In general, a hypothesis test is done using a test statistic. A test statistic is a statistical quantity which can be calculated from random variables and is therefore a random variable itself. The distribution of the test statistic varies for different hypotheses. According to the Neyman-Pearson lemma [110], the most powerful test statistic is a likelihood ratio. In statistical data analysis, random variables are often examined with histograms. To apply the formalism of hypothesis tests to the search for new particles or processes, a null hypothesis $H_0$ and an alternative hypothesis $H_1$ have to be defined. The hypothesis which contains only known processes is used as the null hypothesis or background hypothesis. The hypothesis which includes the searched signal is the alternative hypothesis. Considering a random variable under hypothesis $H_1$, which is filled in a histogram $n = (n_1, \ldots, n_N)$, the expectation value for the $i$-th bin can be written as

$$E[n_i] = \mu s_i + b_i.$$  \hspace{1cm} (7.1)

In this equation, the signal strength parameter $\mu$ has been additionally introduced to vary the strength of the signal process. In the final step of the statistical procedure, a upper limit is set on this parameter. The quantities $s_i$ and $b_i$ are the expectation values of the examined random variable from signal and background process, respectively. The models for signal and background in general depend on nuisance parameters, here denoted by $\theta$. Thus, the probability distribution functions naturally depend on these nuisance parameters as well. These parameters influence the shape of the distributions and are considered as floating values within some uncertainty, which reflects the prior knowledge about the considered parameter, in the later fit. The expected number of entries in the $i$-th bin for signal and
background are then given as the integrals of their probability density functions:

\[
    s_i = s_{\text{tot}} \int_{\text{bin } i} f_s(x; \theta_s) \, dx \quad (7.2)
\]

\[
    b_i = b_{\text{tot}} \int_{\text{bin } i} f_b(x; \theta_b) \, dx \quad (7.3)
\]

Moreover, the total number of expected background events \( b_{\text{tot}} \) is often used as a nuisance parameter as well. To decrease the uncertainty of the nuisance parameters, supplementary measurements can be done in control regions, enriched with background events. The bin contents of some random variable filled in a histogram \( m = (m_1, \ldots, m_M) \) are consequently dependent on the nuisance parameters with their expectation values \( E[m_i] = u_i(\theta) \). The values \( u_i(\theta) \) are the expected number of events in the \( i \)-th bin, contributed by the background processes in the control region. Naturally, they depend on the nuisances as well. Since the entries in a histogram bin are Poissonian distributed around the mean value, the binned likelihood function becomes

\[
    L(\mu, \theta) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} e^{-u_k}. \quad (7.4)
\]

With this likelihood function, the profile likelihood ratio is calculated as

\[
    \lambda(\mu) = \frac{L(\mu, \hat{\theta})}{L(\hat{\mu}, \hat{\theta})}. \quad (7.5)
\]

Here, the numerator \( L(\mu, \hat{\theta}) \) is called the profile likelihood, which is obtained by maximizing the likelihood function under the assumption of a fixed value of \( \mu \). The parameter values at this maximum are \( \hat{\theta} \). The denominator \( L(\hat{\mu}, \hat{\theta}) \) is the global maximum with respect to all parameters, including \( \mu \). The parameter values at this global maximum are \( \hat{\mu}, \hat{\theta} \).

With this likelihood ratio, different test statistics can be defined, which are suited for different kind of problems. To quantify the discovery of a new signal, the test statistic

\[
    q_0 = \begin{cases} 
        -2 \ln(\lambda(0)) & \hat{\mu} \geq 0 \\
        0 & \hat{\mu} < 0 
    \end{cases} \quad (7.6)
\]

can be used, where \( \mu = 0 \) is set to test the background-only hypothesis. The special case for \( \hat{\mu} < 0 \) takes into account the fact that one assumes the contribution of the signal to be positive. Hence, only for an upward excess of the data, the test statistic points towards an incompatibility between data and background-only prediction. The p-value can then be defined as the probability of finding a value of the test statistic with greater disagreement than the one observed in data.

Another test statistic,

\[
    q_\mu = \begin{cases} 
        -2 \ln(\lambda(\mu)) & \hat{\mu} \leq \mu \\
        0 & \hat{\mu} > \mu 
    \end{cases}, \quad (7.7)
\]
can be used for setting upper limits on the signal strength parameter $\mu$. This test statistic is designed to interpret an upward fluctuation to not be incompatible with the data but only as an underestimation of the strength of the signal. Analogous, the p-value can then be calculated as

\[ p_\mu = \int_{q_{\mu,\text{obs}}}^{\infty} f(q_\mu | \mu) dq_\mu \]  

(7.8)

with the probability density function $f(q_\mu | \mu)$ of the test statistic $q_\mu$ under the assumption of a signal strength $\mu$ and the observed value of the test statistic $q_{\mu,\text{obs}}$ calculated with the measured data. To calculate upper exclusion limits on the signal strength parameter, the method of modified frequentist construction [111, 112] is applied. For this method, two probabilities, $\text{CL}_b$ and $\text{CL}_{s+b}$, are defined. They are defined the same way as p-values.

\[ \text{CL}_{s+b} = P(q_\mu \geq q_{\mu,\text{obs}} | \mu \cdot s + b) = p_\mu \]  

(7.9)

\[ \text{CL}_b = P(q_\mu \geq q_{\mu,\text{obs}} | b) = p_0 \]  

(7.10)

With these two probabilities,

\[ \text{CL}_s = \frac{\text{CL}_{s+b}}{\text{CL}_b} \]  

(7.11)

can be calculated. For exclusion of the nominal ($\mu = 1$) signal process at $1 - \alpha$ confidence level (CL), $\text{CL}_s$ has to be found smaller or equal than $\alpha$. To derive an upper limit on $\mu$ at $1 - \alpha$ confidence level, the value of $\mu$ has to be varied until $\text{CL}_s = \alpha$.

If $q_{\mu,\text{obs}}$ and the resulting p-values are calculated with measured data, the upper limit on $\mu$ is called observed limit. However, it is also possible to calculate $q_{\mu,\text{obs}}$ and the p-values with simulated data from the background-only hypothesis. From the MC simulation, several toy datasets are generated and with each background-only toy dataset, the upper limit at $1 - \alpha$ CL is calculated. This results in a distribution of upper limits. Following this, the expected upper limit can be obtained by calculating the median of this distribution. The expected upper limit at $1 - \alpha$ confidence level can be interpreted as a measure of the analysis sensitivity for the sought process. For this thesis, $\alpha = 0.05$ is chosen. In addition, the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties are calculated as the 16%, 84% and 2.5%, 97.5% quantiles of the distribution.

Since the repeated generation of toy datasets and calculation of limits is a computationally expensive task, another method is used in this thesis. This method uses the Asimov dataset. This dataset is defined for a given signal strength modifier $\mu$ with the assumption that the expected background yields and the nuisance parameters are at their nominal values. With this dataset, it is possible to describe the distribution of $f(q_\mu | \mu \cdot s + b)$ and $f(q_\mu | b)$ analytically in the asymptotic limit of large samples. Consequently, the analytical description can be used to find the value of $\mu$ for which $\text{CL}_s = 0.05$. The found value of $\mu$ is then quoted as the asymptotic expected upper limit at 95% confidence level. The $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties can be obtained with this method as well.
7.2. Systematic Uncertainties

As already mentioned, systematic uncertainties are considered by introduction of nuisance parameters $\theta$, which influence the number of events in the bins of the considered distributions. A short explanation of the detailed procedure proposed in [113] is given. An optimal estimator $\hat{\theta}$ for a nuisance parameter can be assigned conceptually by an imaginary or real measurement. Different a-posteriori probability distribution functions (PDFs) $\rho(\theta|\hat{\theta})$ can be used for the nuisance parameters. They represent the probability or degree of belief in the value of $\theta$ given the measurement $\hat{\theta}$. If there is no knowledge at all for a nuisance parameter it is assigned a flat hyper-prior $\pi(\theta)$. Consequently, a-priori PDFs $p(\hat{\theta}|\theta)$ can be obtained for the measurements by using Bayes’ theorem.

$$
\rho(\theta|\hat{\theta}) \sim p(\hat{\theta}|\theta) \cdot \pi(\theta)
$$

(7.12)

The a-priori PDF (or likelihood) represents the probability of measuring the value $\hat{\theta}$ under the assumption of $\theta$ being the true value of the nuisance. In this thesis, two kinds of systematic uncertainties, called rate and shape uncertainty, are used. A rate uncertainty influences the overall number of events due to a specific process. An example of a rate uncertainty would be the theoretical uncertainty on a cross section or the uncertainty on the measured integrated luminosity. Such a theoretical cross section uncertainty is, amongst others, applied on the heavy-flavour background processes in the $t\bar{t}H(b\bar{b})$ analysis.

The assigned posterior for rate uncertainties usually is the log-normal PDF

$$
\rho(\theta|\hat{\theta}) = \frac{1}{\sqrt{2\pi} \ln \kappa} \exp\left(\frac{-1}{2(\ln \kappa)^2} \ln\left(\frac{\theta}{\hat{\theta}}\right)^2\right) \frac{1}{\theta}.
$$

(7.13)

Here, $\kappa$ is called the width of the distribution and $\hat{\theta}$ is the best estimate. A shape uncertainty influences the shape of a probability distribution or its respective histogram and cannot be taken into account with just a number varying the normalization. Instead, new templates/histograms have to be created for which the nuisance parameter is varied up and down one standard deviation. The assigned probability density function for shape nuisance parameters is a Gaussian distribution with unit width and centered around zero. For shifts other than one standard deviation, the templates are interpolated vertically in a linear (shifts greater than one standard deviation) or quadratic (smaller than one standard deviation) way.

The systematic uncertainties taken into account in this thesis are listed in Tab. 7.1 and were adapted from [13]. The systematic uncertainties can have an impact on the yield of the signal and background processes, the shape of the final discriminants, or on both at the same time. For the luminosity, an uncertainty of 6.2% is used [114]. The uncertainties on the electron and muon identification together with the trigger efficiency were obtained by comparing the measured efficiency between data and MC simulation in a high-purity $Z$ boson decay sample. The uncertainty was found to be around 2-4%. The uncertainty on the number of pile-up events was estimated by varying the cross section for pile-up events in MC by $\pm 5\%$. Furthermore, the uncertainty on the jet energy scale [115] was...
Table 7.1: A list of the systematic uncertainties taken into account in this thesis. Their source is given and the type defines if the uncertainty is treated as a rate or shape uncertainty.

<table>
<thead>
<tr>
<th>Source</th>
<th>Type</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminosity</td>
<td>rate</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>Lepton ID/trigger eff.</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>Pile-up</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b-tag HF fraction</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b-tag HF stats (linear)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b-tag HF stats (quadratic)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b-tag LF fraction</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b-tag LF stats (linear)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b-tag LF stats (quadratic)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b-tag charm (linear)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>b-tag charm (quadratic)</td>
<td>shape</td>
<td>Signal and all backgrounds</td>
</tr>
<tr>
<td>QCD scale (tH)</td>
<td>rate</td>
<td>Scale uncertainty of NLO tH prediction</td>
</tr>
<tr>
<td>QCD scale (t)</td>
<td>rate</td>
<td>Scale uncertainty of NLO t prediction</td>
</tr>
<tr>
<td>QCD scale (t + hf)</td>
<td>rate</td>
<td>Add. scale uncertainty of NLO t + hf prediction</td>
</tr>
<tr>
<td>QCD scale (t)</td>
<td>rate</td>
<td>Scale uncertainty of NLO single t prediction</td>
</tr>
<tr>
<td>QCD scale (V)</td>
<td>rate</td>
<td>Scale uncertainty of NNLO W and Z prediction</td>
</tr>
<tr>
<td>QCD scale (VV)</td>
<td>rate</td>
<td>Scale uncertainty of NLO diboson prediction</td>
</tr>
<tr>
<td>pdf (gg)</td>
<td>rate</td>
<td>PDF uncertainty for gg initiated processes except tH</td>
</tr>
<tr>
<td>pdf (gg tH)</td>
<td>rate</td>
<td>PDF uncertainty for tH</td>
</tr>
<tr>
<td>pdf (qg)</td>
<td>rate</td>
<td>PDF uncertainty for qg initiated processes (W, Z)</td>
</tr>
<tr>
<td>pdf (qg)</td>
<td>rate</td>
<td>PDF uncertainty for qg initiated processes except (singlet)</td>
</tr>
<tr>
<td>$Q^2$ scale (t)</td>
<td>shape</td>
<td>Renormalization and factorization scale uncertainties of the tf ME generator, independent for additional jet flavours</td>
</tr>
</tbody>
</table>
calculated by scaling the energy scale correction of all jets in the signal and background predictions by one standard deviation. The uncertainties on the scale factors of the CSV b-tagging algorithm have to be calculated specifically for the different jet flavours, which are heavy (HF), light (LF), and charm flavour. In the control sample used to evaluate the b-tagging scale factors, the background contamination and the jet energy scale uncertainty as well as the statistical uncertainties of the scale factor evaluation are shifted by one standard deviation. The statistical uncertainty can be parametrized as a sum of a linear and a quadratic term. The QCD scale and pdf uncertainties account for the theoretical uncertainties on the cross sections of the different processes. Those uncertainties are due to uncertainties on the renormalization and factorization scale as well as the parton-density-functions.

### 7.3. Boosted Decision Trees

In high energy physics, so-called multivariate analysis (MVA) methods are often used to enhance the sensitivity of difficult analyses by constructing variables which are sensitive to the given problem due to optimisation/training. In this $t\bar{t}H(b\bar{b})$ analysis, boosted decision trees (BDTs) are used to construct discriminators which can later be used in the fit to obtain the upper limit on the signal strength modifier as explained above. Other analyses use, for example, artificial neural networks or so-called deep-neural networks which have risen to popularity in the last years. Every classifier has its own favourable or difficult properties which have to be considered in the choice for the classifier to be used. A standard textbook on machine learning [116] and the manual of the “Toolkit for Multivariate Data Analysis with ROOT” (TMVA) [117] was used to create this section. The latter is the software framework in which the BDTs were trained and tested for the $t\bar{t}H(b\bar{b})$ analysis.

For each event considered in a multivariate search for a new process in data, the question whether the event is signal-like or background-like has to be answered. In the simplest way, this is a binary question and the corresponding event is assigned with the corresponding class after examination of its properties. Such a classification procedure can be built with a decision tree. The examination of the properties is structured in a sequential way in form of a tree. For every step in the tree, one variable of the observed phase space is considered and a binary decision (for continuous input variables, this is usually a cut) is made upon this variable. The events corresponding to the two possible outcomes, then move on to the next layer of the trees. The number of layers is called the depth of the tree and the subsets of events fulfilling the different cuts within a layer are visualized as nodes in the tree, see Fig. 7.1. Graphically, this can be imagined as slicing the phase space into rectangular pieces, cf. Fig. 7.1 for a multi-classification example.

Before a tree can be used, it has to be built and optimized/trained for the problem at hand. In this analysis, two classes, signal and background, have to be distinguished. For the training of the tree, the number of its layers and the precision of the cuts on the input
variables can be varied. The training is usually done until a predefined stopping-criterion is reached. The nodes for which the stopping-criterion is reached are called leaf nodes. These leaf nodes, or analogous the assigned slices of phase space, are used to determine the classes of the considered events by evaluating which class made up the majority in the considered leaf node or had the highest purity during the training. Only events whose classes are known can be used for training. Therefore, the training in the $t\bar{t}H(b\bar{b})$ analysis is performed with MC events. As already mentioned, the tree can be optimized to get a better classification. The separation can be improved by minimizing the number of wrong classifications in the leaf nodes or by increasing the purity of the subsamples in the nodes. In the following, a few examples of a measure of separation are given, where $p$ denotes the purity $\left(\frac{N_S}{N_S+N_B}\right)$ of the node and $N_S,N_B$ is the number of signal and background events, respectively.

