Universität Bonn

Physikalisches Institut

Studies of $t\bar{t}H(\tau\tau)$ event topologies and kinematics with the ATLAS experiment

Birgit Stapf

The $t\bar{t}H$ process grants one of the few opportunities to access the Yukawa coupling between the Higgs field and top quarks on tree level. This thesis examines possibilities for more elaborate analysis methods going beyond the cut and count approach in the $t\bar{t}H(\tau\tau)$ channel, with a focus on the semileptonic decay of the top quarks and the tau pair. Experience has shown that in the $H \to \tau\tau$ channel the shape of the reconstructed invariant mass can be a powerful tool for signal extraction. In the channel considered reconstructing the invariant $\tau\tau$ mass becomes especially challenging as an ambiguity due to two charged leptons in the final state – only one of which is a product of the $H \to \tau\tau$ decay- arises. Moreover, the missing momenta of the neutrinos cannot be reconstructed easily. To address these problems, the specific topologies and kinematics of $t\bar{t}H(\tau\tau)$ events are studied. As simple, topological considerations are found to be insufficient for the intended purpose, a likelihood based kinematic fit to fully reconstruct the $t\bar{t}$ system is employed in order to associate the two leptons correctly with their respective decay systems. Finally, a new likelihood, based on the transverse masses of the leptonically decaying top quark and W boson, is defined and its potential is investigated.

Physikalisches Institut der
Universität Bonn
Nussallee 12
D-53115 Bonn

BONN-IB-2015-03
Mai 2015

Universität Bonn

Physikalisches Institut

Studies of $t\bar{t}H(\tau\tau)$ event topologies and kinematics with the ATLAS experiment

Birgit Stapf

Dieser Forschungsbericht wurde als Masterarbeit von der
Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen.

Angenommen am:	27.05.2015
1. Gutachter:	Prof. Dr. Norbert Wermes
2. Gutachter:	Prof. Dr. Jochen Dingfelder

Contents

CHAPTER 1

Introduction

One of the most important achievements of modern physics is the development of the so-called *Standard Model* of particle physics. This quantum field theory enables a precise description as well as categorization of all subatomic particles that are as of yet considered to be *elementary*, i.e. without further substructure, and their interactions. To put it into simpler terms, the Standard Model represents the current knowledge on how matter is made up out of indivisible building blocks.

Since its development in the 1960s the Standard Model has been very successful in accurately describing numerous phenomena as well as predicting the discovery of new particles. Nonetheless it is considered incomplete as there are several observations that it does not account for. For example, it is not able to explain the existence of dark matter and energy as experimentally indicated by cosmology. Moreover, it includes no theory of gravity, which is of little relevance on subatomic level but of utmost importance for macroscopic matter. Bringing all of these insights into accordance with each other requires extensions of the Standard Model. Out of such extensions the most well known and promising is the theory of *supersymmetry*, which postulates the existence of a *superpartner* to each currently known particle.

Much progress has been made in recent years concerning a previously experimentally unconfirmed part of the Standard Model, namely the so-called *Brout–Englert–Higgs mechanism* according to which all elementary particles obtain their finite masses by interacting with a scalar field commonly referred to as the Higgs field. A quantum excitation of this field, the Higgs boson, was already predicted in the 1960s [1]. Yet its presumptive discovery was only made in 2012, when a previously unknown boson was observed in proton-proton collisions at the Large Hadron Collider and recorded by both the ATLAS [2] and CMS [3] experiment. An explanation for the long delay between theoretical prediction and experimental confirmation can be found in the fact that massive, exotic particles such as the Higgs boson can only be produced at very high energy densities and the technology to achieve this has only been available in recent years. Another factor is that the mass of the Higgs boson could not be predicted by theory. However, it is now found to be about 125 GeV [4].

Knowing the mass of the Higgs boson, very precise predictions of all its other properties can be made. This includes the coupling strengths between the Higgs field and other Standard Model particles. Experimentally the couplings can be directly determined from measuring the branching fractions of the different decay modes, i.e. the probability for the Higgs boson to decay into the particle type in question. Additionally a measurement of the rate, or cross section, at which the Higgs boson is produced in association with another type of particle

gives direct access to the coupling between them. To fully confirm the Standard Model nature of the discovered Higgs boson measuring all of the possible couplings as precisely as possible is now necessary.

A special case is the coupling between the Higgs field and the heaviest known elementary particle, the top quark. As the mass of the Higgs boson of 125 GeV is smaller than that of top quark pair, which have a mass of 173 GeV each, no decay of the Higgs to top quark pairs is possible. So the only direct way to measure this coupling on tree-level is the associated production of Higgs and either a single top quark or a pair of differently charged top quarks, the latter of which is the $t\bar{t}H$ process. As is common practice the analysis of $t\bar{t}H$ events is further divided into different channels, according to the different modes for the Higgs boson to decay. All of these individual analyses can be optimized for the given situation and are then ideally combined into one measurement, here of the production cross section of $t\bar{t}H$ reactions. Currently none of the $t\bar{t}H$ analyses have found a statistically significant signal, but to draw a more definite conclusion larger datasets are necessary. Fortunately, with the increase of the center of mass energy in the second run of the LHC - starting at the time of this writing - $t\bar{t}H$ processes become much more likely. Additionally this channel has the unique advantage that the cross section of the most dominant background process, namely the production of top quark pairs, does not increase in accordance, resulting in an a-priori improvement of the signal to background ratio. Hence a strong incentive for investigating possible ways in which the current $t\bar{t}H$ analyses can be improved is given.

Especially the approach employed as of yet in the analysis of $t\bar{t}H(\tau\tau)$ events is rather simplistic, as such processes are relatively rare and the available data of the first LHC run do not support more complex methods. A signal region is defined by certain selection requirements on different variables and the number of events found within this region is counted. So far no detailed information on the event topologies and kinematics expected of the signal is exploited. Yet, from the analyses performed in the $H \to \tau\tau$ channel, reconstructing for example the invariant $\tau\tau$-mass can be expected to provide separation between signal and background events. Thus, for Run 2, it is fitting to investigate whether and how this approach can be applied to $t\bar{t}H(\tau\tau)$ events. A first look into this, based on Run 1 Monte Carlo events at $\sqrt{s} = 8$ TeV, is the idea behind this thesis.

More details on the necessary theoretical and experimental basics are given in Chapters 2 and 3. As $t\bar{t}H(\tau\tau)$ processes are very complex, a particular decay channel leading to a final state containing 2 leptons and 1 τ-jet is chosen for the studies presented. Its properties as well as the corresponding selection of events are described in Chapter 4. This chapter includes an outline of the problem encountered when reconstructing the $\tau\tau$-mass in this channel - namely, the combinatorial background due to the selection of two leptons. Simple approaches to address this problem are investigated and discussed in Chapter 5. As an interlude studies of the truth level topologies of signal events are presented in Chapter 6 to gain insight into the apparent shortcomings of the simple approaches. Possible ways of applying more complex methods, which reconstruct the full event kinematics, are studied and tested in Chapter 7. While the first of these approaches is based on an established tool called KLFitter and investigated more extensively, a new idea to simplify the situation is implemented as a short outlook in the second part of the chapter. Chapter 8 summarizes the findings and gives an outlook for further studies.

CHAPTER 2

Theory

2.1 The Standard Model

The Standard Model of particle physics is a quantum field theory with which all as of yet known elementary particles and their interactions can be described and categorized. According to current understanding, there are two types of particles distinguished by their intrinsic angular momentum and quantum number of spin. Firstly, all matter is made up out of so called *fermions* carrying half-integer spin. Interactions, on the other hand, are mediated by particles carrying integer spin, which are called *bosons*. An overview of the particles included in the Standard Model as well as their masses and electromagnetic charges is shown in Fig. 2.1.

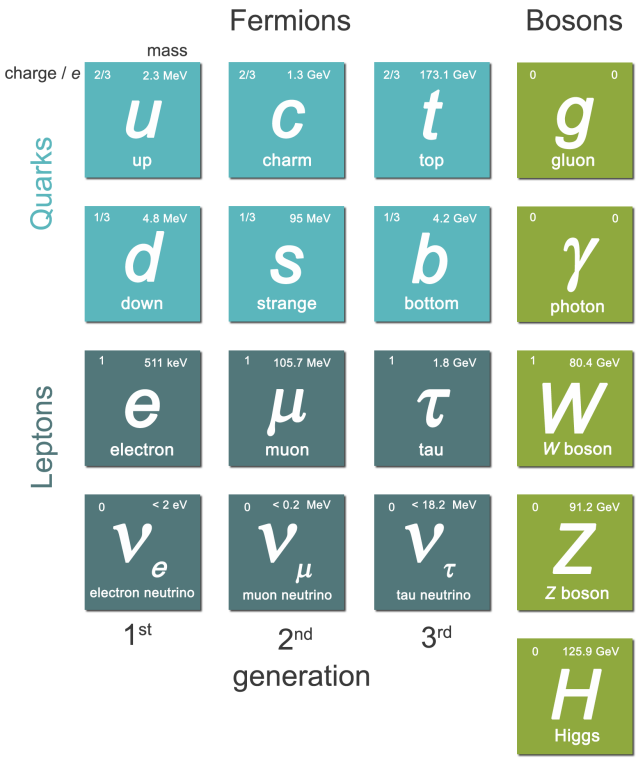


Figure 2.1: Overview of Standard Model particles (values from [5]).

As can be seen fermions occur in two different types, namely *quarks* and *leptons*, and each

of these types also contains six different *flavours* of particles. Within each group two states of electromagnetic charge are possible: the three charged leptons carry one negative elementary charge e each and the electrically neutral leptons are called neutrinos. Quarks of the *up-type* are positively charged with $\frac{2}{3}e$ while *down-type* quarks carry $-\frac{1}{3}e$. Two fermions of the same type but different charge are paired up into a so-called *generation* and with increasing number of generation the mass of the particles in it increases as well. As only the lightest fermions of the first generation are stable, ordinary matter - such as atoms - is exclusively made up out of them, while the higher generation fermions can form what is called exotic matter but only at a finite lifetime before decaying into a stable state. Since for each fermion an anti-fermion with exactly opposite quantum numbers exist, there are in total 12 different leptons. Additionally, quarks occur in three different states of *colour charge* - red, green or blue - resulting in a total number of 36 quarks [6].