- Gini-Index: $G = p \cdot (1 - p)$
- Misclassification error: $\beta = 1 - \max(p, 1 - p)$
- Statistical significance: $\alpha = \frac{N_S}{\sqrt{N_S+N_B}}$

The Gini-Index is a measure of impurity in a node. If the node contains a large fraction of signal events compared to background events or reversed, the impurity is low because most of the events in the node belong to the same class. However, if the numbers are comparable and the node is in a mixed state, the impurity is high. The misclassification error essentially gives the fraction of events which are misclassified in the node. The statistical significance compares the number of signal events in the node with the Poissonian uncertainty of the total number of events in the node. Thus, to find the best combination of variable and cut, the decrease in impurity (imp) from the parent node to the two daughter nodes is
calculated by

\[ \text{imp(parent node)} - (w_1 \cdot \text{imp(daughter node 1)} + w_2 \cdot \text{imp(daughter node 2)}) \]  

(7.14)

and subsequently maximized. Here, \( w_1 \) and \( w_2 \) are the fraction of events in daughter node 1 and 2 compared to the number of events in the parent node.

A single tree has the problem that it cannot handle statistical fluctuations, due to the limited MC statistics, very well. A classifier which was trained only with a small number of training events compared to its complexity, will in general learn statistical fluctuations of the training dataset instead of the features of the underlying distributions. A procedure named boosting, with the aim of getting a more stable classifier, can be applied to decision trees. Its principle is to build several classifiers considering and combining the response of each one in the final classification. The single decision trees are in general weak learners (especially with low depth). Weak learners are classifiers which are constructed in a simple way and consider only few properties/variables of the events. A weak learner, in general, does not have a high classification power but can be evaluated quickly. However, an ensemble of weak learners can be combined to a stronger learner. The low depth of the single trees has the advantage of a higher robustness against statistical fluctuations of the training dataset. There are several boosting algorithms. The training of the BDTs in this analysis uses gradient boosting. The idea behind the boosting is to minimise the difference between the classification output of the trained model and the true class. During the procedure, the binary outputs of the decision trees are combined into a continuous output. The algorithm works iteratively, meaning that it adds a new estimator (another decision tree) to gain a better model in every step. This is essentially done by using a gradient descent algorithm on a defined loss function, which shows the discrepancy between prediction and true classification.

In addition, it is important to train and test the BDTs on statistically independent samples to check its behaviour regarding overtraining. Overtraining describes the already mentioned effect that a classifier learned the specific properties of the training dataset. This effect can be checked with a testing dataset. In general, the set of all available training events is split into a statistically independent training and testing dataset. An overtrained classifier should consequently perform worse on the testing dataset than on the training dataset. Furthermore, the performance of a classifier can be visualized in a simple way by using the receiver operating characteristic (ROC). A cut on the output of the BDT is made and based on this cut, the events are classified as signal or background. Then, with the probability distributions of the BDT outputs for signal and background training samples, it is possible to calculate the signal efficiency and the background rejection for a given cut. By varying the cuts, the ROC curve can be drawn. By integrating the curve to obtain the ROC integral, a figure of merit for the separation provided by the variable is given. However, it is not possible to completely predict the influence of an input variable on the statistical limit of the analysis based solely on the ROC integral.
8. Theory of Spin Correlations

As was explained in section 4.1.1, the top quark is the most massive elementary fermion with a mass of approximately 173 GeV, which leads to a decay width $\Gamma_t \sim G_F m_t^3 \approx 1$ GeV. Its decay width is therefore significantly larger than the typical scale of hadronisation ($\Lambda_{\text{QCD}} \sim 0.1$ GeV) or the associated decorrelation scale at which the strong interaction affects the spin ($\Lambda_{\text{QCD}}^2/m_t \sim 0.1$ MeV) [119]. This has the consequence that the spin of the top quark is distributed to its decay products before the influence of strong interactions can have an effect on its spin. In addition, the angular distributions of the decay products of the top quark are correlated with the top quark’s spin axis. If two top quarks are considered, for example in $t\bar{t}$-production, the combination of the two aforementioned effects lead to a correlation between the decay products of the two top quarks. These correlations can consequently be measured in the angular distributions of the final-state particles originating from the top quarks.

8.1. Introduction

At the LHC, top-quark pair production is dominated by gluon-gluon fusion. In the low-$x$ regime, the PDF of the proton is made up almost entirely of gluons, see Fig. 2.2. Although the momentum and therefore energy fractions of these gluons are small compared to the protons, the energy of the two protons is still high enough to produce the masses of the two top quarks out of the two partons. As a result, the major part of the cross section originates from $gg \rightarrow t\bar{t}$ and an additional small part from quark-antiquark annihilation $q\bar{q} \rightarrow t\bar{t}$. For two special cases, defined as $m_{t\bar{t}} \rightarrow 2m_t$ and $m_{t\bar{t}} \rightarrow \infty$, there are simple arguments how the spins or helicities should behave in $t\bar{t}$-production, see [120]. If the invariant mass of the $t\bar{t}$-pair is very low, such that the energy was just enough to create the masses of the two top quarks (also called threshold-production), the angular momentum of the system vanishes and the total angular momentum only consists of spin. Since the initial gluons are physical states, they have (as massless particles) only transverse polarization states. Because of this, the resulting spin state of the $t\bar{t}$-system has to be

$$\frac{1}{\sqrt{2}} (|+,-\rangle - |-, +\rangle)$$

with a total spin of zero. The spin axis is chosen parallel to the momentum axis (and in the same direction for both quarks) in the $t\bar{t}$ center-of-mass system leading to opposite spin or
8. Theory of Spin Correlations

Figure 8.1.: The differential cross section for the different $t\bar{t}$-production channels (dashed and dotted lines) and the total (solid line) dependent on the invariant mass of the top-pair [122]. As was mentioned in the text, for low/high invariant mass pairs of the top-quark pair the like/unlike-helicity gluons dominate the cross section.

Same helicity states for the top quark and the top antiquark. Complementary to this, if the invariant mass of the top-quark pair is very high, the masses of the top quarks can be neglected and the helicity is conserved leading to same spin or opposite helicity states. To quantify the correlations in the whole process of $t\bar{t}$-production, the spin-correlation coefficient

$$C_{t\bar{t}} \equiv \frac{\sigma_{\uparrow\uparrow} + \sigma_{\downarrow\downarrow} - \sigma_{\uparrow\downarrow} - \sigma_{\downarrow\uparrow}}{\sigma_{\uparrow\uparrow} + \sigma_{\downarrow\downarrow} + \sigma_{\uparrow\downarrow} + \sigma_{\downarrow\uparrow}}$$

is often used [120, 121], where, for example, $\sigma_{\uparrow\downarrow}$ denotes the cross section for $t\bar{t}$-production with the top-quark spin having the same direction as the top-quark spin axis and the top-antiquark spin being oppositely aligned to the top-antiquark spin axis. For the two simplified cases explained above, this means a 100% correlation ($C_{t\bar{t}} = 1$) for threshold-production and 100% anticorrelation ($C_{t\bar{t}} = -1$) for the ultrarelativistic limit. In general, the value of the correlation depends on the invariant mass of the top-quark pair and on the choice of the spin axes, which were chosen parallel to the movement direction in this example. Other choices will change the correlation coefficient.

8.2. Spin Correlations in Top-Quark Pair Production

This part follows [121] for a short summary of the different processes that take part in $t\bar{t}$-production and how they influence the spin properties of the top-quark pair without restricting to the two limiting cases described in the introduction. The differential cross section for different $t\bar{t}$-production processes, separated into their helicity components and dependent on the invariant mass of the top-quark pair, can be found in Fig. 8.1. The following summary shows the impact of the spin-axis choice. It is common to define the
8.2. Spin Correlations in Top-Quark Pair Production

Figure 8.2.: The direction of the spin unit-vector \( \hat{s} \) in the top-quark rest frame [121]. The spin vector is defined with an angle \( \xi \) clockwise to the flight direction of the top antiquark. In the box in the right bottom corner, the situation is shown in the center-of-mass system of the top-quark pair, where the movement axis of the top quark is at an angle \( \theta \) with respect to the gluon axis.

spin vector \( \hat{s} \) (direction of the spin axis) in the rest frame of the considered particle. The spin vector of the top quark is then defined in such a way that it is at an angle \( \xi \) clockwise with respect to the momentum of the top antiquark, see Fig. 8.2. The same is done for the spin of the top antiquark in its rest frame and the corresponding angle \( \xi' \). If \( \xi = \xi' \), the spin vectors are antiparallel in the center-of-mass frame of the two top quarks. In the following sections, \( \beta \) denotes the relativistic velocity.

8.2.1. Spin Structure of Quark-Antiquark Annihilation

By using the same angles \( \xi = \xi' \) in combination with \( \theta \) being the scattering angle in the center-of-mass system (see Fig. 8.2) as well as using the relation

\[
\tan \xi = \gamma^{-1} \tan \theta \quad \text{with} \quad \gamma \equiv \frac{1}{\sqrt{1 - \beta^2}}
\]

(8.3)

to set the basis, the up-up and down-down spin components vanish. Only the mixed states up-down and down-up remain. This basis is called the off-diagonal basis and it incorporates the beamline basis (\( \cos \xi = \cos \theta \)) for small invariant masses of the top-quark pair (low \( \beta \)) and turns into the helicity basis (\( \cos \xi = -1 \)) for very high invariant masses (high \( \beta \)), where the spin direction of the top quarks is then measured in the same direction as their momentum vector. This is formalized by the following equations:

\[
|M(q_R \bar{q}_L \rightarrow t_\uparrow \bar{t}_\uparrow \text{ and } t_\downarrow \bar{t}_\downarrow)|^2 = |M(q_L \bar{q}_R \rightarrow t_\uparrow \bar{t}_\uparrow \text{ and } t_\downarrow \bar{t}_\downarrow)|^2 = 0
\]

(8.4)

\[
|M(q_R \bar{q}_L \rightarrow t_\uparrow \bar{t}_\downarrow \text{ or } t_\downarrow \bar{t}_\uparrow)|^2 = |M(q_L \bar{q}_R \rightarrow t_\uparrow \bar{t}_\downarrow \text{ or } t_\downarrow \bar{t}_\uparrow)|^2 \sim \left(1 \mp \sqrt{1 - \beta^2 \sin^2 \theta}\right)
\]

(8.5)

The corresponding behaviour is illustrated in Fig. 8.3, where \( \Omega \equiv \xi \). The spin axes have the angle \( \Omega \) with respect to the red arrows representing the momentum vectors of the top quarks. The green arrows actually represent the spin states of the top quarks, not their spin axes. Thus, the final state is \( t_\uparrow \bar{t}_\downarrow \) on the left side and \( t_\downarrow \bar{t}_\uparrow \) on the right side.
8. Theory of Spin Correlations

Figure 8.3.: Illustration of the off-diagonal basis [123]. The spin axes of the top quarks have an angle $\Omega$ with respect to the top momenta in the center-of-mass system of the top-quark pair. The green arrows represent the spin state of the respective quark, so on the left the spins are up-down and on the right they are down-up.

8.2.2. Spin Structure of Unlike-helicity Gluon-Gluon Fusion

The spin structure for the unlike-helicity gluon-gluon fusion is analogous to the quark-antiquark annihilation. Therefore, the off-diagonal basis can again be chosen so that

\[ |M(g^R_L \rightarrow t^R \bar{t}^L \& t^L \bar{t}^R)|^2 = |M(g^R_R \rightarrow t^R \bar{t}^L \& t^L \bar{t}^R)|^2 = 0 \]  
\[ |M(g^R_L \rightarrow t^R \bar{t}^L \& t^R \bar{t}^L)|^2 = |M(g^R_R \rightarrow t^R \bar{t}^L \& t^L \bar{t}^R)|^2 \sim \left(1 \pm \sqrt{1 - \beta^2 \sin^2 \theta^2}\right)^2. \]  

This situation is analogous to the one presented in Fig. 8.3.

8.2.3. Spin Structure of Like-helicity Gluon-Gluon Fusion

For this process, the off-diagonal basis is not the optimal choice, but the helicity basis is more appropriate. In this basis, the matrix elements are:

\[ |M(g^L_L \rightarrow t^L \bar{t}^L \& t^R \bar{t}^R)|^2 = |M(g^R_R \rightarrow t^R \bar{t}^L \& t^L \bar{t}^L)|^2 = 0 \]  
\[ |M(g^L_L \rightarrow t^L \bar{t}^L \& t^L \bar{t}^R)|^2 = |M(g^R_R \rightarrow t^R \bar{t}^R \& t^L \bar{t}^L)|^2 \sim \gamma^{-2}(1 \pm \beta)^2. \]  

These equations hold for all $\beta$ within this like-helicity gluon-gluon fusion process contrary to the usual conjecture that helicity is only appropriate for highly relativistic scenarios. The situation is visualized in Fig. 8.4.
8.3. Top-Quark Decay

A top quark decays weakly by coupling to a W boson and a bottom quark. The W boson subsequently decays leptonically or hadronically. The relation between the angle of the decay products of polarized top quarks and the chosen spin axis is given by

\[
\frac{1}{\Gamma_T} \frac{d\Gamma}{d\cos \chi_i} = \frac{1 + \alpha_i \cos \chi_i}{2},
\]

where \( \Gamma_T \) denotes the total decay width of the top quark. The coefficient \( \alpha_i \), often called spin analyzing power, can take values between \(-1\) and \(1\) and differs for different particles, see Tab. 8.1. The angle \( \chi_i \) is measured between the direction of fermion \( i \) and the chosen spin axis of the top quark in its rest frame [121]. The charged leptons and the down-type quarks have the largest correlation with the axis of the top-quark spin because they have the largest spin analyzing power coefficient. It is also possible to use the quark coming of the hadronic W-decay with the lowest energy in the top-quark rest frame. This is a pragmatic approach because experimentally it is hardly possible to distinguish between

Table 8.1.: Spin analyzing power \( \alpha \) of several decay products in the top-quark decay for different orders of perturbation theory [119, 124]. The spin analyzing power of the corresponding antiparticles can be obtained by reversing the sign of the values.

<table>
<thead>
<tr>
<th></th>
<th>( \ell^+ )</th>
<th>( \bar{d} ) quark</th>
<th>( \bar{\nu}_l )</th>
<th>u quark</th>
<th>b quark</th>
</tr>
</thead>
<tbody>
<tr>
<td>LO</td>
<td>1</td>
<td>1</td>
<td>-0.31</td>
<td>-0.31</td>
<td>-0.41</td>
</tr>
<tr>
<td>NLO</td>
<td>0.998</td>
<td>0.966</td>
<td>-0.314</td>
<td>-0.317</td>
<td>-0.393</td>
</tr>
</tbody>
</table>
8. Theory of Spin Correlations

Table 8.2.: Numerical values for the spin-correlation operators taken from [126] with $\mu$ being the factorization and renormalization scale. The expectation values of the operators also depend on the center-of-mass energy of the collider.

<table>
<thead>
<tr>
<th>Operator</th>
<th>$m_t/2$</th>
<th>$m_t$</th>
<th>$2m_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$O_{\text{spin}}$</td>
<td>-0.239</td>
<td>-0.236</td>
<td>-0.234</td>
</tr>
<tr>
<td>$O_{\text{hel}}$</td>
<td>0.327</td>
<td>0.328</td>
<td>0.331</td>
</tr>
</tbody>
</table>

jets initiated by up- and down-type quarks. The spin analyzing power for this quark would be 0.51 (0.47) at LO (NLO) [125] combining the great spin analyzing power of the $\bar{d}$-quark and the fact that it is the least energetic one more than half of all times [120].

8.4. Spin-Induced Angular Distributions

As was shown in the previous sections, QCD processes correlate the spins of the two top quarks depending on the production mechanism. However, the spin of one top quark cannot be determined. Only the polarization of an ensemble of top quarks can be measured within angular distributions. The polarization of the top quarks in strong production is very small (even zero on tree level) but the relative polarization, or in other words the correlation, is clearly visible. In this section, a short summary of the operators and the interesting angular distributions is given following [126].

There are essentially two operators of concern, which are

$$O_{ab} = \left\langle 4(\vec{s}_t \cdot \vec{a})(\vec{s}_{\bar{t}} \cdot \vec{b}) \right\rangle \quad \text{and} \quad O_{\text{spin}} = \left\langle \frac{4}{3} \vec{s}_t \cdot \vec{s}_{\bar{t}} \right\rangle .$$  \hspace{1cm} (8.11)

The vectors $\vec{a}$ and $\vec{b}$ are reference directions and therefore unit-vectors. The operator $O_{ab}$ describes the correlation of the spins of the top quarks with reference directions used as the spin axes. The various choices for the spin axes of the previous chapters then read:

$$\vec{a} = -\vec{b} = \vec{k} \rightarrow O_{\text{hel}}$$  \hspace{1cm} (8.12)

$$\vec{a} = \vec{b} = \vec{p} \rightarrow O_{\text{beam}}$$  \hspace{1cm} (8.13)

$$\vec{a} = \vec{b} = \vec{d} \rightarrow O_{\text{off-diag}},$$  \hspace{1cm} (8.14)

with $\vec{k}$ being the flight direction of the top quark in the center-of-mass frame of the top-quark pair, $\vec{p}$ the beam line in the same frame and $\vec{d}$ corresponds to the off-diagonal basis explained above. The operator $O_{\text{spin}}$ is defined independent of the choice of the spin axes and describes the correlations between the spins of the top quarks themselves. For numerical values of the operators, see Tab. 8.2.
Angular distributions which incorporate the spin-correlation effects will be presented in the following. The given functional forms are only valid if no kinematic cuts are applied.

First, the double-differential distribution

\[ \frac{1}{\sigma} \frac{d\sigma}{d \cos \theta_1 d \cos \theta_2} = \frac{1}{4} (1 + B_1 \cos \theta_1 + B_2 \cos \theta_2 - C \cos \theta_1 \cos \theta_2) \quad \text{with} \quad C = \alpha_1 \alpha_2 O_{ab}, \]

(8.15)

where \( \theta_1 (\theta_2) \) is the angle between the flight direction of lepton \( l^+ (l^-) \) or jet \( j_1 (j_2) \) from \( t (\bar{t}) \) decay (in its rest frame) and the direction \( \mathbf{a} (\mathbf{b}) \). The procedure to obtain the angles is illustrated for the helicity basis in Fig. 8.5 for the two leptons. The coefficients \( B_1, B_2 \) are due to polarization by weak and higher-order QCD interactions of the top quarks, but they are very small \([128, 129]\) and usually can be neglected. The coefficient \( C \), which is proportional to the spin operator \( O_{ab} \), describes the impact of the spin correlations on this double-differential distribution. In addition, the observed decay products have an influence on the magnitude of the coefficient \( C \) in form of their spin analyzing power \( \alpha_1 \) and \( \alpha_2 \). Without spin correlations and polarization, this is a two-dimensional flat distribution whereas in case correlations are present, the projections on one axis become linear distributions. This can be seen in Fig. 8.6.
Moreover, the two-dimensional distribution from Eq. 8.15 can be integrated to obtain a one-dimensional distribution in \( \cos \theta_1 \cdot \cos \theta_2 \) [131]. This leads to the distribution
\[
\frac{1}{\sigma} \frac{d\sigma}{d \cos \theta_1 \cos \theta_2} = \frac{1}{2} (1 - C \cos \theta_1 \cos \theta_2) \ln \frac{1}{|\cos \theta_1 \cos \theta_2|}.
\] (8.16)

It is also possible to look at the related one-dimensional distributions of the helicity angles to analyze the already mentioned polarization and the corresponding coefficients \( B_1 \) and \( B_2 \) from Eq. 8.15. These distributions have the form
\[
\frac{1}{\sigma} \frac{d\sigma}{d \cos \theta_{1,2}} = \frac{1}{2} (1 + B_{1,2} \cos \theta_{1,2}),
\] (8.17)

where the angles are defined in the same way as above.

The one-dimensional distribution
\[
\frac{1}{\sigma} \frac{d\sigma}{d \cos \theta} = \frac{1}{2} (1 - D \cos \theta) \text{ with } D = \alpha_1 \alpha_2 O_{\text{spin}}
\] (8.18)
is of importance as well. In the latter equation, \( \theta \) denotes the angle between the momentum direction of the lepton \( l^+ \) or jet \( j_1 \) and \( l^- \) or \( j_2 \). The flight directions of the decay particles are hereby defined in the rest frame of their parent top quark. The coefficient \( D \) is directly proportional to the spin correlation operator \( O_{\text{spin}} \). Because of this, the operator influences the slope of the corresponding distribution directly. The spin analyzing power of the observed decay products influence the magnitude of \( D \) as well and additionally its sign. This distribution and the influence of spin correlations are visualized in Fig. 8.7 together with the integrated two-dimensional distribution, which is dependent on the product of the two cosines. The left side shows a small asymmetry in the integrated distribution. In case of no spin correlations, it would be completely symmetric instead. The right side shows even more clearly the transformation from a flat to a linear distribution caused by spin correlations.

The final distribution is
\[
\frac{1}{\sigma} \frac{d\sigma}{d \Delta \Phi},
\] (8.19)
8.5. Spin Flip by Gluon Radiation

Figure 8.7.: The distribution of $\cos \theta_1 \cdot \cos \theta_2$ of the two leptons from Eq. 8.16 (left) and the three-dimensional opening angle between the two leptons from Eq. 8.18 (right) are shown in the helicity frame with activated (solid) and neglected (dashed) spin correlations [126, 132]. The left plot only shows the case with spin correlations.

where $\Delta \Phi$ is the azimuthal difference of one top and one antitop decay product. This distribution is usually measured in the laboratory frame. Unfortunately, a functional form could not be found.

As already mentioned, the analytic form of these distributions and their coefficients are only valid in the absence of cuts. The shapes of these distributions will be deformed in a real analysis with kinematic cuts. Still, it is proposed that fits can be used in trying to determine the correlation coefficients [126]. In addition, several estimators which can be used for the determination of the coefficients, are given.

$$\hat{C} = -9 \langle \cos \theta_1 \cos \theta_2 \rangle$$  
$$\hat{D} = -3 \langle \cos \theta \rangle$$  
$$\hat{B}_1 = 3 \langle \cos \theta_1 \rangle$$  
$$\hat{B}_2 = 3 \langle \cos \theta_2 \rangle$$

If kinematic cuts are not applied $\hat{C} = C$, $\hat{D} = D$, $\hat{B}_1 = B_1$, and $\hat{B}_2 = B_2$.

8.5. Spin Flip by Gluon Radiation

The behaviour of the spin correlations can be changed if the spin of a particle is flipped. This is theoretically possible by the radiation of a gluon. In general, the radiation of a gluon can flip the spin of an off-shell top quark via its chromomagnetic moment, but this
is suppressed because of the high top-quark mass. The transition rate for such a process, which is the QCD analogue of a M1 transition, is given by [125]

$$\Gamma(M1) \sim \alpha_s E_g^3 m_t^2$$  \hspace{1cm} (8.24)

with $\alpha_s(m_t) = 0.1$ and $E_g, m_t$ being the energy of the radiated gluon and on-shell top-quark mass, respectively. For the average of gluon energies, the timescale for this process to happen is much longer than the lifetime of the top quark. Therefore, the influence of such processes is expected to be rather small.

### 8.6. Spin Behaviour in Associated Higgs Boson Production with a Top-Quark Pair

The influence of the Higgs boson on the spin correlations of the top quarks is explained in this section following [12]. The main difference between $t\bar{t}$+jets and $t\bar{t}H$ is the radiation of the additional Higgs boson within the top-Higgs coupling. As explained in the section about the electroweak theory and the Higgs mechanism, this interaction is introduced as a Yukawa coupling and has, after spontaneous symmetry breaking, the form

$$L_{\text{Higgs interaction}} = -\frac{m_f}{v} \left( \bar{f}_L f_R + \bar{f}_R f_L \right) H \hspace{1cm} (8.25)$$

for a fermion $f$ with mass $m_f$, the vacuum expectation value for the Higgs field $v$ and the Higgs field $H$. This Lagrangian directly shows that the radiation of the Higgs boson leads to a possible chirality flip of the top quark by coupling left- and right-chiral states. Because the chirality is related to the helicity of a particle, a qualitative difference in the spin correlations of the final-state top quarks is intuitively expected. For the ultrarelativistic case $m_{t\bar{t}} \to \infty$ that was presented in the introduction to this chapter, this also implies a flip of the helicity states because for highly relativistic particles, chirality and helicity become the same quantity. As a consequence, the helicities in this case are left-right and right-left, meaning 100% anticorrelated spin states (spin axes defined in movement direction), opposite to the $t\bar{t}$-only case. Unfortunately, the limit case explained above is more for illustration purposes because it is highly unlikely to produce these heavy particles in an ultrarelativistic state with the energies the LHC can provide. As a consequence of the significant top-quark mass, a chirality flip does not have to imply a spin flip, hence the spin correlations cannot be deduced in a simple illustrative way but have to be calculated. In Fig. 8.8, the integrated transverse momentum distribution of the hardest top quark separated into like- and unlike-helicity configurations is shown for $t\bar{t}$ on the left and for $t\bar{t}H$ on the right side. The distributions show that the behaviour of the helicity states varies significantly over the production phase space and exhibits a qualitative difference in the composition of the configurations. For $t\bar{t}$, the two simplified cases for low (mainly same helicity states) and high (opposite helicity states) invariant masses of the top-quark
8.6. Spin Behaviour in Associated Higgs Boson Production with a Top-Quark Pair

Figure 8.8.: Integrated top-quark $p_T$ distribution for the like- and unlike-helicity states of top-quark pairs in $t\bar{t}$ and $t\bar{t}H$ at a center-of-mass energy of 14 TeV in the laboratory frame. On the $x$-axis is the minimum $p_T$ of the hardest top quark.

Pairs can be found in the left distribution. However, the opposite helicity configurations which would be expected for high invariant masses in $t\bar{t}H$, due to the spin flip, cannot be observed because of the massive Higgs boson which takes a significant amount of energy to create. This means that the ultrarelativistic or chiral limit for both top quarks in the final state is expected to be reached for higher leading top-quark $p_T$. This is checked in a simple way for the dileptonic $t\bar{t} + \text{jets}$ and $t\bar{t}H(b\bar{b})$ sample. Only events with a leading top-quark transverse momentum greater than 200 GeV were selected and for those events the transverse momentum of the subleading top quark was examined. Then, the fraction of events with a subleading top-quark transverse momentum smaller than 200 GeV of the events with arbitrary subleading top-quark $p_T$ was calculated. For $t\bar{t} + \text{jets}$, this fraction was found to be around 37% and for $t\bar{t}H(b\bar{b})$ around 49%. This simple test shows that the additional Higgs boson leads to a higher amount of subleading top quarks with smaller momenta and thus confirms that chiral limit will be reached only for higher leading top-quark momenta.

Because the final state helicity composition depends significantly on the production phase space area, the angular distributions affected by this composition do as well. This can be observed especially well in the helicity frame as will be shown in section 9.3. This also means that by cutting on the final production phase space, for example, by demanding minimum transverse momenta of the top quarks, the spin correlations can be influenced and, for example, be reversed. Therefore, a separation criterion could be created exploiting this different phase space behaviour between $t\bar{t}H(b\bar{b})$ and $t\bar{t} + \text{jets}$. This might be used in the boosted regime where the top quarks have high transverse momenta.
9. Validation of the Modelling of Top-Quark-Antiquark Spin-Correlations at Generator Level

To validate the modelling and observe the spin correlations of top quarks, the exclusive dileptonic $13\text{TeV} \, \ttbar$ MC sample, cf. Tab. 5.1 in chapter 5, is used. Because kinematic cuts have a major influence on the distributions of the angular variables in the helicity frame, see section 11.2, no selection criteria are applied on the events for this chapter. However, the MC events are always generated with a set of basic kinematic cuts. Hence, small deviations from the theoretical calculations can be expected. Furthermore, only generator level quantities are used to validate the procedure implemented to obtain the distributions of angular variables sensitive to spin correlations. It will be checked if the expected phenomenological behaviour explained in chapter 8 is reproduced. The dileptonic $\ttbar$ sample is used because every interesting variable considering spin correlations in the $\ttbar$ system can be obtained with dileptonic events. From the MC sample, the generated quantities on parton level can directly be extracted. Therefore, no detector resolution, acceptance or efficiency effects have to be taken into account. In the following, the angular variables will be shown in the helicity frame and in the laboratory frame. The distributions are always normalized and the corresponding ratio plots are always calculated with respect to $t\bar{t}H(b\bar{b})$.

9.1. Validation in the Helicity Frame

In this section, the modelling of the angular variables in the helicity frame by the MC simulation will be compared to the phenomenological expectation. After the extraction of the four-momenta (in the laboratory frame) of the top quark, top antiquark, bottom quark, bottom antiquark, lepton, and antilepton from the MC event, a Lorentz boost into the center-of-mass system of the two top quarks is performed. Starting there, the procedure outlined in Fig. 8.5 is applied.

The first variable which will be examined is the three-dimensional opening angle between one top-quark and one top-antiquark decay product. The expected decay distribution is given in Eq. 8.18 and, in case of two leptonically decaying top quarks, visualized by the
Figure 9.1.: Visualization of the three-dimensional opening angle distributions in the helicity basis for different combinations of $t\bar{t}$ decay products. The distributions are normalized to unit area. The fit parameter $D$ describes the slope of the linear distributions. The coefficients $\alpha_{ij} = \alpha_i \cdot \alpha_j$ are the products of the spin analyzing power coefficients of the decay products $i$ and $j$. The expected values of $\alpha_{ij}$ in LO (NLO) perturbation theory are given in the bottom right corner. The distributions resulting from the MC simulation are shown in Fig. 8.7. The distributions according to Eq. 8.18 are used to calculate these coefficients. For the $\alpha_{ll}$ case, the values are set to the NLO theoretical values because otherwise the resulting system of ratio equations cannot be solved. In the upper left corner of Fig. 9.1, the results of the linear fits according to Eq. 8.18 are used to calculate these coefficients. For the $\alpha_{ll}$ case, the values are set to the NLO theoretical values because otherwise the resulting system of ratio equations cannot be solved. The value of $\chi^2/\text{ndof}$ for the fits is given as well.

The influence of the decay product choice due to the different analyzing powers (resulting in different slopes) is seen very clearly in the three shown distributions. The results of the fits confirm the phenomenological expectations in a very convincing way. The values of the fitted $\alpha_{ij}$ are compatible with the expected values to a good precision.
9.1. Validation in the Helicity Frame

addition, the $\chi^2/\text{ndof}$ values of the fits indicate a very good compatibility of the fitted theory model to the distribution of the MC data. Furthermore, the estimator $\hat{D} \approx 0.2261$ (with an uncertainty of $O(10^{-5})$) from Eq. 8.20 is also in agreement with the fitted value to a high precision. Finally, comparing the obtained value by the MC simulation with the full theory calculation ($D = \alpha_s^2 \text{O}_{\text{spin}} = 0.998^2 \cdot 0.236 \approx 0.235$) in NLO, see Tab. 8.2 and 8.1, under consideration that these values were calculated for 14 TeV, no conflicts are observed with a deviation of around 4%.