At the subatomic level three fundamental forces act between the particles of the Standard Model which carry the respective charge. The *electromagnetic interaction* is imparted by the exchange of a photon which couples to the electromagnetic charge. Therefore all quarks as well as the three charged leptons are subject to electromagnetic interaction. Since the photon itself is massless, the range of the electromagnetic force is infinite. Even though its mediator the gluon is also massless, the *strong interaction* is only effective at small distances in the order of femtometres. The respective charge is the colour charge, which is not only carried by quarks but also by gluons themselves. The latter results in an anti-screening of a free colour charge, so that the effective charge is only decreased with decreasing distance. In this way, it is not possible for free quarks to exist, instead they are *confined* due to their colour charge. According to the theory of the strong interaction, quantum chromodynamics, gluons exist in eight different states of colour. The weak interaction is mediated by the electrically neutral Z as well as the charged W boson and acts upon all fermions of the Standard Model, even the neutrinos which do not participate in any of the other interactions. Again this force is short-ranged due to the large mass of its exchange particles and therefore only relevant at microscopic distances.

From a theory point of view, all three fundamental interactions are derived from the requirement of local gauge invariance within the Standard Model. The corresponding gauge group is

$$SU(3) \otimes SU(2) \otimes U(1) \tag{2.1}$$

where $SU(3)$ corresponds to the symmetry group of the strong interaction, and the electromagnetic and weak interaction are unified into the electroweak interaction described by $SU(2) \otimes U(1)$. The interacting fermions are then represented by left handed Weyl fields of different representations [7]. However when formulating the theory in this way, the gauge bosons mediating the electroweak interaction are required to be massless. This is obviously not the case for the W and Z boson, which are experimentally found to exhibit masses of $m_W = 80.4\,\mathrm{GeV}$ and $m_Z = 91.2\,\mathrm{GeV}$ [5]. This incompatibility can be solved introducing another field, namely the Higgs field, into the theory. Its quantum excitation, the Higgs boson, is then the last boson contained in the Standard Model.

2.1.1 The Higgs boson

In particular the Higgs field takes the shape of a doublet containing the complex scalar fields Φ^+ and Φ^0

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \Phi_1 + i\Phi_2 \\ \Phi_3 + i\Phi_4 \end{pmatrix} \tag{2.2}$$

where Φ_i describe real fields. This doublet Φ is invariant under $SU(2)$ and its invariant product is given by $\Phi^\dagger\Phi = \frac{1}{2}(\Phi_1^2 + \Phi_2^2 + \Phi_3^2 + \Phi_4^2)$. The corresponding scalar potential is

$$V(\Phi) = \mu^2|\Phi^\dagger\Phi| + \lambda|\Phi^\dagger\Phi|^2 \tag{2.3}$$

where $\lambda > 0$ is given to bound the potential from below [8]. The shape of this potential depends on the sign of the parameter μ^2: if μ^2 is positive, it has exactly one minimum found at $< 0|\Phi|0 >= \Phi_0 = 0$. In this case the gauge symmetry is conserved and the interaction of a massless vector boson and a massive charged scalar particle is described [9]. On the other hand, if $\mu^2 < 0$, an infinite number of energy-degenerated minima is found. They satisfy the equation

$$\Phi_0 = |\Phi^\dagger\Phi| = \frac{-\mu^2}{2\lambda} \tag{2.4}$$

which can be solved by setting $\Phi_1 = \Phi_2 = \Phi_4 = 0$ and $\Phi_3 = \frac{-\mu^2}{2\lambda} \equiv v^2$. A sketch of the two cases is shown in Fig. 2.2.

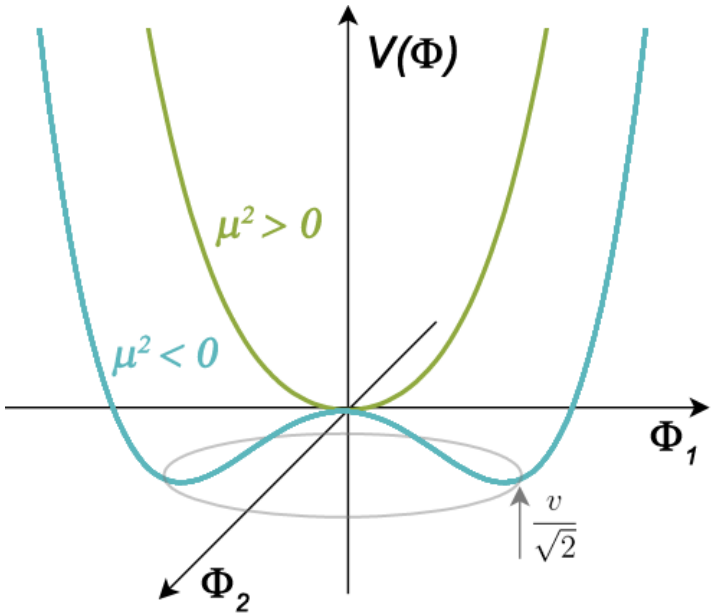


Figure 2.2: Sketch of the Higgs-potential $V(\Phi)$ for different signs of the parameter μ^2.

In the case of $\mu^2 < 0$ the vacuum expectation value - or ground state - of Φ is not found at 0; rather the point $\Phi = 0$ is unstable. Instead out of the possible minima one particular ground state must be chosen and the potential can then be expanded in the vicinity of

Decay channel	Branching ratio	Rel. uncertainty
$H \to \gamma\gamma$	2.28×10^{-3}	$^{+5.0\%}_{-4.9\%}$
$H \to ZZ$	2.64×10^{-2}	$^{+4.3\%}_{-4.1\%}$
$H \to W^+W^-$	2.15×10^{-1}	$^{+4.3\%}_{-4.2\%}$
$H \to \tau^+\tau^-$	6.32×10^{-2}	$^{+5.7\%}_{-5.7\%}$
$H \to b\bar{b}$	5.77×10^{-1}	$^{+3.2\%}_{-3.3\%}$
$H \to Z\gamma$	1.54×10^{-3}	$^{+9.0\%}_{-8.9\%}$
$H \to \mu^+\mu^-$	2.19×10^{-4}	$^{+6.0\%}_{-5.9\%}$

Table 2.1: Branching fractions of the predicted Higgs boson decays [5].

it. Since in doing so the original gauge symmetry is not conserved, this is called *spontaneous symmetry breaking*. Following the procedure through, explicit mass terms for the interacting particles arise [7]. In this way the electroweak symmetry breaking is responsible for elementary particles obtaining their finite mass. The mechanism was first postulated in the 1960s ([1, 10, 11]) but the experimental evidence for a previously unknown boson, which is consistent with the predicted Higgs boson so far, was only found in 2012 ([2], [3]). Recently measurements of its performed by both ATLAS and CMS were combined, yielding a rather precise value of $m_H = 125.09 \pm 0.21 \pm 0.11\,\text{GeV}$, where the first error refers to the statistical and the second one to the systematic uncertainty. Whether or not this boson fully confirms all expectations of a Standard Model Higgs boson is still to be validated.

One important factor in this is the analysis of its couplings to other Standard Model particles. To probe the Higgs couplings in principle different possibilities exist. For one thing they are reflected in the branching fractions of the different decay modes, an overview of their predicted values for a Standard Model Higgs boson is given in Table 2.1. Currently significant evidence for a Higgs boson coupling to other bosons, for example in decays of the form $H \to \gamma\gamma$ and $H \to ZZ$ used in [4], has been found. However there are observations of it also coupling to fermions, for example in the form of $H \to \tau\tau$ decays [12].

Moreover the cross sections of the different Higgs production mechanisms in pp-collisions for a given center of mass energy are predicted by theory, as depicted in Fig. 2.3. While the most likely case corresponds to the production of only the Higgs boson, for example and most prominently out of the fusion of two gluons, there are multiple ways in which the Higgs boson is produced in association with other Standard Model particles. Out of these one special case is the production of Higgs boson and a top quark pair, $t\bar{t}H$, which this thesis focuses on. A review of the current measurements of the Higgs couplings, combining the results achieved in different channels so far, can be found in [13].

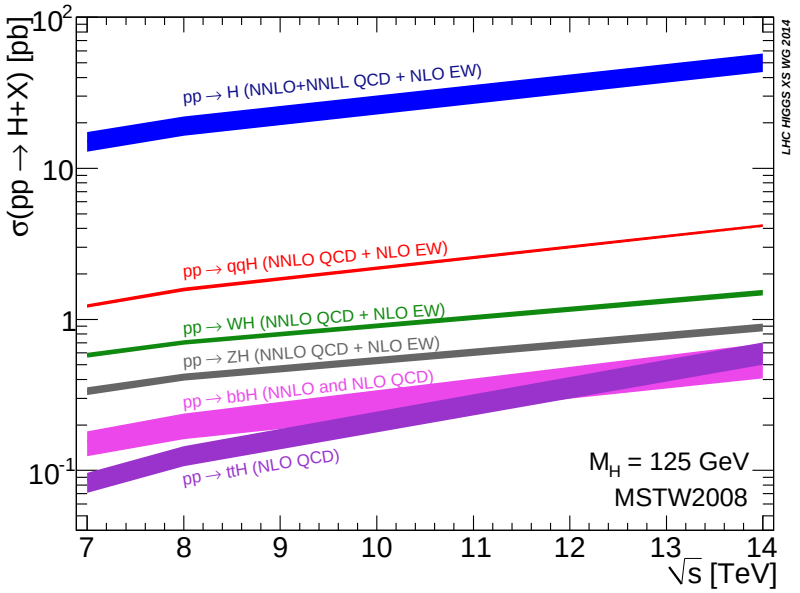


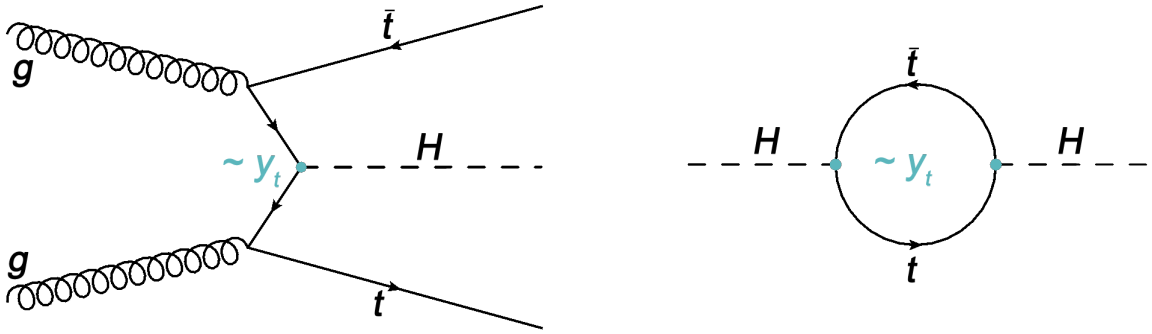
Figure 2.3: Higgs production cross sections in proton-proton collisions at different values of the center of mass energy $\sqrt{s}$.