The next variables considered are the helicity angles and the coefficients $B_1$ and $B_2$ from Eq. 8.16. These variables basically represent the correlation of the spins with a given axis, meaning they are a measure of polarization. For a vanishing polarization, the distribution is expected to be flat. As already mentioned, the polarization is expected to be very small, because there is only a small contribution due to weak corrections and higher order QCD. The investigated dileptonic $t\bar{t}$ events were generated by the matrix element generator “Powheg” at NLO QCD. Thus, a small polarization should be expected. The resulting distributions for the helicity angles are shown in Fig. 9.2. These almost flat distributions confirm the theoretical predictions and show only a tiny polarization. In addition, the values obtained from linear fits according to Eq. 8.17 and the estimator values from Eq. 8.20, $\hat{B}_1 = 0.00489 \pm 0.00007$ and $\hat{B}_2 = 0.00538 \pm 0.00007$, are compatible within their statistical uncertainties.

Next, the two-dimensional distribution from Eq. 8.15 is investigated. The shape of the distribution, see Fig. 9.3, again shows good compatibility between the theoretical prediction
Figure 9.3.: The two-dimensional distribution of the helicity angles is shown. The shape observed is in high accordance with theory prediction. The fitted values are shown in the upper left corner with the $\chi^2$/ndof value of the two-dimensional fit.

and the simulated distribution from the MC sample. Again, a fit is performed to determine the correlation parameters. The obtained values for $B_1$ and $B_2$ are compatible with their expected values, with the one-dimensional fits to the helicity angle distributions from before, and their estimators. The obtained value for $C$ is very close to the expected value ($C = \alpha_2^2 \theta_{hel} = 0.998^2 \cdot 0.328 \approx 0.327$) with a deviation of about 4% and the $\chi^2$/ndof value of the fit does not indicate any conflict of the MC simulation with the theoretical model.

As was explained before, the two-dimensional distribution can be integrated to obtain a one-dimensional distribution in the product of the two cosines of the helicity angles. This distribution is generated as well and displayed in Fig. 9.4. The shape clearly shows the asymmetry, the fitted value for $C$, and the estimator $\hat{C} = 0.31366 \pm 0.00004$ shows good agreement with the value obtained by the two-dimensional fit.

The conclusion of this section is that all the distributions of the observed angular variables behave as theoretical calculations predict and the parameters used in the validation are all compatible to the theoretical values. The parameters themselves determined with multiple methods are all compatible within their statistical uncertainties. Because of the significant dependence of the distributions on cuts, differences in the resulting values of the correlation parameters are expected within different analyses/calculations using different cuts. The results shown throughout this section confirm that the implemented boosting procedure and the calculation of the variables work correctly. In addition, they work
9.1. Validation in the Helicity Frame

Figure 9.4.: The one-dimensional distribution corresponding to Eq. 8.16 obtained by integrating the two-dimensional distribution of the helicity angles. The distribution shows the expected asymmetry created by the spin correlations.

independently from the retrieval of the four-vectors of the particles and will therefore be used at reconstruction level as well.

However, since the aim of this thesis is to search for potential separation power of the angular variables, qualitative differences in the distributions of \( t\bar{t} \) and \( t\bar{t}H \) have to be found. Concerning the separation of \( t\bar{t} \) and \( t\bar{t}H \) with the variables investigated until this point, Fig. 9.5 shows that there is no separation possible between those two processes using spin correlation sensitive distributions. This can be explained as follows. Looking at the transverse momentum distribution of the top quarks for \( t\bar{t} + \text{jets} \) and \( t\bar{t}H \) in Fig. 9.6, it is observed that only a small part of the whole distribution is situated above 200 GeV. This is however roughly the region where a significant difference between \( t\bar{t} + \text{jets} \) and \( t\bar{t}H \), regarding the composition of \( \text{LL} + \text{RR} \) and \( \text{LR} + \text{RL} \) configurations, begins according to Fig. 8.8. In \( t\bar{t}H \) events, the \( \text{LL} + \text{RR} \) configuration makes up around 60% and the \( \text{LR} + \text{RL} \) around 40% for a range of \( p_{T}^{\text{top}} \ll 400 \text{ GeV} \). Complementary to this behaviour, \( t\bar{t} + \text{jets} \) has reversed configurations for \( p_{T}^{\text{top}} \gg 200 \text{ GeV} \) after traversing a point of even composition. In summary, this means that only for a very small fraction of events a complementary behaviour is expected because most of the events populate the area significantly below 200 GeV transverse momentum of the leading top quark. In this region, the helicity composition is comparable for the two processes. On the other hand, this means that by looking at the boosted regime a better distinction might be possible.
Figure 9.5.: Comparison of $t\bar{t} + \text{jets}$ and $t\bar{t}H$: On the left side the distribution of the three-dimensional opening angle between the two leptons of the top quarks is shown in the helicity frame. On the right side is the integrated distribution of the two helicity angles. In both distributions no significant separation can be found as the ratio plot indicates.

Figure 9.6.: The distribution of the transverse momentum of the leading top quark in different processes [133].
9.2. Validation in the Laboratory Frame

In this section, some angular variables will be shown in the laboratory (lab) frame. The problem with the lab frame is the fact that kinematic properties of the two processes $t\bar{t}$ and $t\bar{t}H$ generally differ much stronger than the effects expected by spin correlations. In addition, if the $t\bar{t}(b\bar{b})$ single-lepton analysis already uses variables which are sensitive to these kinematic differences, the expected gain of angular distributions in the lab frame is expected to be reduced.

The first variable to look at is the three-dimensional opening angle between one decay product from the top quark and one decay product from the top antiquark. Only considering kinematic arguments, it is expected that a small angle between the two decay products is overall favoured. However, a smaller accumulation of events around large angles compared to medium ones is expected as well. This can be understood by looking at the boost of the center-of-mass system of the top-quark pair.

In the center-of-mass system of the two top quarks, their momenta are antiparallel. If a particle decays with a high momentum in some direction, its decay particles are expected to move in the same direction. This can be seen in Fig. 9.7 where the angle between the lepton/antilepton and its parent top quark is shown. Together with the back-to-back flight direction of the top quarks, this leads to the fact that the angles between the decay particles of different top quarks are accumulated around large angles, see Fig. 9.8. Nevertheless, there is also a significant probability for decreasing angles. This is because of the high mass of the top quark and therefore the high possible momentum of the decay particles, which can compensate or overcome the boost of the top quarks. For $t\bar{t}H$ the qualitative behaviour is mostly the same. However, the distributions show that in $t\bar{t}H$ events the decay products follow the top quark directions more closely. This might be attributed to the additional boost of the $t\bar{t}$ system in $t\bar{t}H$ events due to the radiation of the Higgs boson.

In general, the top-pair center-of-mass system is additionally boosted in beam direction relative to the laboratory frame because of the different momentum fractions of the two colliding partons. This means that the angles of the top quarks change from 180 degrees to smaller angles. Because of the boost of the top quarks, the decay products are likely to follow. This can be seen in Fig. 9.9 for several combinations of decay products. The largest probability is now found at smaller angles. Because of the large probability at large angles in the first place, a small accumulation of events at large angles is left compared to the range of medium angles. The small angles now occur with the highest probability. This also emphasizes that the center-of-mass system of the top-quark pair is actually heavily boosted in the lab frame.

In $t\bar{t}H$ events, the $t\bar{t}$ center-of-mass system boost in the lab frame is dampened by the additional Higgs boson absorbing some of the boost of the $t\bar{t}H$ center-of-mass system. This leads to a higher number of events at medium angles and fewer at small angles.
9. Validation of the Modelling of Top-Quark-Antiquark Spin-Correlations at Generator Level

Figure 9.7.: The distribution of $\cos \theta$ is shown for $t\bar{t}$ and $t\bar{t}H$, where $\theta$ is the three-dimensional opening angle between lepton/antilepton and its parent top quark in the center-of-mass system of the top-quark pair. It can be seen that the decay particles favour small angles to their parent top quarks.
Figure 9.8.: The distribution of $\cos \theta$ is shown for $t\bar{t}$ and $t\bar{t}H$, where $\theta$ is the three-dimensional opening angle between the lepton and antilepton in the center-of-mass system of the top-quark pair. It can be seen that the decay particles angles still have their maximum around the opening angle of the top quarks.
Figure 9.9.: The distribution of cos θ is shown for $t\bar{t}$ and $t\bar{t}H$, where θ is the three-dimensional opening angle between one top-quark and one top-antiquark decay product in the laboratory frame.
compared to $t\bar{t}$ events. The distributions of $t\bar{t}$ in the laboratory frame can be compared with a theoretical prediction shown in Fig. 9.10. The theoretical distribution confirms the observed distributions of the MC simulation in Fig. 9.9.

The shape of the angular distributions in Fig. 9.9 compared within one process, $t\bar{t}$ or $t\bar{t}H$, is very similar, meaning different combinations of decay products do not change the distribution significantly. This might be due to the high mass of the top quark compared to its decay products leading to a domination of purely kinematic properties. Nevertheless, the distribution with mixed particles, meaning lepton and bottom quark or antilepton and bottom antiquark, show a small increase of probability for larger angles and a small decrease for smaller angles compared to the other two distributions. This could be due to the mixed couplings compared to the similar ones.

Another set of interesting variables to look at are the differences in the pseudorapidities $\Delta \eta_{ij} = |\eta_i - \eta_j|$ of the decay particles. The results of this can be observerd in Fig. 9.11. These distributions show a better discrimination power. For small differences in pseudorapidity, the probability of $t\bar{t}H$ is up to 10% higher than $t\bar{t}$. For larger differences, the situation is turned around. There, $t\bar{t}$ is more probable up to around 35%. Another measure of separation is the integral of the ROC curve, cf. section 7.3. Its value is approximately 0.53 for all four distributions shown in Fig. 9.11.

The last variable of interest is the difference in the azimuthal angle $\Delta \Phi_{ij} = |\Phi_i - \Phi_j|$ between two particles. In the case of the two leptons, the predicted distribution is shown in Fig. 9.12 together with the obtained results from the MC sample. The shapes of these two distributions are, despite differences in their generation (13 TeV and 14 TeV, different
Figure 9.11.: The distribution of $\Delta \eta$ is shown for $t\bar{t}$ and $t\bar{t}H$, where $\Delta \eta$ is the difference in the pseudorapidity of one top-quark and one top-antiquark decay product in the laboratory frame.
Figure 9.12.: The distribution of $\Delta \Phi$ is shown for $t\bar{t}$ and $t\bar{t}H$ on the right-hand side, where $\Delta \Phi$ is the difference in the azimuthal angle of one top-quark and one top-antiquark decay product in the laboratory frame. On the left-hand side, there is the theoretically predicted shape for $t\bar{t}$ from [126] for 14 TeV. The solid line includes spin correlations, the dashed line neglects them. Only the solid line is of importance for this comparison. It has to be kept in mind that this distribution was generated with another set of kinematic cuts, which can be found in the paper.
kinematic cuts), comparable. In addition, there is a small separation visible between $t\bar{t}$ and $t\bar{t}H$ leading to a ROC integral of around 0.51.

 Altogether, some separation gain between $t\bar{t} + \text{jets}$ and $t\bar{t}H(\text{bb})$ due to angular variables in the laboratory frame is indicated. In addition, good agreement between theoretical expectations and MC simulation is found. In the next chapter, only the irreducible $t\bar{t}b\bar{b}$ heavy flavor subset of the $t\bar{t} + \text{jets}$ events is compared to $t\bar{t}H(\text{bb})$ because the background due to $t\bar{t}$ events with additional light flavour jets can generally be reduced by demanding additional b-tags.

### 9.3. Phase Space Dependence of the Spin Correlations

As was mentioned in the last chapter, the composition of the helicity configurations depends on the phase space considered. Since the helicity configurations affect the angular distributions in the helicity frame, this behaviour should be observable by cutting on the transverse momenta of the top quarks. By looking at the left graph in Fig. 8.8, it is possible to see that the composition of the left-left and right-right configurations starts with an initial composition considering all events without cuts. Then, increasing cuts can be applied and the composition passes a point where both configurations are equally probable. This means that no correlation is found at all at this point. After that point, the composition is reversed and increases closing in to the chiral limit. The expectation is that with a variable sensitive to the correlations, for example the three-dimensional opening angle between the two leptons from the top quarks in the helicity frame, this behaviour should also be observed. Looking at Eq. 8.11 and 8.18, it is expected that this change in correlation results in a change of the slope of the distribution $\cos \theta_{ll}$ in the helicity frame. The slope starts with the value shown in Fig. 9.1 and should consequently get smaller because the probability of the two configurations get closer to each other. When the point of same probability is reached and the correlation disappears, the distribution should be flat. After that, the slope should again emerge but with a reversed sign because of the reversed correlations. This behaviour is visualized in Fig. 9.13 for a few points of production phase space defined by $p_T$ cuts on the leading top momentum.

For $t\bar{t}H$, shown in the right graph in Fig. 8.8, the behaviour is different. Because the composition stays the same for a large phase space region, the slope of the distribution should not change much as well. Only for a high $p_T$ cut the distribution should become flat, but for such a severe cut on the phase space, the number of events passing this cut is very small. Nevertheless, the expected behaviour can be observed in Fig. 9.13 until an approximately flat distributions is reached. The expected negative slope could not be reached because the number of events is too low to make a significant statement.

This behaviour can also be investigated, for example, for the distribution of $\Delta \Phi$ between the negatively charged lepton and the bottom quark. The corresponding distribution is
Figure 9.13.: The three-dimensional opening angle for the two leptons created in the decay of the top-quark pair is shown in the helicity frame for $t\bar{t} + \text{jets}$ on the left-hand side and $t\bar{t}H$ on the right-hand side. Different cuts on the transverse momentum of the leading top quark are applied to visualize the dependence on the production phase space. On the left-hand side the expected behaviour for $t\bar{t} + \text{jets}$ is observed. The initially positive slope shrinks for increasing cuts on the leading top $p_T$, then passes the zero correlation point with a vanishing slope and afterwards has a negative slope. For $t\bar{t}H$ the slope is reduced much slower.

displayed in Fig. 9.14. By imposing cuts on the momentum of the leading top quark the $\Delta\Phi$ distribution of $t\bar{t}$ changes significantly. Since there is no analytical form given for $\Delta\Phi$, it is not possible to check if the behaviour is exactly as expected. However, the $t\bar{t}H(b\bar{b})$ distribution behaves as expected as well. Since these cuts do not change the helicity composition of $t\bar{t}H$ significantly, the distribution is not expected to change in a major way. Neglecting the low number of events, this behaviour is approximately confirmed by the right plot.
9. Validation of the Modelling of Top-Quark-Antiquark Spin-Correlations at Generator Level

Figure 9.14: The azimuthal difference between the lepton from the top antiquark decay and the b-quark from the top decay is shown in the lab frame (left) and in the helicity frame (right). The cuts influence the distributions of $t\bar{t} +$ jets in the helicity frame more significantly than the $t\bar{t}H(b\bar{b})$ distributions.
10. Separation of Signal and Irreducible Background using Angular Variables at Generator Level

As was explained in chapter 4, the most difficult background to the $t\bar{t}H(b\bar{b})$ analysis is the irreducible $t\bar{t}b\bar{b}$ background because the same particles are found in the final state. This prevents the possibility to use the number of $b$-tags as a separator between those two processes. The $t\bar{t}b\bar{b}$ background features several different topologies with varying spin correlations. In this chapter, the various angular distributions typically used in spin correlation analysis are investigated with respect to their potential in improving the separation of the $t\bar{t}b\bar{b}$ irreducible background and the signal $t\bar{t}H(b\bar{b})$. However, the selection of $t\bar{t}b\bar{b}$ events from the $t\bar{t}$ + jets sample reduces the dataset enormously to approximately 0.5% of the inclusive $t\bar{t}$ sample. In Tab. 10.1, the number of events before and after the $t\bar{t}b\bar{b}$ selection is displayed. In addition, the corresponding two additional generator $b$-jets in $t\bar{t}b\bar{b}$ and $t\bar{t}H(b\bar{b})$ not coming from the top-quark decay are required to fulfill some kinematic cuts. This is done to operate in comparable kinematic regions of $t\bar{t}b\bar{b}$ and $t\bar{t}H(b\bar{b})$ to search for features in an area where a simple separation, for example by the invariant mass, is not possible. The cuts on the generator $b$-jets not coming from the top-quark decay were adapted from [12]:

\[
P_{T}^{b,\bar{b}} > 20 \text{ GeV} \tag{10.1}
\]
\[
|\eta^{b,\bar{b}}| < 2.4 \tag{10.2}
\]
\[
m_{b\bar{b}} > 100 \text{ GeV} \tag{10.3}
\]
\[
\Delta R_{b\bar{b}} > 0.4, \tag{10.4}
\]

Table 10.1.: The number of events in the semi- and dileptonic $t\bar{t}$ MC sample is shown before and after the selection of $t\bar{t}b\bar{b}$ events. In addition, the fraction of the number of events after the selection divided by the number of events before the selection is given.

<table>
<thead>
<tr>
<th># events</th>
<th>before $t\bar{t}b\bar{b}$ selection</th>
<th>after $t\bar{t}b\bar{b}$ selection</th>
<th>fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>dileptonic $t\bar{t}$ + jets sample</td>
<td>75383000</td>
<td>328978</td>
<td>0.44%</td>
</tr>
<tr>
<td>semileptonic $t\bar{t}$ + jets sample</td>
<td>157387260</td>
<td>758164</td>
<td>0.48%</td>
</tr>
</tbody>
</table>
where $m_{bb}$ denotes the invariant mass of the system of the two b-jets. These cuts reduce the number of $t\bar{t}b\bar{b}$ events in Tab. 10.1 further to approximately one quarter of their number. The initial number of approximately 4 million $t\bar{t}H(b\bar{b})$ events generated by “Powheg”, cf. Tab. 5.1, is reduced by approximately 20% by the cuts in Eq. 10.1. The distributions of the inclusive $t\bar{t} + \text{jets}$ events will also be shown in the following figures to get a better understanding of the differences.

### 10.1. Helicity Frame

The same variables which were used for validation in the previous section are now examined and compared with the results from [12] concerning a possible usage in separating $t\bar{t}H(b\bar{b})$ and $t\bar{t}b\bar{b}$. It has to be examined how the additional b-jets, which are mainly produced by initial state radiation of a gluon [135], affect the spin correlations of the $t\bar{t}$ system. Then, the differences compared to $t\bar{t}H(b\bar{b})$ events have to be found.

Again, starting with the three-dimensional opening angle between one top-quark and one top-antiquark decay product in the helicity frame, the results are shown in Fig. 10.1. First of all, the immensely reduced statistics combined with the small scale of the y-axis is observed. Nevertheless, it is visible that there is a small separation emerging in the helicity frame in case of the angle between the leptons (bottom right distribution). The slope of the $t\bar{t}b\bar{b}$ distribution is visibly smaller than in the inclusive $t\bar{t} + \text{jets}$ or $t\bar{t}H(b\bar{b})$ case. This means that the spin correlations are dampened for the $t\bar{t}b\bar{b}$ events. To highlight the reduction effect, in Fig. 10.2, this angle is shown again for $t\bar{t}H(b\bar{b})$ and $t\bar{t}b\bar{b}$ with a linear fit superimposed and the fitted correlation parameter $D$, cf. Eq. 8.18. The $\chi^2$/ndof values of 1.758 and 1.182 are increased because of the larger uncertainties but still do not indicate incompatibility. The fitted value of the parameter $D$ for $t\bar{t}b\bar{b}$, nevertheless, shows the reduction of the slope compared to the value for $t\bar{t} + \text{jets}$ in Fig. 9.1 or $t\bar{t}H(b\bar{b})$ in Fig. 10.2. The ROC integral between the two distributions, $t\bar{t}H(b\bar{b})$ and $t\bar{t}b\bar{b}$, is approximately 0.51. This effect is probably due to the fact that $t\bar{t}b\bar{b}$ events on average have higher top momenta, see Fig. 10.3. As was already shown, this different phase space results in a different composition of helicity states and in case of $t\bar{t}$ events leads to reduced spin correlations. The same behaviour was found in [12] and is also shown in Fig. 10.2 on the right-hand side. This difference in slopes might be used, by a multivariate classifier, to distinguish between signal and irreducible background.

The angles between one lepton and one bottom quark seem to indicate different slopes as well comparing $t\bar{t}b\bar{b}$ and $t\bar{t}H(b\bar{b})$, cf. Fig. 10.1. The value of the Kolmogorov-Smirnov (KS) test between $t\bar{t}b\bar{b}$ and $t\bar{t}H(b\bar{b})$ supports this statement by indicating incompatibility between the two distributions. However, the ratio plots show that this difference is very small because of the lower spin analyzing coefficient of the b-quark and will most certainly not have a significant influence. If the angle between the two bottom quarks from the top-quark decay is considered, the KS tests and the ratio plots show that no separation is
Figure 10.1.: The distribution of $\cos \theta$ is shown for $t\bar{b}b$ and $t\bar{H}$, where $\theta$ is the three-dimensional opening angle between one top-quark and one top-antiquark decay product in the helicity frame.
10. Separation of Signal and Irreducible Background using Angular Variables at Generator Level

Figure 10.2.: The distribution of $\cos \theta$ is shown on the left-hand side for $t\bar{b}b$ and $tH$, where $\theta$ is the three-dimensional opening angle between the two leptons from the top quarks in the helicity frame. The fitted parameter $D$ describing the slope of the distribution, cf. Eq. 8.18, and the goodness of fit values are shown as well. On the right-hand side, the same distribution is shown from [12].

Figure 10.3.: The transverse momentum distribution of the top quarks within the semileptonic top-quark pair events is shown. On the left-hand side, the distribution is shown for top-quark pair events with none or only additional light flavour jets ($t\bar{t} + l\bar{l}$). On the right-hand side, only $t\bar{b}b$ events are shown. It is clearly visible that the maximum of the distribution for $t\bar{b}b$ is at a higher value of the transverse momentum than in the light flavour case. In addition, the mean value and the tail of the distributions show the same behaviour.
Figure 10.4.: The distribution of the integrated helicity angles of the two leptons is shown.

provided by this variable. This is due to the combined low spin analyzing power of the two b-quarks.

The next distribution to consider is the one-dimensional distribution of the product of the two lepton helicity angles, cf. Eq. 8.16. It is shown on the left in Fig. 10.4. It can be observed that the asymmetry created by the spin correlations is reduced leading to a ROC integral of 0.52 between $t\bar{b}b\bar{b}$ and $t\bar{b}bH(b\bar{b})$. This is the same behaviour which was observed in the three-dimensional opening angle in the helicity frame. Therefore, the behaviour of reduced correlations seems to be consistent. The effect leads to a separation between $t\bar{b}b\bar{b}$ and $t\bar{b}bH(b\bar{b})$ visible in the ratio plot and confirmed by the KS test.

The difference in pseudorapidity shows no separation power in the helicity frame. Thus, the plots are shown in Fig. A.1 in the Appendix. However, in the center-of-mass frame of the top quarks the variable indicates some separation power. These plots are shown as well in the Appendix in Fig. A.2.

Regarding the azimuthal difference $\Delta \Phi$, only the variables including two leptons or a lepton and a bottom-type quark show a considerable difference in behaviour, see Fig. 10.5. The decay-particle combination with two bottom quarks has a combined spin analyzing power which is too small to observe visible differences. This observation is analogous to the behaviour found for the three-dimensional opening angles.
10. Separation of Signal and Irreducible Background using Angular Variables at Generator Level

Figure 10.5.: The distribution of $\Delta \Phi$ is shown for $t\bar{t}b\bar{b}$ and $t\bar{t}H(b\bar{b})$, where $\Delta \Phi$ is the azimuthal difference between one top-quark and one top-antiquark decay product in the helicity frame.
The angular variables in the helicity frame at generator level show that there are some qualitative and also quantitative differences in the behaviour of the observed distributions. The main reason behind the emerging differences of $t\bar{t}b\bar{b}$ and $t\bar{t}H(b\bar{b})$ is the different phase space behaviour of the spin correlations of $t\bar{t}$ and $t\bar{t}H$. Especially the dileptonic variables allow observation of this behaviour because their combined spin analyzing power is largest out of the considered decay products. In the single-lepton analysis where this thesis is placed, these dileptonic variables cannot be used at reconstruction level because of the applied selection criteria filtering events with two leptons. Instead, only the variables relying on $b$-quarks and on a lepton of the top-quark decay can be used. Moreover, to apply the Lorentz boosts to the helicity frame at reconstruction level, the complete top-quark system has to be reconstructed. However, the event reconstruction with a high number of jets and $b$-tags is a difficult task as will be shown in chapter 11.

10.2. Laboratory Frame

In the lab frame, a change of the distributions is investigated as well. The changes due to spin correlations are expected to be small in the lab frame for the three-dimensional opening angle, see, for example, Fig. 8 in [12]. If a separation is found, this should be due to kinematic differences.

The three-dimensional opening angles are shown in Fig. 10.6. The two processes, $t\bar{t}b\bar{b}$ and $t\bar{t}H(b\bar{b})$ show a separation. A difference occurs for large angles where $t\bar{t}b\bar{b}$ has a considerably smaller probability than $t\bar{t}H(b\bar{b})$ compared to the former case (Fig. 9.9) where no additional heavy flavour jets were required and the two distributions were more similar. At small angles, the $t\bar{t}b\bar{b}$ events occur at a higher probability than $t\bar{t}H(b\bar{b})$ and $t\bar{t} + \text{Jets}$ events. This could be due to the fact that most of the bottom quarks in $t\bar{t}b\bar{b}$ are created by hard initial-state gluon radiation. This additional hard radiation of a high transverse momentum gluon would lead to a boost of the $t\bar{t}$ system mainly in transverse direction to compensate the transverse momentum of the gluon. The strong transverse boost would then collimate the decay products of the top-quark pair system leading to an increase in smaller and a decrease in larger angles. The left plot in Fig. 10.7 showing how an angle of two top-quark pair decay products changes by requiring two additional high-$p_T$ $b$-quarks supports this theory. For higher momenta of the additional $b$-jets, small angles get more and large angles less probable. On the right-hand side in Fig. 10.7, the spectra of transverse momenta of the two additional $b$-quarks in $t\bar{t}b\bar{b}$ are shown. The two-dimensional histogram visualizes that a major part of the events tagged as $t\bar{t}b\bar{b}$ have momenta for the additional $b$-quarks even higher than 30 GeV and therefore fulfilling the initial hypothesis. This effect is enhanced additionally by the applied cuts on the additional $b$-quarks. Figure 8 in [12] confirms the observed behaviour.

In addition, the $\Delta \Phi$ variable shows an increased separation power throughout all decay products, see Fig. 10.8. This can be understood by reconsidering the already shown three-
Figure 10.6.: The distribution of $\cos \theta$ is shown for $t\bar{b}b$ and $t\bar{H}$, where $\theta$ is the three-dimensional opening angle between one top-quark and one top-antiquark decay product in the laboratory frame.
dimensional angles of the top-pair decay products where the conclusion was that $t\bar{t}b\bar{b}$ events favor smaller angles more than the inclusive $t\bar{t} + \text{jets}$ distributions. Since the threedimensional opening angle between two particles depends on the azimuthal difference and the difference in pseudorapidity, the behaviour of favouring small angles should impact the distributions of these variables as well. The most natural assumption is that smaller opening angles result in general in smaller $\Delta \Phi$ and $\Delta \eta$ compared to $t\bar{t} + \text{jets}$ events. This is observed in Fig. 10.8 and 10.9.

### 10.3. Training and Testing of a Boosted Decision Tree

In the previous sections, the variables with separation power useful for a multivariate classifier were shown. This classifier combines the information of all these variables and predicts if the event is a more signal- or a background-like process. Thus, its purpose is to separate the irreducible background $t\bar{t}b\bar{b}$ from the signal process $t\bar{t}H(b\bar{b})$. The multivariate classifier which is used in this section is a boosted decision tree (BDT) together with gradient boosting (BDTG), see section 7.3. The configuration of the BDT is shown in Tab. 10.2 and the input variables can be found in Tab. 10.3. This BDT only uses information at generator level to evaluate the maximum separation potential obtained by qualitative features of the distributions. It is expected that this performance will degrade if a similar approach is applied at reconstruction level.

For training and testing, the $t\bar{t}b\bar{b}$ events from the dileptonic $t\bar{t} + \text{jets}$ MC sample and the dileptonic events from the $t\bar{t}H(b\bar{b})$ sample are used, because these kinds of events contain all possible angular information including the variables with two leptons, which are not
10. Separation of Signal and Irreducible Background using Angular Variables at Generator Level

Figure 10.8.: The distribution of $\Delta \Phi$ is shown for $t\bar{t}b\bar{b}$ and $t\bar{t}H$, where $\Phi$ is the azimuthal difference between one top-quark and one top-antiquark decay product in the laboratory frame.
Figure 10.9.: The difference in pseudorapidity between the various decay products of the \( t\bar{t} \) system in the top-quark pair center-of-mass frame.
10. Separation of Signal and Irreducible Background using Angular Variables at Generator Level

Table 10.2.: The parameters of a gradient boosted decision tree which was used to evaluate the maximum separation power between $t\bar{t}b\bar{b}$ and $t\bar{t}H(b\bar{b})$ only using angular variables at generator level.

<table>
<thead>
<tr>
<th>#trees</th>
<th>shrinkage</th>
<th>bagged sample fraction</th>
<th>#cuts</th>
<th>maximum depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>0.025</td>
<td>0.15</td>
<td>40</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 10.3.: The variables which were used as input for the training of the boosted decision tree at generator level.

<table>
<thead>
<tr>
<th>frame</th>
<th>variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>lab frame</td>
<td>$\cos \theta_{ll}$, $\cos \theta_{lb}$, $\cos \theta_{ll}$, $\Delta \phi_{ll}$, $\Delta \phi_{lb}$, $\Delta \phi_{ll}$, $\Delta \phi_{bb}$</td>
</tr>
<tr>
<td>helicity frame</td>
<td>$\cos \theta_{ll}$, $\cos \theta_{lb}$, $\cos \theta_{ll}$, $\cos \theta_{ll}$, $\cos \theta_{ll}$, $\Delta \phi_{ll}$, $\Delta \phi_{ll}$, $\Delta \phi_{bb}$</td>
</tr>
<tr>
<td>tt cm frame</td>
<td>$\Delta \eta_{ll}$, $\Delta \eta_{lb}$, $\Delta \eta_{ll}$, $\Delta \eta_{bb}$</td>
</tr>
</tbody>
</table>

accessible in semileptonic events. A disadvantage of this method is the lower number of events compared to the semileptonic channel. The BDT is then trained and tested on statistically independent subsets of the MC samples.

The output shape of the BDT is given in Fig. 10.10 on the left-hand side where the distributions for the test and training sample are superimposed. The separation between signal ($t\bar{t}H(b\bar{b})$) and background ($t\bar{t}b\bar{b}$) is clearly visible and the distributions for training and testing dataset behave similar neglecting statistical fluctuations. The number of trees is quite low comparing to the number of trees used later within the BDTs in the analysis. The reason for this is the small number of events fulfilling the selection criteria. The number was chosen to this low value because otherwise the BDT showed significant overtraining. This is because there is only limited information gain in variables which are related by the same observed particles or by Lorentz boosts.