2.2 Associated production of the Higgs boson with a top quark pair

Even though the associated production of a Higgs boson with a top quark pair is a very rare occurrence, it is of interest for precision measurements of Higgs properties. This is due to the fact that for a Higgs boson with $m_H = 125\,\text{GeV}$ a decay into a real top quark pair is energetically forbidden. Therefore the coupling between the Higgs field and top quarks, referred to as the *Yukawa coupling* y_t, can only be accessed directly and on tree-level in the associated production of Higgs and top quarks. While events in which a single top quark is produced along the Higgs boson can also be studied (see e.g. [14]), this thesis focusses on the production along with a top-antitop quark pair, as depicted in Fig. 2.4.(a). The point of interaction is commonly referred to as a *vertex*. Additionally the coupling y_t contributes indirectly to production and decay rates of different different Standard Model processes in the form of loop corrections, sketched in Fig. 2.4.(b) [15].

Hence one reason to investigate the potential of observing $t\bar{t}H$ processes is the possible manifestation of new, beyond Standard Model physics: whether the measured value of y_t is compatible with Standard Model expectations, is one step further in establishing if indeed a Standard Model Higgs boson has been discovered. The values of y_t obtained from tree- and loop-level measurements may not agree provided that new physics processes are contributing on either of the levels or even on both, but in different ways.

Another reason to study $t\bar{t}H$ events is that a precise knowledge of the coupling between even a Standard Model Higgs and top quark is of special importance as it may help in understanding more about the mass giving mechanism of electroweak symmetry breaking. In fact, it may give insight into why the top quark is much heavier than all other fermions, as according to theory the mass of a particle should be proportional to its coupling with the Higgs field [16].



(a) (b)

Figure 2.4: Processes involving the coupling between Higgs boson and top quarks, y_t : (a) on tree level in the form of $t\bar{t}H$ production and (b) in the form of loop corrections.

Despite its branching fraction being significantly lower the channel in which the Higgs boson subsequently decays into a pair of τ leptons can add substantial information to complement analyses in the more frequent $t\bar{t}H(b\bar{b})$ channel. In particular, as presented in [17], it provides good opportunity for a model independent determination of Higgs couplings to the particles of the Standard Model as well as its width. An important feature lies in the fact that in this channel the Yukawa coupling between Higgs boson and τ lepton, y_τ, is also directly accessible. Combining this with precision measurements in the $t\bar{t}H(b\bar{b})$ channel, the ratio of the Higgs couplings to bottom quark and τ-lepton, $\frac{y_b}{y_\tau}$, can be determined without relying on any predictions made by the Standard Model or other theoretical models going beyond it. Indications for the latter may then be found if the value of ratio $\frac{y_b}{y_\tau}$ deviates from the expected one, e.g. due to non Standard Model loop corrections. The advantage of considering a ratio of couplings is that in this case the dependency on theory uncertainties is reduced compared to calculating individual couplings. Thus there is ample motivation for facilitating the observance of $t\bar{t}H(\tau\tau)$ events.

CHAPTER 3

Experiment

3.1 The Large Hadron Collider

The **L**arge **H**adron **C**ollider is a synchrotron-type particle accelerator located underground at the site of the European Organization for Nuclear Research (CERN) in Geneva, Switzerland. It consists of two rings of 26.7 km circumference each, which facilitates the acceleration of two beams consisting of particles of the same type in opposite directions. There is four different points where the beams can be crossed and the resulting particle collisions recorded by the respective detectors. A schematic view of this is shown in Fig. 3.1, where the collision points are indicated by stars.

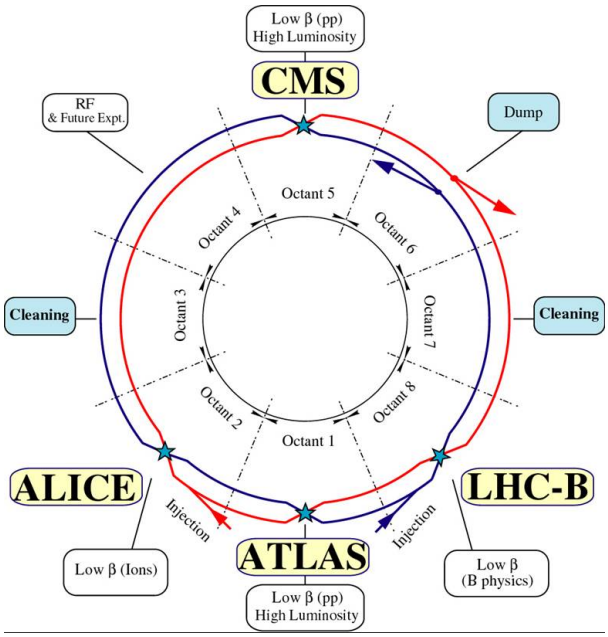


Figure 3.1: Schematic view of the LHC layout [18]. The direction of acceleration for each hadron beam is indicated by the arrows.

More specifically the tunnel which houses the LHC consists of eight straight sections alternated with eight arcs, dividing it into octants. Particles are accelerated by RF cavities along part of the straight sections, and deflected onto a circular path by superconducting dipole-magnets in the arcs. Additionally an elaborately designed structure of superconducting quadrupole-

magnets ensures sufficient cooling, focussing and stabilisation of the beam.

The four collision points and experiments located in four of the straight sections have different objectives. While CMS and ATLAS are general purpose detectors, LHCb focusses on physics processes involving bottom quarks. These experiments focus on analysing proton-proton collisions. The ALICE experiment is supplied with collisions of heavy ion beams in order to investigate the behaviour of hadronic matter at high energies and densities.

Currently the LHC is the world's most powerful accelerator, colliding particles at unsurpassed center of mass energies. While it ran succesfully in 2012 at $\sqrt{s}$ = 8 TeV, it is now upgraded for the second run with collisions at $\sqrt{s}$ = 13 (and eventually 14) TeV. Reaching such high energies is made possible by the choice of beams consisting of hadrons, rather than much lighter leptons. The latter suffer from great energy losses due to *synchrotron radiation*, the scale of which is proportional to the mass of the accelerated particles. However, the downside of using non-elementary particles such as protons is that it is unknown which of the partons contained in the colliding protons are involved in the inelastic scattering reaction. The longitudinal momentum components of the interacting partons are inaccessible but their transverse momenta are negligible before the collision. Therefore, most quantities used in the analyses are reconstructed in the transverse plane.

The beams at the LHC consist of bunches composed of $1.15 \cdot 10^{11}$ protons each. In this way during the crossing of such bunches multiple collisions can happen simultaneously, which increases the rate of possible scattering reactions. However within one recorded event several unrelated processes can take place. This effect is called *pile-up* [19].

The LHC is also designed for maintaining high reaction rates, quantifiable by the *luminosity* $\mathcal{L}$. Its design instantaneous luminosity of $\mathcal{L} = 1 \times 10^{34}\,\mathrm{cm}^{-1}\,\mathrm{s}^{-2}$ is due to be reached during the Run 2 operation from 2015 - 2017, while during the course of 2012 a data sample corresponding to an integrated luminosity of $20.3\,\mathrm{fb}^{-1}$ was recorded at 8 TeV [20]. As the spacing between proton bunches is decreased and the beam optics are improved for Run 2, a larger data set containing more high quality events can be obtained in the future.

3.2 The ATLAS detector

One of the main experiments at the LHC is the multi-purpose detector ATLAS. It is comprised of multiple components which are layered concentrically around the collision point. In this way the cylindrical detector covers almost the full solid angle, with the exception of the direct beam direction. An overview of its structure is depicted in Fig. 3.2. It can be divided into three main systems, dependent on their function. The *inner detector* serves as a basis for precise measurement of the tracks of charged particles and is succeeded by the *calorimeter system*. Finally the outermost layer is comprised by the *muon spectrometer*. As different particles cause different reactions within each of these three parts, particles can be identified by the individual signature they leave when traversing the detector. A brief summary of the detector sub-systems and their functionalities is given in the following, more details can be found in [21].

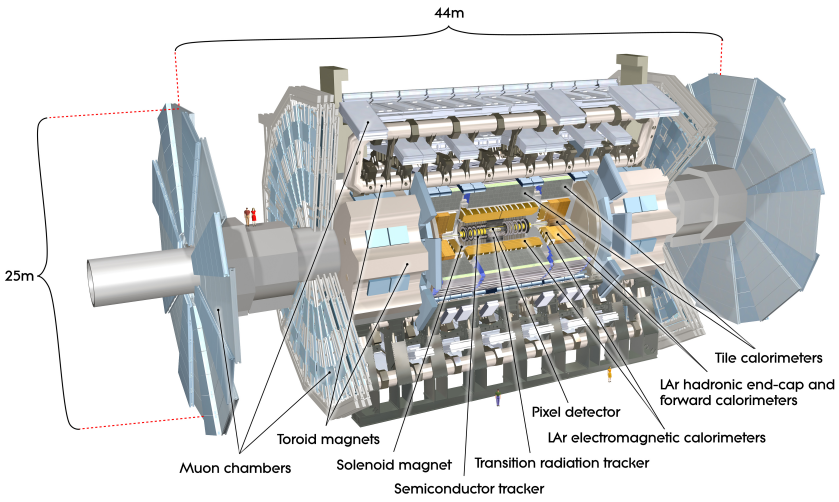


Figure 3.2: Overview of the ATLAS detector and its components [21].