The corresponding ROC curve in Fig. 10.10 on the right, with an integral of 0.565, quantifies the possible separation for different cuts on the signal efficiency. The ROC integral corresponds to an increase of around 13% compared to a value 0.5, which corresponds to no additional information gain.

This study at generator level indicates that there is some potential to separate $t\bar{t}H(b\bar{b})$ and $t\bar{t}b\bar{b}$ by looking at angular variables typically used in spin correlation analyses. If some of this information could be preserved during the step from generator quantities to reconstructed ones, this could help with the discrimination of $t\bar{t}H(b\bar{b})$ from its irreducible background $t\bar{t}b\bar{b}$. In chapter 11, the step to reconstruction level is performed and it has to be evaluated how the variables influence the sensitivity of the analysis.
10.4. Improvements in the Boosted Regime

The boosted regime is the area of phase space where the top quarks and the Higgs boson have a high transverse momentum. This leads to a collimation of the decay products. The decay products of the hadronically decaying top quark and the Higgs boson are then analyzed with dedicated tools. The collimated jets are clustered into so-called fat jets. These fat jets are then also analyzed regarding their substructure with dedicated algorithms. The techniques in the boosted regime allow a significantly improved reconstruction of the hadronic top quark and Higgs boson because the jet-parton combinatorics are reduced [133, 136]. Because the reconstruction of the complete $t\bar{t}$ system is necessary for the calculation of the angular variables in the helicity frame, the boosted regime should have a significant advantage. As will be seen shown in the next chapter, the reconstruction of the $t\bar{t}$ system is a major difficulty because, in general, the reconstruction efficiency is not very high of events with a high number of jets and b-tags.

However, the real advantage should emerge if the improved reconstruction efficiency of the boosted category is combined with the behaviour of the $t\bar{t}$ helicity configurations explained in section 9.3. As was shown there, the spin correlations of $t\bar{t}$ are reversed quite fast for cuts on the leading top quark $p_T$ whereas in $t\bar{t}H$ this happens much slower and only for much higher cuts. This behaviour might be exploited as in the boosted regime the correlations should already be changed significantly in $t\bar{t}$ events but not in the $t\bar{t}H$ events. To demonstrate the prospects, the angular variables in the helicity frame are shown for events with a leading top transverse momentum greater than $200$ GeV in Fig. 10.11. There, the expected behaviour is observed. The slope of the $t\bar{t} + \text{jets}$ sample is reversed indicating reversed spin correlations whereas in $t\bar{t}H$ the slope is still roughly unchanged. Of course, the number of events passing this $p_T$ cut applied on the leading top quark is much smaller than the original event number. On the right-hand side of Fig. 10.11, the
Figure 10.11.: The three-dimensional opening angle between the two leptons from the top-quark pair decay on the left-hand side and the product of the two lepton helicity angles on the right-hand side. In both cases, a cut on the leading top transverse momentum greater than 200 GeV is applied to estimate the behaviour in the boosted category.
10.4. Improvements in the Boosted Regime

Figure 10.12.: On the left-hand side the output of the BDT is shown for the test and training samples for signal (tH(bb)) and background (ttbb) only using generator level information and events with a leading top-quark momentum greater than 200 GeV. The corresponding ROC curve is shown on the right-hand side. The integral of the ROC curve is 0.581.

The integrated distribution of the helicity angles shows the expected behaviour as well. The asymmetry of the tt and ttbb distribution is turned around compared to ttH.

Both ratio plots show that there are quality differences concerning the angular distributions. Consequently, another BDT is trained at generator level with the additional cut on the leading top-quark $p_T$. The configuration of the BDT is the same as in Tab. 10.2 except the shrinkage, which is chosen to be 0.015 to prevent overtraining. The corresponding BDT output and its ROC curve with an integral of 0.581 is displayed in Fig. 10.12. The output shape and the ROC integral confirm the expected improvement.

Combining this behaviour with the improved reconstruction efficiency, possibly enabling usage at reconstruction level, the boosted regime seems to be a promising candidate to use the spin correlations for additional separation of not only ttbb but also tt + jets from ttH. It is not easy to predict the impact of these variables on the boosted category, but even a small improvement would be a well understood and physically motivated advantage.
11. Separation of Signal and Irreducible Background using Angular Variables at Reconstruction Level

In Chapter 10, the angular variables used in spin correlation analysis were treated at generator level. The aim was to evaluate the maximum potential of these variables concerning the separation of $t\bar{t}H(b\bar{b})$ and $t\bar{t}b\bar{b}$. It was shown that there is promising separation potential available. In this chapter, a similar approach is performed at reconstruction level. However, the reconstruction level is more challenging than the generator level because of resolution as well as efficiency or acceptance effects. Except for the variables defined in the lab frame, the $t\bar{t}$ system needs to be reconstructed to be able to perform the Lorentz boosts to the other reference frames. This is already not simple for $t\bar{t}$ events but even harder for the reconstruction of $t\bar{t}H(b\bar{b})$ where the reconstruction efficiency is lower because of the higher number of jets and b-tags. As will be shown throughout this chapter, incorrectly reconstructed events have a major influence on the distributions of the angular variables. In addition, due to the already small statistics for $t\bar{t}b\bar{b}$ combined with the low reconstruction efficiency for high-jet multiplicity events, the number of correctly reconstructed events usable for training is very low. Furthermore, it has to be evaluated how much of the angular information is already included in the variables which are used as reference [13] at this point in time. This chapter uses the reconstructed objects and selected events which are also used in the $t\bar{t}H(b\bar{b})$ analysis in the semileptonic channel, cf. chapter 6. Because of this, dileptonic variables are not used.

11.1. Reconstruction with the $\chi^2$ Method

In order to reconstruct the $t\bar{t}$ system in events with at least four jets and two b-tags, a simple $\chi^2$ reconstruction is used. The $\chi^2$ expression is not used further in a kinematic fit (which would, for example, be allowed to change momenta of jets) but only to choose the seemingly right combination of jet and parton assignment. This means that it has to be decided which jets are assigned to the hadronic W boson, which jet is assigned to the leptonic top-decay quark (called leptonic bottom), and which one to the hadronic top-decay quark (hadronic bottom). To find the most probable combination of jet-parton
Separation of Signal and Irreducible Background using Angular Variables at Reconstruction Level

Assignment in an event where a $t\bar{t}$ hypothesis is assumed, the $\chi^2$ expression of the form

$$\chi^2 = \frac{(m_{j_1,j_2} - m_{W_{had,rec}})^2}{\sigma_{W_{had,rec}}^2} + \frac{(m_{j_1,j_2,j_3} - m_{had,rec})^2}{\sigma_{had,rec}^2} + \frac{(m_{l,\nu,j_4} - m_{lep,rec})^2}{\sigma_{lep,rec}^2}$$

(11.1)

is used, see [137] for the same method used in a different analysis. Here, $m_{j_1,j_2}$ is the invariant mass of the two jets which are assigned to the decay quarks of the $W$ boson, $m_{j_1,j_2,j_3}$ is the invariant mass of the aforementioned two jets and the jet corresponding to the bottom quark coming out of the hadronic top-quark decay. Furthermore, $m_{l,\nu,j_4}$ is the invariant mass of the lepton, the neutrino, and the jet assigned to the bottom quark from the leptonic top-quark decay. The quantities $m_{W_{had,rec}}, m_{had,rec}, m_{lep,rec}$ and $\sigma_{W_{had,rec}}, \sigma_{had,rec}, \sigma_{lep,rec}$ are the mean values and standard deviations of the reconstructed invariant masses of the hadronic $W$ boson, hadronic top quark, and the leptonic top quark obtained by a matching procedure between generator partons and reconstructed objects in MC events explained below.

In general, every possible combination of jets which can be assigned to $j_1-j_4$ is considered and the combination with the lowest value of the $\chi^2$ expression is used as the correct combination. However, the more jets an event contains, the more combinations have to be calculated. Without additional information, this leads to 12 combinations for an event with 4 jets, 60 combinations with 5 jets and 180 combination with 6 jets. Combinations in which the jets from the hadronic $W$ boson are switched, will be treated as one combination because there is no physical distinction possible and the $\chi^2$ expression has the same value. The combinatorics are additionally reduced by using only b-tagged jets as candidates for the b-quarks in the event. Combinations in which b-tagged jets are assigned to a decay quark from the $W$ boson are kept.

To calculate the invariant masses, the four-momenta of the reconstructed jets and of the lepton are used. In addition, the momentum of the neutrino is required. The momentum of the neutrino is estimated with the missing transverse energy $E_T$ and kinematic conservation of the invariant mass of its parent $W$ boson, giving the equation

$$p_W^2 = m_W^2 = (p_l + p_\nu)^2 = (E_l + E_\nu)^2 - (\vec{p}_l + \vec{p}_\nu)^2.$$  

(11.2)

Now the fact that $\vec{p}_\nu = (E_T, p_{\nu,z})$ is used. This leads in general to one or two possible solutions (depending on the argument of the square root emerging in the solution) for the longitudinal momentum of the neutrino. In case of the two solutions, both are used within the combinations, therefore doubling the original combinatorics. For every combination, the $\chi^2$ value is computed and the combination with the lowest value is chosen as the optimal combination.

The values of $m_{W_{had,rec}}, m_{had,rec}, m_{lep,rec}$ and $\sigma_{W_{had,rec}}, \sigma_{had,rec}, \sigma_{lep,rec}$ are generally extracted by using semileptonic $t\bar{t}$ MC samples and assigning the reconstructed objects to their corresponding partons at generator level. This is usually done by demanding a $\Delta R$ distance between reconstructed object and parton which is smaller than a defined value, for example
0.4. These events which are assigned correctly with a high probability are consequently used to calculate these invariant masses with their corresponding reconstructed objects. The mean values and the standard deviations of the resulting invariant mass distributions are then calculated (or fitted) and used in the $\chi^2$ expression. This evaluation of the invariant masses and their uncertainties for the $\chi^2$ expression takes resolution and response effects into account. After finding the optimal combination, the four-vectors of the partons are assigned with the four-vectors of the reconstructed jets. Finally, the angular variables can be calculated with the reconstructed objects.

After applying the selection for the number of jets and b-tags as well as the lepton selection, cf. chapter 4 and 6, the explained reconstruction is performed on the remaining events. In Tab. 11.1, the reconstruction efficiencies are tabulated. The reconstruction is tagged as correct if the reconstructed four-vectors of the top quark, the top antiquark, and the two bottom-type quarks from the $t\bar{t}$-system are within a $\Delta R = 0.4$ cone radius of their respective generator partons. Because of this, it does not matter if the two jets from the W boson are switched. Thus, the reconstruction efficiency is defined as the fraction of correctly reconstructed events and the number of events where a correct $\Delta R = 0.4$ jet-parton assignment on MC events is possible at all (such events are often called matched events [138]). In this thesis, matched events are defined as events in which for every final state parton of the $t\bar{t}$-system, a jet can be found within a radius of $\Delta R = 0.4$.

The events were reconstructed with different hypotheses. First, the reconstruction was done with a $t\bar{t}$ hypothesis corresponding to Eq. 11.1. Another procedure using varying hypotheses was tested as well. First, the reconstruction is performed with a $t\bar{t}H(bb)$ hypothesis. In this case, another term corresponding to the invariant mass of the Higgs boson is added to the $\chi^2$ expression. The $t\bar{t}H(bb)$ hypothesis is only applied if six jets, four of them b-tagged, are present in the event. If this was not the case, the reconstruction falls back to the $t\bar{t}$ hypothesis. In Tab. 11.1, this procedure is named “mixed hypothesis”. In principle, this combines the two hypotheses to have the best possible reconstruction efficiency in almost all categories. This leads to the fact that in all categories below six jets, the $t\bar{t}$ hypothesis is used. For categories containing six or more jets, the $t\bar{t}H(bb)$ hypothesis is used. This has the advantage that this hypothesis has significantly higher reconstruction efficiencies for $t\bar{t}H(bb)$ and $t\bar{t}b\bar{b}$ events with six or more jets and more than two b-tags.

Tab. 11.1 shows that the reconstruction efficiency drops rapidly as expected with a higher number of jets and b-tags. It is worth to noting that the reconstruction can handle one additional jet without losing efficiency comparing the categories 4j2t and 5j2t.

11.2. Impact of the Event Selection

In this chapter, only events fulfilling the event selection criteria are considered. The shape of the distributions of the angular variables therefore are expected to show some differences
Table 11.1.: Reconstruction efficiencies for semileptonic $t\bar{t}$ + jets and $t\bar{t}H(b\bar{b})$ events. The definition of the reconstruction efficiency can be found in the text.

<table>
<thead>
<tr>
<th>mixed hypothesis</th>
<th>4j2t</th>
<th>4j3t</th>
<th>4j4t</th>
<th>5j2t</th>
<th>5j3t</th>
<th>5j4t</th>
<th>6j2t</th>
<th>6j3t</th>
<th>6j4t</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}$ + jets</td>
<td>34%</td>
<td>27%</td>
<td>15%</td>
<td>34%</td>
<td>25%</td>
<td>15%</td>
<td>13%</td>
<td>9%</td>
<td>8%</td>
</tr>
<tr>
<td>$t\bar{t}b\bar{b}$</td>
<td>15%</td>
<td>10%</td>
<td>9%</td>
<td>12%</td>
<td>11%</td>
<td>6%</td>
<td>10%</td>
<td>11%</td>
<td>12%</td>
</tr>
<tr>
<td>$t\bar{t}H(b\bar{b})$</td>
<td>8%</td>
<td>5%</td>
<td>1%</td>
<td>9%</td>
<td>10%</td>
<td>5%</td>
<td>10%</td>
<td>13%</td>
<td>17%</td>
</tr>
<tr>
<td>$t\bar{t}$ hypothesis</td>
<td>4j2t</td>
<td>4j3t</td>
<td>4j4t</td>
<td>5j2t</td>
<td>5j3t</td>
<td>5j4t</td>
<td>6j2t</td>
<td>6j3t</td>
<td>6j4t</td>
</tr>
<tr>
<td>$t\bar{t}$ + jets</td>
<td>34%</td>
<td>27%</td>
<td>15%</td>
<td>34%</td>
<td>25%</td>
<td>15%</td>
<td>15%</td>
<td>11%</td>
<td>9%</td>
</tr>
<tr>
<td>$t\bar{t}b\bar{b}$</td>
<td>15%</td>
<td>10%</td>
<td>9%</td>
<td>12%</td>
<td>11%</td>
<td>6%</td>
<td>11%</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>$t\bar{t}H(b\bar{b})$</td>
<td>8%</td>
<td>5%</td>
<td>1%</td>
<td>9%</td>
<td>10%</td>
<td>5%</td>
<td>9%</td>
<td>7%</td>
<td>6%</td>
</tr>
</tbody>
</table>

to the ones from chapter 10 where no event selection was applied. In this section, these differences are investigated by looking at the generator distributions. This allows to isolate the effects of the selection criteria which reduced the number of events from around 150 million semileptonic $t\bar{t}$ events to approximately 11 million. The number of semileptonic $t\bar{t}H(b\bar{b})$ events was reduced from approximately 4 million events to around 500000. The number of $t\bar{t}b\bar{b}$ events in the semileptonic $t\bar{t}$ sample was affected by a reduction from around 760000 to approximately 170000 events.