Inner detector The inner detector is comprised of three layers all specially designed for high-precision measurements of particle tracks. They are based on different techniques but the same general principle: when a charged particle traverses the detector material, the latter is ionized and the track of the particle can be reconstructed from the free charges created in the process. Moreover the whole system is enclosed within a solenoid creating a nominal magnetic field of 2 T. Due to this field, charged particles are forced onto a bent trajectory, the curvature of which is proportional to the particle's momentum. Additionally particle tracks are bent in different directions depending on the sign of their electromagnetic charge, allowing for charge separation.

Closest to the interaction point and providing the highest resolution of track measurements, the *Silicon Pixel Detector* is found. This component in turn consists of three different layers, the innermost of which is called the *B-layer*. Due to its very close proximity to the beam pipe at a distance of $\geq$ 4.55 cm, especially this component is affected strongly by radiation damage. Therefore a fourth layer, called the *Insertable B-Layer* (IBL), was added between the beam pipe and the original B-layer in preparation for Run 2. In this way continued high performance of the pixel detector - even at higher track densities due to the increase in luminosity - is ensured [22].

The pixel detector is followed by the *Silicon Microstrip Tracker* (SCT). As the name indicates this component consists of longs strips of silicon detectors rather than pixels spread over a total of eight layers.

Finally, the outermost part of the inner detector is the *Transition Radiation Tracker* (TRT). It consists of gas-filled straw tubes. Transition radiation is emitted during the passage of charged particles though the TRT. As the intensity of the radiation is correlated to the ratio of kinetic energy and rest mass of a particle, with the help of this component heavy particles can be distinguished from light particles - i.e. pions from electrons - at a given momentum.

Calorimeter system Calorimeters are used to measure the energy of incoming particles of high energy. They are comprised of high-density material into which the particles induce cascades of secondary particles, called *showers*. In order to measure the full energy deposited by the original particle, enough material must be present to stop all of the particles created in the shower. As more material is required to stop hadronic showers, the ATLAS calorimeter system is divided into the electromagnetic and the hadronic calorimeter. While the former is enough to absorb the energy of electromagnetically interacting particles such as electrons and photons, the latter is necessary to also stop strongly interacting particles which create hadron showers. Muons on the other hand do not deposit much of their energy into either of the calorimeters due to their large mass. Both calorimeters are of the sampling variety - layers of active material for signal readout are alternated with layers of passive material which serve for shower induction. The ATLAS calorimeters use a combination of liquid argon (LAr) and lead/stainless steel.

Muon spectrometer Since muons are typically not stopped by the calorimeter system but are able to traverse the rest of the detector, the muon spectrometer which serves for their detection is the outermost component of ATLAS. It mainly uses monitored drift tubes but also other techniques to ensure high-precision muon tracking. Additionally a magnetic field for muon momentum measurement is provided by toroid magnets.

Data Acquisition Another important part of the detector is its *Trigger and Data Acquisition* system. Running at its nominal values the event rate at the LHC is 40 MHz, corresponding to 1 billion collisions per second. Obviously this is too much to store and process, so to decrease it to more manageable quantities so-called *triggers* filter the raw data for interesting physics events [23].

Coordinate system Due to the specific detector geometry a commonly used coordinate system is defined, which is also used throughout this thesis. The origin of the coordinate system is found in the center of the detector, at the nominal collision point. It is a right-handed coordinate system, with the z-axis defined along the beam direction. The positive z-direction points towards the LHCb experiment, so counter-clockwise along the ring. The xy-plane is then transverse to the beam direction, with the x coordinate pointing to the center of the ring and y pointing upwards. Within the xy-plane the azimuthal angle ϕ is used, running from $-\pi$ to π. A ϕ value of 0 corresponds to the (positive) direction of the x-axis. To describe the distance from the beam and therefore the z-axis the polar angle θ is used. Its range is from 0 to π and a value of $\theta = 0$ gives the positive z-direction. However θ is commonly replaced by the lorentz-invariant rapidity $y = \frac{1}{2} \ln\left(\frac{E+p_z}{E-p_z}\right)$. In the ultra-relativistic limit of massless particles the pseudorapidity $\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$ can be used instead. Within this coordinate system distances between particle directions can then be quantified by the angular separation $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

3.3 Particle reconstruction

Particles can be identified and their properties reconstructed from the raw detector data. As different particles leave individual signatures in the detector, different reconstruction algorithms are used. The only particles that cannot be detected directly are neutrinos, as they only interact weakly and therefore cause no reaction in the detector materials. However, a variable for their indirect detection can be reconstructed.

The reconstruction procedures of the particles considered in this thesis are briefly summarized in the following.

Jets Due to the nature of the strong interaction, no free quarks or gluons exist but instead each quark or gluon, also referred to as parton, involved in a scattering reaction inevitably produces a collimated beam of further particles. This process is called *hadronization* and the resulting object, which is detected, is called a *jet*. Jets can also be initiated by a gluon which is radiated off a quark. The only quark that decays before it can hadronize is the top quark. In the samples used for this thesis jets are reconstructed using the *anti-k_t algorithm* [24]. Its approach is based on a pairwise clustering of objects, which are either single particles or pseudojets already constructed out of multiple particles. For a pair of objects i and j the distance d_{ij} between them as well as the distance of the object i to the beam axis d_{iB} is calculated according to the following definitions:

$$\begin{aligned} d_{ij} &= \min(k_{T,i}^{-2} k_{T,j}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \\ d_{iB} &= k_{T,i}^{-2} \end{aligned} \tag{3.1}$$

where $k_{T,i}$ refers to the transverse energy of object i and R is a constant resolution parameter, quantifying the distance at which distance from each other jets are resolved. In this way the pair of objects which is closest to each other is found and its distance d_{ij} compared to d_{iB}. If d_{ij} is smallest, the objects i and j are combined into a pseudojet and if d_{iB} is smallest, the jet is considered complete and removed from the list of objects. The procedure is repeated until no more objects are left. For the studies presented in thesis $R = 0.4$ is chosen.

b-tagging Jets that are initiated by bottom quarks exhibit unique characteristics that make their special identification, referred to as *b-tagging*, possible. The B-hadrons which are found within a b-jet have a large mass as well as a relatively long lifetime, meaning they can travel a small distance (in the order of a fraction of a mm) before the decay. In this way displaced, secondary vertices can be reconstructed from the tracks associated to the jet in question. Mutivariate techniques are used to exploit this, more details can be found in [25]. The efficiency of the b-tagging mechanism can be chosen by its *working point*. Here a working point corresponding to an efficiency of 70% of correctly b-tagged jets is chosen, which has the advantage of low mistagging rate of differently flavoured jets around 1%.

Electrons As electrons are charged particles, they leave a track in the inner detector before depositing their energy in the electromagnetic calorimeter. Therefore electron identification is

performed by reconstructing energy clusters in the calorimeter and matching them to nearby particle tracks reconstructed in the inner detector. If an energy cluster can be successfully matched to at least one track, this object is considered as an electron candidate. Not all of these correspond to primary electrons, as for example hadrons could leave a similar signature. To minimize misidentification, multiple discriminant variables, taking into account shower shapes and track information, are used as identification criteria. Due to the detector acceptance, $|\eta| < 2.47$ is required for electron candidates and the transition region between calorimeters of $1.37 < |\eta| < 1.52$ has to be excluded. Candidates with a transverse momentum of $p_{\mathrm{T}} < 10$ GeV are disregarded. Lastly candidates are required to be isolated from other objects, in order to exclude cases of non-prompt electrons, which are produced and contained within jets. More details on the procedure, including the specifics of the isolation variables, as well as the resulting identification efficiencies can be found in [26].

Muons Muons are not stopped in the calorimeters but leave tracks in both the inner detector as well as the muon spectrometer. Similarly to electrons, they are then reconstructed by matching track segments found in the muon system to those found in the tracking components. The type of muon reconstruction algorithm employed for the events studied in this thesis is described in [27]. Again only isolated muon candidates with $|\eta| < 2.5$ and $p_{\mathrm{T}} > 10$ GeV are considered.

τ-jets Since they have a very short lifetime in the order of 10^{-13} s corresponding to a decay length of only a few µm τ leptons typically decay before they reach the detector. Therefore they can only be measured indirectly in form of their decay products. Being the heaviest leptons τ leptons can decay to either of the two lighter charged leptons as well as into light hadrons. Each of these decay modes involves at least one neutrino, which is invisible to the detector. Tau decays are therefore reconstructed from the *visible tau* - which corresponds either to a charged lepton or a jet of hadrons. A specialized τ-jet reconstruction algorithm is applied to distinguish the latter case from quark or gluon initiated jets. It is described in detail in [28]. In this approach jets reconstructed as described above and fulfilling the $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$ criteria are considered as potential candidates of a visible, hadronically decaying tau, referred to as τ_{had} in this thesis. To distinguish whether a candidate most likely corresponds to a τ decay or a jet initiated by a quark/gluon, multiple discriminating variables are combined into a *Boosted Decision Tree*. Again, three different working points - *loose*, *medium* and *tight* - can be chosen from, which each exhibit different reconstruction efficiencies as well as misidentification rates.

Neutrinos and $E_{\mathrm{T}}^{\mathrm{miss}}$ As neutrinos do not interact with the detector material, information on their energy and momentum can only be approximated by an indirect measurement of the *missing transverse energy*, $E_{\mathrm{T}}^{\mathrm{miss}}$. Since the protons are only accelerated along the longitudinal axis, i.e. the beam direction, it can be assumed that the scattering reaction occurs at balance in the transverse plane. According to momentum conservation the sum of transverse momenta of all final state particles must therefore vanish. The vectorial momentum sum of the neutrinos can then be approximated as the negative of the vectorial momentum sum of all visible final

state objects i:

$$E_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{i} \vec{p}_{\mathrm{T}}^{(i)} \tag{3.2}$$

Since this quantity is reconstructed using energy deposits in the calorimeter cells, it is referred to as an energy here, even though technically it is based on transverse momentum balance. Approximating neutrino momenta from $E_{\mathrm{T}}^{\mathrm{miss}}$ can only be correct, if indeed the momenta of all visible final state particles are measured with sufficient precision. The matter is further complicated by the vectorial nature of $E_{\mathrm{T}}^{\mathrm{miss}}$- if multiple neutrinos are produced in non-parallel directions, their momenta compensate each other at least partially. Therefore it can also be useful to calculate the scalar sum of of all transverse cell energies, $\sum E_{\mathrm{T}}$. Since measurement uncertainties scale with the number of calorimeter cells into which energy is deposited, $\sum E_{\mathrm{T}}$ is also a quantifier of the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution. More technical details on the reconstruction procedure of these two quantities are provided in [29].

Overlap removal To further minimize the possibility of object misidentification, a procedure called overlap removal is applied. If two fully identified objects are found in overlapping cones of ΔR one of them is removed. The order and ΔR cut for each combination of objects is given in Tab. 3.1.

Static object	Removed object	ΔR cut
Muon	Electron	0.1
Electron	Jet	0.3
Jet	Muon	$0.04 + 10\,\mathrm{GeV}/\ p_{\mathrm{T}}$
Electron	Tau	0.2
Muon	Tau	0.2
Tau	Jet	0.3

Table 3.1: Cuts applied for object overlap removal.

CHAPTER 4

Analysis strategy

In this chapter the tools and concepts used for the following studies are described. Before turning to that however, a brief summary of the results achieved by the current ATLAS analysis of $t\bar{t}H(\tau\tau)$ processes shall be given, to outline the starting point of this thesis.

4.1 Current $t\bar{t}H(\tau\tau)$ analysis in ATLAS

A search for $t\bar{t}H$ processes leading to a multi-lepton final state, including $t\bar{t}H(\tau\tau)$, has been performed on data taken with $\sqrt{s}$ = 8 TeV at the LHC in 2012. It is presented in detail in [30]. Due to the very low event rate of $t\bar{t}H$ processes a simplistic cut and count approach is employed. Different categories, or also called channels, of expected signal regions are defined based on the number of leptons and τ-jets found in the final state. For each of these categories a set of cuts is optimized to enhance the signal to background ratio. By counting the number of observed data events and comparing to the expected number of background events a potential signal is determined in each category. Its strength μ is given by the ratio of observed $t\bar{t}H$ events over the Standard Model expectation $\mu = \sigma_{t\bar{t}H,\text{obs}}/\sigma_{t\bar{t}H,\text{SM}}$. The best-fit value of μ is found as the result of a maximum likelihood fit to the event yields in each category. The case of $\mu = 1$ corresponds to the Standard Model hypothesis, meaning that the observed $t\bar{t}H$ production cross section is found to take its exact theoretical value. The resulting values of μ in each category are shown in Fig. 4.1. As can be seen, they are compatible with the Standard Model expectation of $\mu = 1$ within the bounds of their uncertainties. Due to the low event count, the latter are predominantly of statistical nature for almost all channels.

Combining these results with what the analyses of $t\bar{t}H(b\bar{b})$ [31] and $t\bar{t}H(\gamma\gamma)$ processes have found, evidence for $t\bar{t}H$ production in pp-collisions is indicated [13] but as of yet not significant. While all three analyses are statistically limited, especially $t\bar{t}H(\tau\tau)$ processes are very rare. For example, in the Run 1 data used for the results above only one event was observed in the $2\,\ell, 1\,\tau_{\text{had}}$ channel. So for this purpose the simple methods employed were sufficient. Nonetheless with the start of the second run of the LHC at an increased collision energy of $\sqrt{s}$ = 13 TeV the incentive to perform more complex studies of $t\bar{t}H(\tau\tau)$ events becomes very strong. Hence the starting point of this thesis is to examine possibilities for more elaborate methods going beyond the current analysis, where the focus is on reconstructing the invariant mass of the τ-pair. Since the signal topology is very complex, all further studies are focussed on the $t\bar{t}H(\tau\tau)$ $2\,\ell, 1\,\tau_{\text{had}}$ channel only, which is described in more detail in the following.

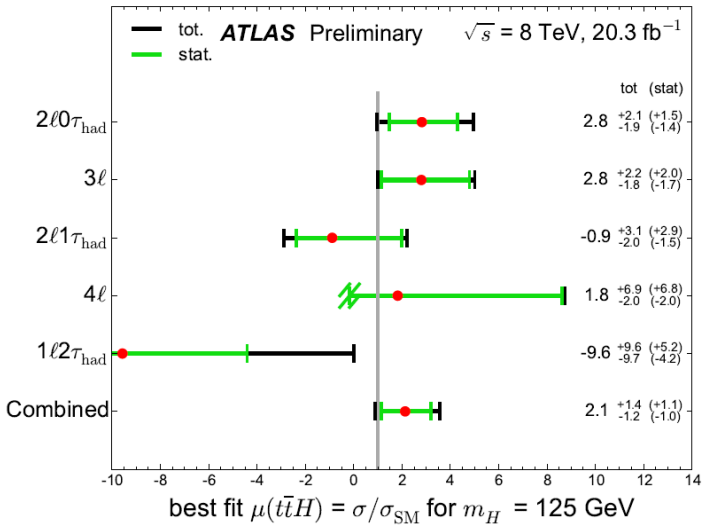


Figure 4.1: Results of the ATLAS analysis of $t\bar{t}H$ processes leading to a multi-lepton final state: Best-fit values of the signal strength parameter μ for each category [30].

4.2 The $t\bar{t}H(\tau\tau)$ $2\,\ell, 1\,\tau_{\mathrm{had}}$ channel

Considering the different possibilities for the top quarks and the τ leptons to decay, there are nine different channels of $t\bar{t}H(\tau\tau)$ events. With a probability of nearly 100% a top quark decays into a bottom quark and a W boson. This W boson can then subsequently decay in two ways: either into two light quarks or into a charged lepton along with its neutrino. Hence there are three different decay modes for a $t\bar{t}$ pair, which are characterized by their number of charged leptons. If neither of the top quarks produces a lepton, the decay is *fully hadronic*. In case one of them decays into a charged lepton, the decay is called *semileptonic*. And lastly, there is also the possibility that both top quarks produce a charged lepton, which is then a *dileptonic* process. As τ leptons decay into a τ neutrino and a W boson, exactly the same is true for the τ-pair. The respective branching fractions are given in Table 4.1.

Decay mode	$t\bar{t}$	$\tau\tau$
Fully hadronic	44%	41%
Semileptonic	45%	46%
Dileptonic	8%	12%

Table 4.1: Branching fractions of $t\bar{t}$ and $\tau\tau$ decay modes.

Combining them yields $3 \cdot 3 = 9$ different possible final states of the $t\bar{t}H(\tau\tau)$ system. Again these can be distinguished from each other by the number of light leptons as well as τ-jets. Since a semileptonic decay is the most likely mode for both $t\bar{t}$ and $\tau\tau$ pairs, the combination of these modes is also the most likely for $t\bar{t}H(\tau\tau)$ events. It is the $2\,\ell, 1\,\tau_{\mathrm{had}}$ channel. Its advantage is a clear signature due to the charged leptons as well as some amount of $E_{\mathrm{T}}^{\mathrm{miss}}$ caused by the

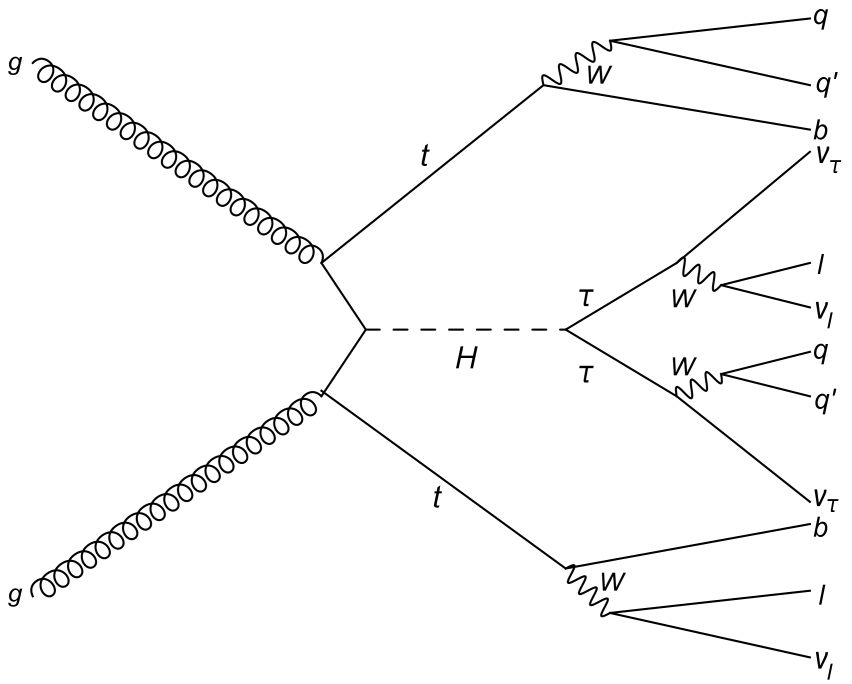


Figure 4.2: Feynman diagram of the $t\bar{t}H(\tau\tau)$ process resulting in a $2\,\ell, 1\,\tau_{\text{had}}$ final state

neutrinos. While the fully hadronic mode is also very frequent, it only exhibits jets in the final state and is much harder to distinguish from QCD background events. A decay mode only including leptons would be even more recognizable, but is very rare. Therefore the $2\,\ell, 1\,\tau_{\text{had}}$ channel offers a good compromise in recognizability and branching fraction. In the following a short description of the complex final state is presented, as well as a summary of the most important backgrounds, before turning to the generation and selection of events.