In Fig. 11.1, the distributions for $\cos \theta_{b\bar{b}}$ and $\cos \theta_{lb}$ is shown for all events at generator level in the lab frame. The distribution of $t\bar{t}$ shows a strong dip for small angles. Since the minimum number of $b$-tags for events considered in the analysis is two and because it is clear that two $b$-quarks with a very small angle are very unlikely to be clustered into two separate jets, the number of events having such a small angle between the $b$-quarks should be very small. Because the additional $b$-jets created in $t\bar{t}b\bar{b}$ and $t\bar{t}H(b\bar{b})$ events can compensate for the missing $b$-jet, those distributions are not affected as much by this cut on the minimum number of $b$-tags. However, the $t\bar{t}b\bar{b}$ events are even less affected because by selection two additional separated generator $b$-jets within the acceptance are required for those events by the $t\bar{t}b\bar{b}$ tagging-mechanism. An analogous behaviour can be observed on the right-hand side of Fig. 11.1. In this case, the $b$-jets are discarded by the object selection if a lepton is within a radius of $\Delta R = 0.4$ of a jet which is the case if the lepton and the bottom quark have very small angles between them. The additional $b$-jets in the $t\bar{t}b\bar{b}$ and $t\bar{t}H(b\bar{b})$ events have a compensating effect for this distribution as well. For angular areas other than small angles, the distributions in the lab frame stay comparable to chapter 10.

The variables in the helicity frame are especially sensitive to kinematic cuts leading to severely distorted distributions. Thus, the impact of the event selection is examined by looking at the changes of the expected distributions of the three-dimensional opening angle. In Fig. 11.2, this angle is shown for the two bottom quarks from the top-quark pair decay in $t\bar{t}$ + jets events and the for the angle between the lepton and the bottom quark. In addition, several cuts from the event selection are applied to observe their impact on the
11.3. Impact of Incorrect Reconstruction

Due to the non-vanishing resolution of the invariant mass, incorrect combinations can actually have a smaller $\chi^2$ value and can therefore be selected instead of the correct combination. Hence, the reconstruction of an event can be incorrect, i.e. obtain a wrong jet-parton assignment. This not only has a influence on the calculated variables but can impact the distributions also in form of incorrect Lorentz boosts. This leads to the fact that the variables in frames other than the lab frame are expected to be affected more drastically by incorrect reconstructions. This is especially problematic for the $t\bar{t}H(b\bar{b})$ and $t\bar{t}b\bar{b}$ events: Because of the huge combinatorics in these events, their reconstruction efficiency is low. To visualize the effect, the distributions from some of the variables are shown for correctly reconstructed events and for events in which the reconstruction is incorrect. The fraction of correct and incorrect reconstructions for all events which fulfilled the selection criteria are given in Tab. 11.2. The results in the table confirm that the reconstruction efficiency for
11. Separation of Signal and Irreducible Background using Angular Variables at Reconstruction Level

Figure 11.2: The distributions of the angle between the two bottom quarks (left) and between the bottom and the lepton (right) from the top-quark decay in the helicity frame at generator level. Several cuts adapted from the event selection are applied to highlight the impact on the distributions.

Table 11.2: The fraction of correctly and incorrectly reconstructed events with respect to all events for the different samples. Only events fulfilling the event and object selection are considered.

<table>
<thead>
<tr>
<th>events</th>
<th>t\bar{t}</th>
<th>t\bar{t}b</th>
<th>t\bar{t}H(b\bar{b})</th>
</tr>
</thead>
<tbody>
<tr>
<td>correct reconstruction</td>
<td>26%</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>incorrect reconstruction</td>
<td>74%</td>
<td>93%</td>
<td>93%</td>
</tr>
</tbody>
</table>
As it turns out, the distributions of incorrectly reconstructed events show a completely different behaviour in case of the generator variables and the reconstructed ones. The plots in Fig. 11.3 show the distributions of $\cos \theta_{bb}$ in the lab frame at generator level and reconstruction level for events which were reconstructed correctly. The events which were reconstructed correctly show the same behaviour at reconstruction level as the generator level distributions. This is expected because of the correctly chosen combination. Hence, only resolution effects create small differences between them.

For the incorrectly reconstructed events in Fig. 11.4, the behaviour is completely different. The generator distributions on the left-hand side in Fig. 11.4 behave comparable to the distributions in Fig. 10.6, keeping in mind that the event selection was used. However, the reconstructed distributions in the same figure are changed completely. The corresponding ratio plots show a highly complementary behaviour.

The maximum of the $t\bar{t}b\bar{b}$ distribution is now situated at large angles. This is probably due to the fact that events with very small angles are in general hard to reconstruct because the jets cannot be separated correctly. This can have the consequence that the two bottom
Figure 11.4.: The generator (left) and reconstructed (right) distribution of the angle between the two bottom quarks from the top decay in the lab frame for incorrectly reconstructed events.

Jets from the top-quark decays are clustered into one jet. The second jet considered in the $\chi^2$ reconstruction could be an additional bottom jet unlikely to be close to the first one. This can, for example, be attributed to the already mentioned initial-state high-$p_T$ gluon generating most of the additional b-jets. As already explained, the high transverse momentum of the gluon leads to a boost of the tt system in opposite transverse direction. The direction of the, consequently collimated, decay products (leading to merged jets in the first place) should be roughly opposite as well, therefore too resulting in larger angles. Moreover, the additional b-jets create more combinatorics resulting in a higher number of wrong jet-parton assignments. These are possible reasons for incorrect reconstructions and why combinations with larger angles are preferred. This is supported by the generator distributions of the correctly reconstructed events in Fig. 11.3, where at small angles only few of the events are found and the distributions grow towards larger angles.

The signal process tH(b¯b) shows a quite different behaviour. This might be attributed to the fact that the additional b-jets from the Higgs boson are created solely within the final state. The kinematic distributions of the b-jets from the Higgs boson and the tt pair should hence be similar [135]. Consequently, there is a qualitative kinematic difference between tH(b¯b) and tbb although a similar kinematic selection is applied on the additional b-jets, cf. Eq. 10.1. The fact that this effect does not depend on the observed top decay products would support such a kinematic argument. Thus, the $\chi^2$ reconstruction is affected differently.
11.3. Impact of Incorrect Reconstruction

Figure 11.5.: The migration of angles caused by the reconstruction is shown. Events were selected which have an angle within a specific range at generator level. For those events, the reconstructed angles are shown in the histogram for $t\bar{t}$ + jets (top left), $t\bar{t}$ + $b\bar{b}$ (top right), and $t\bar{t}H(b\bar{b})$ (bottom)

The migration of events/probability regarding the three-dimensional opening angles due to the reconstruction is visualized in Fig. 11.5. There, the reconstructed angles of events which have a specific angle at generator level are shown. All distributions feature edges where the cut at generator level is put and show a significant amount of migrated events. However, the shape of the migration indicates qualitative distinctions between the different processes. The inclusive $t\bar{t}$ + jets events indicate a quite similar up and down migration. Only into areas of small angles, fewer events migrate. For very large angles, a small increase of migrating events is observed. Because these effects are small, the overall shape of the reconstructed $t\bar{t}$ + jets events is quite similar to the generator distribution, see Fig. 11.4.

The $t\bar{t}$ + $b\bar{b}$ events show a different behaviour. In this process, a significantly larger fraction of events migrate into the area of larger angles. A possible reason for this was already given above. A smaller accumulation of migrated events is found at very small angles. This might be attributed to events where the bottom-type quarks from the top-quark decay are too far spread and therefore the $\chi^2$ reconstruction uses a combination of one of the b-jets from the top-quarks and one of the additional b-jets or both additional b-jets as the b-jets from the top-quark decay.
Figure 11.6.: The generator (left) and reconstructed (right) distribution of the azimuthal difference between the two bottom quarks from the top-quark decay in the lab frame is shown.

Yet different behaviour is shown by t\bar{t}H(b\bar{b}) events. These distributions show that significantly more events migrate from larger angles to smaller angles than the other way around. This is essentially the reason for the apparent separation between t\bar{t} + jets or t\bar{t}b\bar{b} and t\bar{t}H(b\bar{b}) after the reconstruction. To ensure that this behaviour is correctly simulated, an additional t\bar{t}b\bar{b} and t\bar{t}H(b\bar{b}) sample from a different MC generator was used. With this sample, the distribution of the incorrectly reconstructed events was examined again and the behaviour was reproduced. A large fraction of t\bar{t}b\bar{b} events migrated to larger angles and the t\bar{t}H(b\bar{b}) events have an accumulation at smaller angles, cf. Fig. A.3 in the Appendix.

Since a major part of the t\bar{t}H(b\bar{b}) and t\bar{t}b\bar{b} events are not reconstructed correctly, cf. Tab. 11.2, it has to be decided if the training is performed on correctly reconstructed events with the aim to only learn physical features or it might also be possible to use the incorrectly reconstructed events if the agreement with data is good enough. This would have the advantage of much higher available statistics, the extra separation created by incorrect reconstruction, and a straightforward use within the existing t\bar{t}H(b\bar{b}) single-lepton analysis. Nevertheless, the reconstruction does not have this positive effect on every variable. Considering, for example, the azimuthal difference \Delta \Phi_{b\bar{b}} in Fig. 11.6, it is observed that no separation is created by the reconstruction. The opposite is the case: the separation between t\bar{t}H(b\bar{b}) and t\bar{t}b\bar{b} is reduced significantly. Hence, the incorrectly reconstructed events can also have negative influence on the possible separation.
11.4. Data/MC Agreement

For the classifier to be valid, the variables on which it is trained have to modeled well enough to give a sufficiently accurate description of the shape and correlations of the data. Therefore, the agreement between the simulated and measured data has to be verified. This is done by looking at histograms where the MC prediction for the different processes is stacked and the data is superimposed, see Appendix A.3.

The histograms show that the agreement between the MC prediction and the measured data is good regarding the shape of the angular variables. Its worth noting that the shape-changing behaviour concerning incorrect reconstructions is modeled well. None of the categories shows significant incompatibility between data and MC simulation. However, the overall normalization of the MC prediction does not seem to match the data although this effect is mostly within the uncertainties. This can be seen among others in the ≥ 6 jets and 2 b-tags or the 4 jets and 3 b-tags categories. However, this should not be a serious problem because this can be compensated by nuisance parameters describing rate uncertainties in the later fit. The results lead to the conclusion that the angular variables are described well. Consequently, the classifier, later used in the analysis, will be trained on all events no matter if they are reconstructed correctly or not. This also means that the angular variables can be added in a straightforward way to the categories of the analysis.

The remaining variables which were used as input for the final classifier of the analysis are listed in Tab. 11.3.

11.5. Training and Testing of a Boosted Decision Tree

In section 10.3, a BDT was trained only with the angular variables at generator level to get an approximation of the maximum possible separation. In this section, the same will be done using only reconstructed quantities, which can be calculated with the events used in the tH(bb) single-lepton channel analysis. The variables chosen here are all calculated in the laboratory frame to remove the impact of incorrect Lorentz boosts due to incorrectly reconstructed events. The training will be performed using all events which are reconstructed correctly to investigate the impact from the reconstruction step and the missing dileptonic variables. The resolution effects and the much lower statistics are expected to distort the BDT output shape. Nevertheless, it is expected that the output shape of the BDT trained with correctly reconstructed events should at least somehow resemble the output shape of the one trained at generator level. This can be seen in Fig. 11.7 where the BDT output is shown for a training with only correctly reconstructed events. The parameters of the BDT are shown in Tab. 11.4. Due to the much lower statistics, the output shape is distorted significantly showing large statistical fluctuations. Still, the BDT
Table 11.3.: The input variables for the BDT training in every category. An explanation of the specific variables can be found in [13], Tab. 6. Moreover, a new Particle Swarm Optimization was performed in the 6j4t category considering the angular variables, see section 11.6. The results of this optimisation are given in the last row.

<table>
<thead>
<tr>
<th>category</th>
<th>variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 jets, 3 b-tags</td>
<td>B-tagging likelihood ratio, third highest CSV, HT, $H_1$, sphericity, dev from avg CSV (tags), $\Delta R$(lepton,jet), $(\sum \text{jet p}_{T})/(\sum \text{jet E})$</td>
</tr>
<tr>
<td>(4j3t)</td>
<td></td>
</tr>
<tr>
<td>4 jets, 4 b-tags</td>
<td>$\sum \text{p}<em>{T}$(lepton, jet, met), avg CSV (tags), $H_3$, $(\sum \text{jet p}</em>{T})/(\sum \text{jet E})$ aplanarity, $M_2$ of min $\Delta R$(tag, tag)</td>
</tr>
<tr>
<td>(4j4t)</td>
<td></td>
</tr>
<tr>
<td>5 jets, 3 b-tags</td>
<td>B-tagging likelihood ratio, $\sum \text{p}_{T}$(lepton, jet, met), $M_2$ of min $\Delta R$(tag, tag), $H_2$, sphericity, avg CSV (tags), avg $\Delta R$(tag, tag), jet 3 $p_T$, $\Delta R$(lepton,jet), max $\Delta \eta$(tag, avg tag $\eta$)</td>
</tr>
<tr>
<td>(5j3t)</td>
<td></td>
</tr>
<tr>
<td>5 jets, ≥ 4 b-tags</td>
<td>aplanarity, $M_2$(tag, tag) closest to 125, $\sum \text{p}_{T}$(lepton, jet, met), $M_3$, avg CSV (tags), HT, avg $\Delta \eta$(jet, jet), avg $\Delta R$(tag, tag), $M_2$ of min $\Delta R$(tag, tag)</td>
</tr>
<tr>
<td>(5j4t)</td>
<td></td>
</tr>
<tr>
<td>≥6 jets, 2 b-tags</td>
<td>B-tagging likelihood ratio, $M_2$ of min $\Delta R$(tag, tag), $H_2$, avg CSV (tags), $\sum \text{p}<em>{T}$(lepton, jet, met), aplanarity, $H_1$, $(\sum \text{jet p}</em>{T})/(\sum \text{jet E})$, BDT_common5_input_Mlb, max $\Delta \eta$(tag, tag)</td>
</tr>
<tr>
<td>(6j2t)</td>
<td></td>
</tr>
<tr>
<td>≥6 jets, 3 b-tags</td>
<td>aplanarity, $\sqrt{\Delta \eta(t_{lep}, bb) \times \Delta \eta(t_{had}, bb)}$, $(\sum \text{jet p}<em>{T})/(\sum \text{jet E})$, min $\Delta R$(tag, tag), dev from avg CSV (tags), $\sum \text{p}</em>{T}$(lepton, jet, met), B-tagging likelihood ratio</td>
</tr>
<tr>
<td>(6j3t)</td>
<td></td>
</tr>
<tr>
<td>≥6 jets, 4 b-tags</td>
<td>best Higgs mass, $M_2$(tag, tag) closest to 125, mass(jets,lepton,MET), fourth-highest b-tag, $\sum \text{p}_{T}$(lepton, jet, met), fifth-highest CSV</td>
</tr>
<tr>
<td>(6j4t)</td>
<td></td>
</tr>
<tr>
<td>PSO optimized</td>
<td>best Higgs mass, $M_2$(tag, tag) closest to 125, mass(jets,lepton,MET), fourth-highest b-tag, $\sum \text{p}<em>{T}$(lepton, jet, met), fifth-highest CSV, $\Delta \Phi</em>{lb}$, $\Delta \eta_{lb}$</td>
</tr>
</tbody>
</table>

Table 11.4.: The configuration of the BDT which was trained with correctly reconstructed $t\bar{t}b\bar{b}$ and $t\bar{t}H(bb)$ events.