Signal Figure 4.2 shows a Feynman diagram depicting the $t\bar{t}H(\tau\tau)$ $2\,\ell, 1\,\tau_{\text{had}}$ channel. As can be seen, one of the leptons originates from a top quark decay, while the other one and the hadronic τ-decay belong to the $H \rightarrow \tau\tau$ system. It is important to note that, as they come from different decays, the two leptons can have the same charge. Additionally (at least) four jets are produced by the quarks from the $t\bar{t}$ decay. Two of these jets should ideally be b-tagged. Four neutrinos are produced in total: three of them originate from $H \rightarrow \tau\tau$ and one is produced in the leptonic top-decay. Since neutrinos are not detected, the final state cannot be fully reconstructed.

Backgrounds An overview of the production cross sections of $t\bar{t}H$ as well as the most important background processes is given in Table 4.2 for pp-collisions at center of mass energies of 8 and 14 TeV.

As can be seen top quark pairs are created abundantly in collisions at the LHC, resulting in the largest background contribution. In fact the $t\bar{t}$ production cross section is three magnitudes larger than that of the signal and other backgrounds. As we have already restricted the considerations to one specific final state, only certain $t\bar{t}$ topologies contribute. The most

$\sqrt{s}$	$\sigma(t\bar{t}H)$	$\sigma(t\bar{t})$	$\sigma(t\bar{t}W)$	$\sigma(t\bar{t}Z)$
8 TeV	0.129 pb	253 pb	0.232 pb	0.206 pb
14 TeV	0.611 pb	978 pb	0.769 pb	1.1 pb

Table 4.2: Predicted cross-sections for $t\bar{t}H$ as well as background events in pp-collisions at $\sqrt{s}$ = 8 TeV and 14 TeV (values taken from [30], [32–35]).

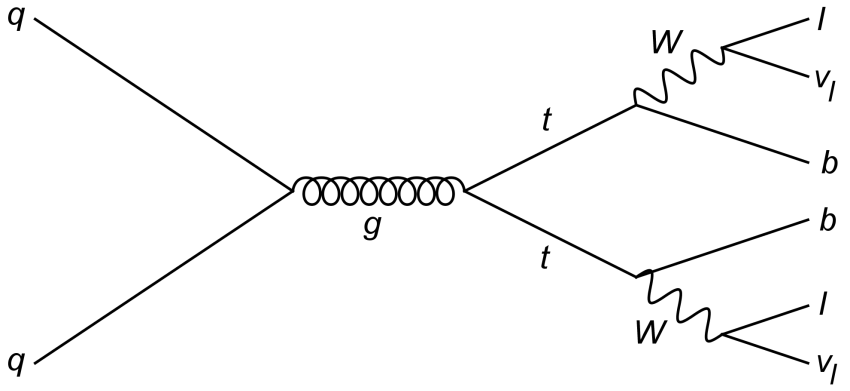


Figure 4.3: Feynman diagram of $t\bar{t}$ pair production with subsequent dileptonic decay.

likely way in which $t\bar{t}$ events can enter the $2\,\ell, 1\,\tau_{\text{had}}$ selection is by both of the top quarks decaying into charged leptons. This process is depicted in Fig. 4.3. There are then essentially two possibilities for these kind of events to contribute: First, the $t\bar{t}$ decay could happen fully dileptonic and the detected τ-jet is a misidentification of another kind of object. Secondly, one of the top quarks decays directly into a light lepton and neutrino, while the second decays into a τ lepton and neutrino, of which the first then subsequently decays hadronically. In this case it would then be the second lepton which would be fake. Therefore in processes of the latter type, the two leptons can also exhibit the same charge. It is important to note that both of these topologies include at least one misidentified object.

Another important background, which even on truth level can yield the same final state of $2\,\ell, 1\,\tau_{\text{had}}$ and is therefore *irreducible*, is the production of vector bosons in association with top quark pairs. Their production cross sections are of about twice that of the signal. As the Z boson boson is electrically neutral it can directly decay into τ-pairs and hence such events leading to a $2\,\ell, 1\,\tau_{\text{had}}$ final state look exactly like the signal process, except with a Z boson instead of the Higgs. Since the W boson is charged, it decays into a charged lepton and its corresponding neutrino. Hence in that case both top quarks as well as the W have to decay into leptons to produce the selected final state. The hadronic τ-decay can then either be a product of the W boson or one of the top quarks. Since these specific decay topologies of $t\bar{t}V$ events require leptonic decays of the vector bosons, which exhibit the lowest branching fractions, the contribution of such irreducible background events is not as relevant as that of $t\bar{t}$ production.

There are several other kinds of events which contribute to the background to the $t\bar{t}H$ signal but as they are not as important as the very frequent $t\bar{t}$ and irreducible $t\bar{t}Z$/ $t\bar{t}W$ production,

they are not described here. Details can be found in [30]. Moreover due to its high cross section in the following studies only the $t\bar{t}$ background process is considered.

4.3 Event generation and selection

Commonly analyses of high energy physics processes obtain their results from a comparison of measured to simulated events. This is especially important for analyses of data from hadron colliders, as the exact activity of such a complex collision cannot be predicted. Rather the typical shape of events is simulated using a phenomenological approach. Additionally, the detector response is also modelled. Therefore using simulated events one has access to event information on three different levels. First, there is the true event as simulated by the generator and as it would be measured with an ideal resolution. This is referred to as the *truth level*. After the simulation of the detector response the events look exactly as they would be measured and reconstructed from raw data. This level is then the *reconstruction level*, where nothing but the stable final state particles are known. As an intermediate between the two one can match reconstructed particles to their truth counterparts in order to obtain additional information, for example the parent particles from which they originate. This is then the *truth matched level*, where the detector resolution is accounted for but the true event shape is accessible.

For this thesis the Monte Carlo generated events created for and used by the ATLAS $t\bar{t}H$ multi-lepton final state analysis as presented in [30] are used. Specifically the signal sample was produced using the POWHEG BOX 1.0 generator [36] in combination with the PYTHIA 8.1 tool [37], which models the showering and hadronization processes. The $t\bar{t}$ sample is generated with a combination of POWHEG and PYTHIA 6.

Out of these simulated events a well defined selection has to be made, in order to optimize the analysis. The basic event selection here corresponds to two selected leptons, which fulfil $p_{\mathrm{T}}^{l0} > 25\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{l1} > 15\,\mathrm{GeV}$, where the leptons are ordered by their p_{T} value $p_{\mathrm{T}}^{l0} > p_{\mathrm{T}}^{l1}$. This is done as the efficiency of the trigger requiring at least one lepton is p_{T} dependent. Additionally for this channel at least one hadronic τ decay is required. To increase the number of events, the "loose" working point of the hadronic tau identification is chosen. These requirements correspond to the basic event selection. It is shown as the starting point in Table 4.3, which lists the resulting number of expected $t\bar{t}H$ signal and $t\bar{t}$ events, weighted by the production cross sections, for each cut applied in the current ATLAS analysis.

Cut applied	$t\bar{t}H(\tau\tau)$	$t\bar{t}$
Basic selection	2.22 ± 0.10	1302.5 ± 18.8
Same sign leptons	0.82 ± 0.05	21.7 ± 2.4
$N_{\text{b-tag}} \geq 1$	0.66 ± 0.05	10.4 ± 1.7
$N_{\text{jet}} \geq 4$	0.39 ± 0.04	1.5 ± 0.6

Table 4.3: Number of expected $t\bar{t}H$ signal and $t\bar{t}$ events after each cut as applied in the ATLAS analysis of the $2\,\ell, 1\,\tau_{\text{had}}$ channel. Numbers are normalized to the respective cross sections at $\sqrt{s} = 8$ TeV.

One of the basic criteria by which background events are excluded in this analysis is the

requirement for the two detected leptons to have the same charge. This removes most of the $t\bar{t}$ background by excluding the possibility for the two leptons to originate from the same process. Only $t\bar{t}$ topologies with a fake lepton are left. However, as can be seen, more than half of the signal events are also removed. Moreover, a cut on the expected number of b-tagged jets can be imposed. To account for the limited tagging efficiency this is set to at least one. Lastly, at least four jets are required. Again these cuts remove sizeable fractions of the number of signal events.

As the used MC samples were generated for an inclusive analysis of rare processes, the number of generated events available is also very low. Using the basic event selection as well as a truth cut to exclude decays of the nature $t \to Wb \to \tau\nu_\tau b$ within the $t\bar{t}$ system[1] only roughly 650 generated events are left. Therefore and seeing that the aim of this thesis was to investigate methods going beyond a cut-and-count procedure, this basic event selection is chosen for the following studies. Specifically even the same charge requirement is dropped.

4.4 Invariant mass reconstruction

The reconstruction of the invariant mass of a decaying particle is a powerful tool, which will be used for two different purposes throughout this thesis. For one, the shape of the invariant mass distribution reconstructed from the $H \to \tau\tau$ decay can help distinguish signal from background. Secondly, the invariant mass distributions of the other particles involved in the process can be used in the definition of likelihoods by which the event kinematics be described. Ideally, a precise measurement of the four momenta of all decay products involved is required, as the invariant mass is obtained from the total four momentum of the decaying particle:

$$m_0^2 = P_0^2 = \left(\sum_{i=1}^{n} P_i\right)^2 = \left(\sum_{i=1}^{n} E_i\right)^2 - \left(\sum_{i=1}^{n} \vec{p}_i\right)^2 \tag{4.1}$$

Here m_0 and P_0 denote the invariant mass and four momentum of a particle decaying into n particles with four momenta P_i, energies E_i and spatial momenta $\vec{p}_i$. Since it is a scalar product of two four vectors, m_0 is invariant under Lorentz-transformations. For an on-shell particle decay it corresponds to its rest mass. In case the decay can also be off-shell, the invariant mass distribution shows a resonance that can be described by a relativistic Breit-Wigner distribution

$$B(m_{\text{inv}}|m_0, \Gamma) \propto \frac{\Gamma^2 m_0^2}{(m_{\text{inv}}^2 - m_0^2)^2 + m_{\text{inv}}^4(\Gamma^2/m_0^2)} \tag{4.2}$$

where Γ refers to the decay width of a particle with rest mass m_0 [38].