<table>
<thead>
<tr>
<th>#trees</th>
<th>shrinkage</th>
<th>bagged sample fraction</th>
<th>#cuts</th>
<th>maximum depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>210</td>
<td>0.005</td>
<td>0.15</td>
<td>10</td>
<td>2</td>
</tr>
</tbody>
</table>
11.6. Angular Variables within the MVA Classification of the \(t\bar{t}H(b\bar{b})\) Analysis

This section describes the usage of the angular variables within the existing \(t\bar{t}H(b\bar{b})\) semileptonic analysis [13]. In this analysis, events are categorized based on their number of jets and number of \(b\)-tagged jets. In these categories, several variables are chosen to separate the signal process \(t\bar{t}H(b\bar{b})\) from the background processes. These variables are combined into the outputs of the BDTs trained for each category. Consequently, all the distributions of the BDT output variables in all categories are fitted simultaneously to the data. From this fit, an observed upper limit on the signal strength modifier can be obtained. The signal strength modifier \(\mu = \sigma_{\text{measured}} / \sigma_{\text{SM}}\) is a measure of the strength of the signal process in the data, cf. section 7.1. This observed upper limit can be compared to the expected upper limit, which can be obtained by using the background prediction from MC simulation instead of measured data in the fit. As was already explained in chapter 7, the expected upper limit is a measure of sensitivity of the analysis towards the sought process. In this thesis, only the (asymptotic) expected upper limit on the signal strength modifier is calculated at 95% confidence level.

To obtain the outputs of the BDTs, they first have to be trained in every category. The decision on which input variables are used in a specific category is automated by the so-
11. Separation of Signal and Irreducible Background using Angular Variables at Reconstruction Level

Table 11.5.: The different BDTs used in the specific categories of the ttH(bb) single-lepton analysis [13]. The parameters were chosen by the Particle Swarm Optimization. The output distributions of each of these BDTs is fitted simultaneously within the final fit. In addition, another Particle Swarm Optimization was performed in the 6j4t category including the angular variables. The results of this optimisation are given in the last row.

<table>
<thead>
<tr>
<th>category</th>
<th>#trees</th>
<th>shrinkage</th>
<th>bagged sample fraction</th>
<th>#cuts</th>
<th>maximum depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>4j3t</td>
<td>1333</td>
<td>0.0380</td>
<td>0.49</td>
<td>38</td>
<td>2</td>
</tr>
<tr>
<td>4j4t</td>
<td>408</td>
<td>0.0237</td>
<td>0.25</td>
<td>34</td>
<td>2</td>
</tr>
<tr>
<td>5j3t</td>
<td>1312</td>
<td>0.0286</td>
<td>0.4</td>
<td>36</td>
<td>2</td>
</tr>
<tr>
<td>5j4t</td>
<td>700</td>
<td>0.0169</td>
<td>0.49</td>
<td>49</td>
<td>2</td>
</tr>
<tr>
<td>6j2t</td>
<td>507</td>
<td>0.036</td>
<td>0.16</td>
<td>61</td>
<td>2</td>
</tr>
<tr>
<td>6j3t</td>
<td>1043</td>
<td>0.0326</td>
<td>0.11</td>
<td>60</td>
<td>2</td>
</tr>
<tr>
<td>6j4t</td>
<td>932</td>
<td>0.0166</td>
<td>0.51</td>
<td>46</td>
<td>2</td>
</tr>
<tr>
<td>6j4t PSO</td>
<td>896</td>
<td>0.0311</td>
<td>0.20</td>
<td>39</td>
<td>2</td>
</tr>
</tbody>
</table>

called Particle Swarm Optimisation (PSO). This algorithm chooses the variables for every category and the best possible parameters of the BDTs based on the achieved ROC-integral performance [139]. Afterwards, the BDTs are trained using the results of the PSO. In this thesis, the results from the PSO performed in the latest ttH(bb) single-lepton analysis [13] are used and can be found in Tab. 11.5. The considered input variables were already shown in Tab. 11.3.

The training is performed in two different ways to compare and evaluate the influence of the angular variables. The first training is performed without the angular variables analogous to the ttH(bb) single-lepton analysis. The results of this training and the subsequently obtained expected upper limit is used as a benchmark and can be found in the column “without angular variables” in Tab. 11.6. Following this, another training is performed including the angular variables. Only variables defined in the lab frame are considered because the large number of incorrectly reconstructed events, cf. section 11.3, would result in a large number of incorrect Lorentz boosts. In a simple approach, all angular variables available in the semileptonic tt decay are added as input variables to the already used ones in Tab. 11.3. The BDTs are then trained with the same parameters as in the previous training, given in Tab. 11.5. The resulting median expected limits at 95% CL can be found in the column “with angular variables” in Tab. 11.6 and the angular input variables for every category in Appendix A.3 together with the BDT output shapes.

Concluding, a positive trend is observable in almost every category. The impact of just adding the angular variables to the existing training variables with the same training parameters is highest in the categories with a lower number of jets where most of the events are found. In these categories, the already explained effect of incorrect reconstructions seems to be most significant as can also be observed in the Data/MC comparison plots, for
Table 11.6.: The median expected limits at 95% CL obtained in each category without using the angular variables and by just adding them as additional input variables to the training in Tab. 11.5 (“with angular variables”). In addition, the “PSO” column shows the results obtained with the new Particle Swarm Optimisation including the angular variables. The new optimisation was only performed in the 6j4t category. The combined limit in this row was obtained by using the results from the “with angular variables” row except in the 6j4t category in which the newly optimized result was used. The given uncertainties indicate the 68% confidence intervals.

<table>
<thead>
<tr>
<th>category</th>
<th>without angular variables</th>
<th>with angular variables</th>
<th>PSO</th>
</tr>
</thead>
<tbody>
<tr>
<td>4j3t</td>
<td>10.45±4.13</td>
<td>9.84±3.97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.93</td>
<td>-2.69</td>
<td></td>
</tr>
<tr>
<td>4j4t</td>
<td>7.72±3.78</td>
<td>7.41±3.63</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-2.40</td>
<td>-2.32</td>
<td></td>
</tr>
<tr>
<td>5j3t</td>
<td>6.91±2.56</td>
<td>6.72±2.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.85</td>
<td>-1.79</td>
<td></td>
</tr>
<tr>
<td>5j4t</td>
<td>4.33±1.95</td>
<td>4.33±1.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.27</td>
<td>-1.27</td>
<td></td>
</tr>
<tr>
<td>6j2t</td>
<td>11.78±5.03</td>
<td>11.75±5.11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-3.47</td>
<td>-3.44</td>
<td></td>
</tr>
<tr>
<td>6j3t</td>
<td>5.61±2.12</td>
<td>5.52±2.13</td>
<td>3.42±1.49</td>
</tr>
<tr>
<td></td>
<td>-1.48</td>
<td>-1.48</td>
<td></td>
</tr>
<tr>
<td>6j4t</td>
<td>3.64±1.55</td>
<td>3.59±1.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-1.02</td>
<td>-1.01</td>
<td></td>
</tr>
<tr>
<td>combined</td>
<td>2.59±1.06</td>
<td>2.52±1.06</td>
<td>2.46±1.03</td>
</tr>
<tr>
<td></td>
<td>-0.74</td>
<td>-0.72</td>
<td></td>
</tr>
</tbody>
</table>

Example in Appendix A.3.2. The final combined limit is improved by approximately 3%. However, there is still potential for further improvement.

Since the Particle Swarm Optimisation which provided the parameters of the training in Tab. 11.5 was performed without considering the angular variables, the training parameters might not be ideal for the usage together with these variables. A new optimisation considering the angular variables would therefore be preferable. However, the PSO has very long run times. A new optimisation in each category was not possible in the scope of this thesis. Thus, a new optimisation was performed only in the signal enriched category with at least six jets and at least 4 b-tags to evaluate its impact. The resulting training parameters are shown in the last row of Tab. 11.5. The input variables chosen by the PSO are given in the last row of Tab. 11.3. They consist of the variables already given by the PSO without considering the angular variables and additional variables chosen by the new optimisation. The corresponding BDT output shape is displayed in the last row of the histograms in Appendix A.3.8. The expected upper limit corresponding to the new optimisation can be found in the column “PSO” in Tab. 11.6.

The new PSO in the category with at least 6 jets and 4 b-tags indicates that there is potential for improvement by newly optimizing the BDTs. The additional optimisation results in an improved upper limit by approximately 5% in the 6j4t category compared to the case in which no angular variables were considered. Another combined limit is calculated using
the results from just adding the variables to the existing training in each category except in the 6j4t category. There, the result obtained due to the new optimization is used. The improvement in the 6j4t category propagates directly to the combined limit resulting in an improvement of approximately 5% compared to the case without angular variables. If the PSO is performed again considering the angular variables in each category, this could further improve the result.

Moreover, the inclusion of a boosted category promises some improvement as well, as was shown in chapter 10. In addition, an improved reconstruction technique could enable the usage of the angular variables in the helicity frame which should also improve the benefit of the angular variables.
12. Conclusion and Outlook

The production of a Higgs boson in association with a top-quark-antiquark pair allows to measure the Yukawa coupling of the top quark to the Higgs boson. With a mass of approximately 173 GeV, the top quark is by far the heaviest elementary fermion known to this point in time and its Yukawa coupling to the Higgs boson is expected to be of $O(1)$. The strength of such a coupling might indicate a special role of the top quark within the Standard Model or for physics beyond the Standard Model. In 2012, ATLAS and CMS discovered a particle compatible to the properties of the Higgs boson in the Standard Model [2–5]. However, the associated production of a Higgs boson and a top-quark-antiquark pair yet has to be discovered. Within the CMS experiment, a search for $t\bar{t}H$ is performed in several decay channels as a combined effort of many groups and institutes. This thesis is situated within the semileptonic $t\bar{t}H(\ell \ell)$ analysis [13].

In this thesis, the possible usage of angular variables, used in spin correlation analysis, within the semileptonic $t\bar{t}H(\ell \ell)$ analysis was investigated. The irreducible background $t\bar{t}b\bar{b}$ needs to be separated from the signal. Therefore, observables had to be found which are able to show physical differences between the two processes. Candidates for such observables are angular variables sensitive to spin correlations of the top-quark-antiquark pair. These angular variables include, for example, the three-dimensional opening angle between the lepton from the leptonically decaying top quark and the bottom-type quark from the hadronically decaying top quark. The additional Higgs boson radiation in $t\bar{t}H$ is expected to change the configuration of the helicity states of the top quarks because of the chiral structure of the Yukawa terms.

The modelling of the angular variables of top-quark-antiquark-pair decay products was validated at generator level and its good compatibility to the expected theoretical behaviour was confirmed. After that, the angular variables were examined regarding their separation power at generator level. Observable and quantitative differences were found due to spin correlations and kinematic effects. These differences were successfully employed in a boosted decision tree showing the combined separation power of the angular variables. In addition, the complementary spin correlation behaviour of $t\bar{t} + \text{jets}/t\bar{t}b\bar{b}$ with respect to $t\bar{t}H(\ell \ell)$ in the boosted regime was highlighted. Another BDT was trained with boosted events showing even better separation power in this phase space region.

In the next step, the variables were investigated at reconstruction level. In the process, several challenges were observed. The association of jets with final-state partons using a simple $\chi^2$ method does not work well for high-jet-multiplicity events like $t\bar{t}b\bar{b}$ or $t\bar{t}H(b\bar{b})$. 
Combined with the especially low fraction of available $t\bar{t}b\bar{b}$ events in the $t\bar{t} + \text{jets}$ sample, the number of correctly reconstructed $t\bar{t}b\bar{b}$ events is very low compared to rest of the $t\bar{t} + \text{jets}$ events. Moreover, it was shown that the incorrectly reconstructed events show a completely different behaviour than the ones reconstructed correctly. However, due to this different reconstruction behaviour, additional separation was observed in the laboratory frame. Furthermore, agreement of the reconstructed distributions in MC simulation and in data was verified. Due to the aforementioned challenges, the evaluation regarding the statistical limit of the $t\bar{t}H(\text{bb})$ analysis was performed only including angular variables defined in the laboratory frame to avoid the influence of incorrect Lorentz boosts into other reference frames. The upper limit on the signal strength modifier showed a positive influence of the angular variables in each category. By adding the angular variables to the existing classifiers of the semileptonic $t\bar{t}H(\text{bb})$ analysis, the combined expected upper limit was improved by approximately 3% from $\mu < 2.59$ to $\mu < 2.52$ at 95% confidence level. After performing a new classifier optimization in the signal enriched category with at least 6 jets and at least 4 b-tags, the combined expected upper limit compared to the case without angular variables was improved by approximately 5% to be $\mu < 2.46$ at 95% confidence level.

There are, however, several possibilities to improve the usage of the angular variables. A better reconstruction of the $t\bar{t}$ system would immensely help in gaining more correctly reconstructed events and therefore increasing the fraction of correctly reconstructed events in the training statistics, which is very low for the $t\bar{t}b\bar{b}$ process until now. This would increase the utilisation of physical differences between the processes and reduce effects due to the reconstruction technique. Moreover, dileptonic events were not used at reconstruction level because of the applied selection criteria in the $t\bar{t}H(\text{bb})$ single-lepton analysis. However, these events have the largest spin analyzing power. Thus, the dileptonic channel is expected to profit even more from these variables. An application in this channel should therefore be considered. Furthermore, the variables should be investigated in the boosted regime at reconstruction level. There, the $t\bar{t}$ reconstruction efficiency should be significantly higher and the provided separation is expected to be larger.

Concluding, the angular variables studied in this thesis show promising and interesting results. Their modelling in MC simulation is found to be accurate and the sensitivity of the $t\bar{t}H(\text{bb})$ single-lepton analysis is improved by including them in the multivariate analysis. Therefore, the angular variables should be used in this analysis. In addition, the improvements mentioned above should be applied. Only then, the full potential of the angular variables can be exploited.
Bibliography


Bibliography


<table>
<thead>
<tr>
<th>Citation</th>
<th>Reference</th>
</tr>
</thead>
</table>


Bibliography


Bibliography


[136] Search for \( \bar{t}tH \) production in the \( H \rightarrow b\bar{b} \) decay channel with \( \sqrt{s} = 13 \text{ TeV} \) pp collisions at the CMS experiment. Tech. rep. CMS-PAS-HIG-16-004. Geneva: CERN, 2016. url: https://cds.cern.ch/record/2139578.


A. Appendix
A. Appendix

A.1. Additional Material for Chapter 10

Figure A.1.: The difference in pseudorapidity between the various decay products of the $t\bar{t}$ system in the helicity frame.
Figure A.2.: The difference in pseudorapidity between the various decay products of the \(tt\) system in the top-quark pair center-of-mass frame.
Figure A.3: The generator (left) and reconstructed (right) distribution of the three-dimensional opening angle (upper row) and the azimuthal difference (bottom row) between the two bottom quarks from the top-quark decay in the lab frame are shown. These distributions were created from the "MG5_aMC@NLO" MC samples, cf. Tab. 5.1.
A.3. Control Plots

This section shows Data/MC control plots of the angular variables considered at reconstruction level in this thesis. Each variable is shown in each category. The last two plots in the categories used in the final fit (all except the 4 jets, 2 b-tags category) labeled with “final discriminator” show the output shapes of the trained boosted decision trees. The BDT output shapes on the left/right are obtained neglecting/considering the angular variables during the training.

A.3.1. 4 jets, 2 b-tags
A. Appendix

A.3.2. 4 jets, 3 b-tags
A.3.3. 4 jets, 4 b-tags
A. Appendix

A.3.4. 5 jets, 3 b-tags
A.3. Control Plots

A.3.5. 5 jets, ≥ 4 b-tags
A. Appendix

A.3.6. $\geq 6$ jets, 2 b-tags
A.3.7. ≥ 6 jets, 3 b-tags

A.3. Control Plots

12.9 fb⁻¹ (13 TeV)

Events vs. \( \Delta \eta_{lb} \) (lab)

Events vs. \( \Delta \phi_{bb} \) (lab)

Events vs. \( \Delta \eta_{bb} \) (lab)

Events vs. \( \cos(\Delta \phi_{\Phi}) \) (lab)

Events vs. final discriminator (ljets_jge6_t3)

Events vs. data/MC for 12.9 fb⁻¹ (13 TeV)
A. Appendix

A.3.8. ≥ 6 jets, ≥ 4 b-tags