Due to the production of multiple neutrinos reconstructing the invariant mass of a $H \to \tau\tau$ system is especially challenging. In this case $E_{\text{T}}^{\text{miss}}$ is only an approximation of the vectorial sum of the transverse momenta of all neutrinos, hence their individual momenta remain unknown. In any case no information on the longitudinal momentum component of the neutrinos is available. The simplest approach to avoid this problem is to only reconstruct the invariant mass from the visible decay products into the so-called visible mass, $m_{\tau\tau}^{\text{vis}}$. For the

[1] This is done as the neutrinos produced in an additional τ-decay would further complicate the situation.

semileptonic τ-pair decay this is then given by

$$m_{\tau\tau}^{\text{vis}} = \sqrt{\left(E_{l_H} + E_{\tau_H^{\text{vis}}}\right)^2 - \left(\vec{p}_{l_H} + \vec{p}_{\tau_H^{\text{vis}}}\right)^2} \tag{4.3}$$

where l_H denotes the lepton and τ_H^{vis} the visible products of the hadronically decaying τ lepton produced in the $H \to \tau\tau$ decay. Since the four momenta carried by the neutrinos are neglected, the visible mass is smaller than the Higgs mass in this case. In the course of the $H \to \tau\tau$ analysis more complex methods to reconstruct the invariant $\tau\tau$-mass have been developed. One of these is based on the *collinear approximation* according to which the final decay products stay collinear to the direction of the two intermediate taus. Since this technique cannot be used if the two taus are produced in opposing directions, it is not applicable for all event topologies. To overcome this shortcoming a more complex algorithm, the *missing mass calculator* (MMC) was developed [39]. It facilitates a full reconstruction of the event kinematics for nearly all topologies. In principle, parametrized distributions for the unknown quantities (i.e. the neutrino momenta) are used to determine the most likely neutrino configuration for each event. It is important to note that both the collinear and the MMC mass are based on the assumption that $E_{\text{T}}^{\text{miss}}$ exclusively results from the neutrinos produced in the $H \to \tau\tau$ decay.

4.5 Visible mass of the $\tau\tau$-system in the $2\,\ell, 1\,\tau_{\text{had}}$-channel

It stands to reason that the reconstruction of the invariant mass of a decaying particle can only be accurate, if the measured objects used by the mass estimator originate from that decay. For the $t\bar{t}H(\tau\tau)$ $2\,\ell, 1\,\tau_{\text{had}}$ channel this is of particular importance: here two leptons are selected, either of which can belong to the $t\bar{t}H(\tau\tau)$-decay. Thus criteria to decide which one of them to use for the mass reconstruction are needed.

Before discussing this problem in more detail, the potential mass resolution and separation power achieved by reconstructing $m_{\tau\tau}^{\text{vis}}$ is investigated. For this purpose, an ideal identification of ℓ_H and τ_H^{vis} is simulated by using a truth match. The reconstructed objects are matched to their truth counterparts via minimal angular distance ΔR and hence linked to their parent particles. No maximum ΔR is used. A truth match is also implemented for $t\bar{t}$ background events requiring the lepton and hadronic τ-decay to originate from a top quark decay. In this way only the $t\bar{t}$ topology including a fake lepton is considered. More detailed studies of different background topologies are also of interest but go beyond the scope of this thesis.

In Fig. 4.4, the resulting mass spectra of signal and $t\bar{t}$ pair production background events are compared. As can be seen, the signal distribution exhibits a fairly narrow peak, while it is broader and tends to higher values for $t\bar{t}$ events. In terms of shape separation as well as mass resolution reconstructing the visible $\tau\tau$-mass already seems promising. A criterion by which to judge the separation potential more quantitatively will be introduced in the following.

However transferring this truth matched case to reconstruction level is not trivial. Two charged leptons are selected in this channel and correctly associating one of them into the $H \to \tau\tau$ system is crucial for the success of the mass reconstruction. The possibility to reconstruct the mass from the lepton which is produced in one of the top decays, ℓ_{top}, creates a large combinational background. This is illustrated by Fig. 4.5 in which the truth matched

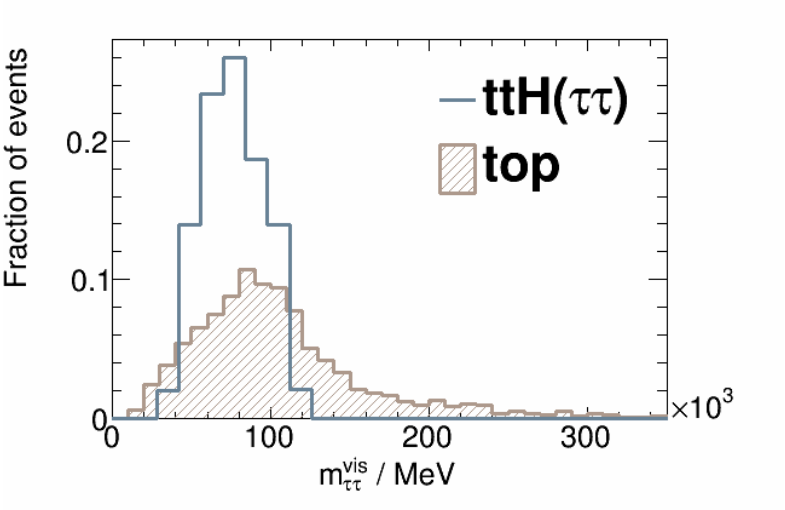


Figure 4.4: Comparison of the visible mass distribution for signal and $t\bar{t}$-pair production background events. Here the truth matched ℓ_H and τ_H^{vis} are used signal events, while for background events lepton and tau are required to originate from a top quark decay.

$m_{\tau\tau}^{\text{vis}}$ distribution is compared to the combination using the wrong lepton, ℓ_{top}. In the latter case the distribution is broadened and it loses its distinct narrow peaked shape, appearing more similar to that of the $t\bar{t}$ pair-production background. For this worst case scenario signal and background distribution would overlap significantly and their shapes would be indistinguishable. It is important to note that only events where one lepton originates from the Higgs and the other one from a top decay are taken into account. The effect of fake leptons that aren't part of the $t\bar{t}H(\tau\tau)$ process is therefore excluded, as it would introduce further ambiguity.

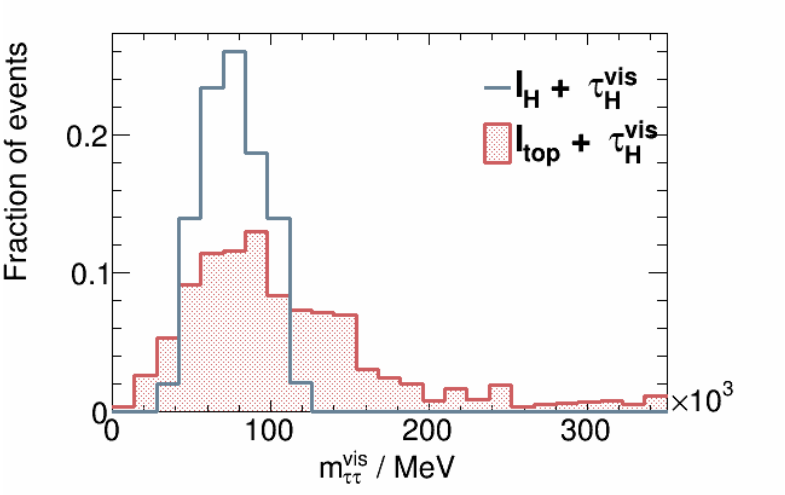


Figure 4.5: Comparison of the true $H(\tau\tau)$ visible mass distribution to the combinatorial background caused by the lepton from the $t\bar{t}$-decay.

In order to estimate to what extend the mass reconstruction is affected by the ambiguity of the lepton, the correct lepton is drawn at random but with a fixed probability. Figure 4.6

shows an overview of the visible mass distributions resulting from choosing the correct lepton ℓ_H with four different probabilities: in Fig. 4.6.(a) the correct lepton is chosen in 30% of the events, in (b) in 50%, in (c) in 70% and in (d) in 90%. Each of these distributions is compared to the truth matched case with a 100% reconstruction efficiency of ℓ_H. Since a choice out of two leptons is made, one would expect to achieve a 50% success rate simply by guessing.

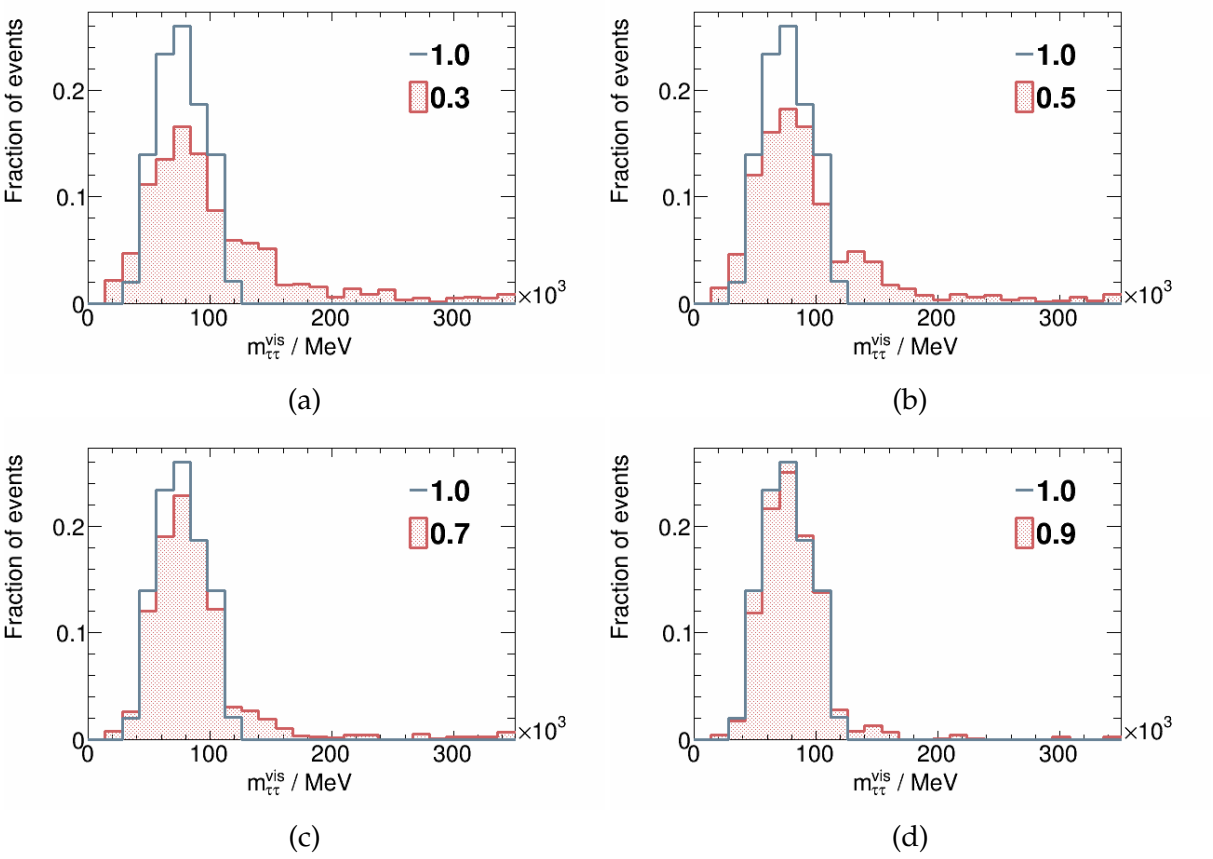


Figure 4.6: Comparison of the truth matched $H(\tau\tau)$ visible mass distribution using ℓ_H and τ_H^{vis} to its combinatorial background due to ℓ_{top} when randomly selecting either of the two leptons with four different probabilities: In (a) ℓ_H is chosen in 30% of the events, in (b) in 50%, in (c) in 70% and in (d) in 90% . The truth matched case corresponds to 100% .

The less often the correct lepton is used, the broader the mass distribution becomes. To quantify this effect, Table 4.4 lists their mean, μ, and standard deviation, σ, estimating their mass value and width respectively. Moving away from the ideal case and decreasing the fraction of events in which the correct lepton ℓ_H is used, the distribution's mean shifts to higher values as it broadens. Additionally the width increases as well, corresponding to a decrease in achieved mass resolution.

Moreover to judge the quality of the mass reconstruction on a substantiated basis the corresponding distribution for background events needs to be considered. Here a good quantifier is the signal to background ratio $\frac{S}{B}$: the number of signal events S found in a certain range of invariant mass values is divided by the number of background events B found in that same

Probability	μ **in** GeV	σ **in** GeV
100%	76.5 ± 0.9	18.9 ± 0.7
90%	81.2 ± 1.4	29.1 ± 1.0
70%	86.6 ± 2.1	42.4 ± 1.5
50%	94.4 ± 2.7	53.6 ± 1.9
30%	101.2 ± 2.9	59.1 ± 2.1
0%	109.3 ± 3.2	64.6 ± 2.3

Table 4.4: Characteristics of the visible mass distributions for different probabilities to choose the right lepton when drawing at random.

range. The respective mass range is defined by $(\mu - \sigma, \mu + \sigma)$ for each signal distribution and thus is flexible. Again the background distribution is kept fixed to the mass reconstructed from a lepton and τ-jet that both originate from a top quark decay. The resulting mass distributions are shown in Fig. 4.7.

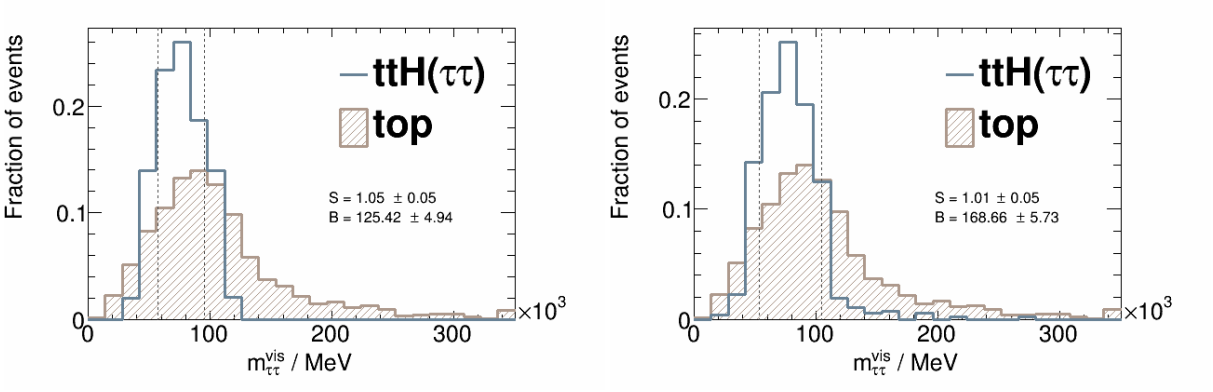
As the mass spectrum broadens, the considered range also becomes larger for each distribution and in turn, more and more background events fall within that range. Thus the signal to background ratio decreases significantly as can be seen in Table 4.5.

Probability	**S**ignal	**B**ackground	**S/B** $\cdot 10^3$
100%	1.05 ± 0.05	125.4 ± 4.9	8.37 ± 0.52
90%	1.01 ± 0.05	168.7 ± 5.7	5.99 ± 0.36
70%	1.37 ± 0.06	263.1 ± 7.2	5.21 ± 0.27
50%	1.30 ± 0.06	277.0 ± 7.3	4.69 ± 0.25
30%	1.28 ± 0.06	292.0 ± 7.5	4.38 ± 0.23
0%	1.21 ± 0.06	297.9 ± 7.6	4.06 ± 0.23

Table 4.5: Ratio of signal over background events within the mass range $(\mu(t\bar{t}H) - \sigma(t\bar{t}H), \mu(t\bar{t}H) + \sigma(t\bar{t}H))$ for different probabilities to choose the right lepton when drawing at random.

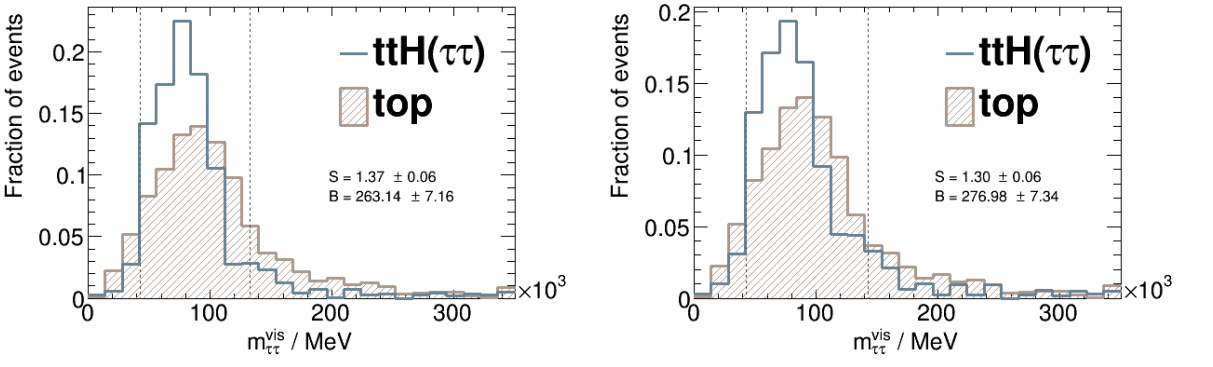
It is important to note that drawing either of the leptons at random while keeping the mass spectrum of background events fixed does not describe realistic conditions. Every method for associating the lepton into a decay system has to be based on certain assumptions regarding measured quantities. If those factor into the mass reconstruction the principle of the method will impact the shape of the mass spectrum of signal and background events in a similar way. The quality of the method can then only be judged properly by also considering how it shapes the background distribution. Nonetheless the studies performed under such simplified conditions already prove the significance of the lepton ambiguity.

All in all, it has become apparent that finding a way by which to associate one of the leptons correctly with the $t\bar{t}H(\tau\tau)$ system is of utmost importance. This problem needs to be ad-



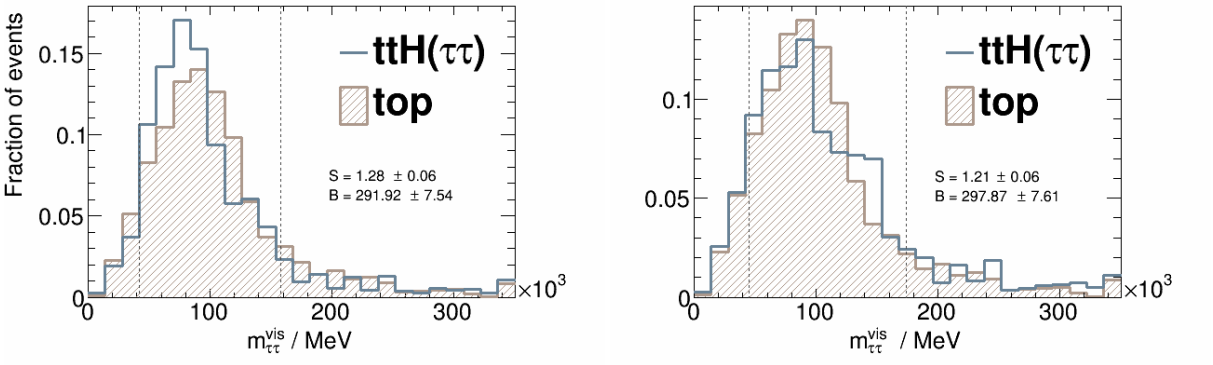
(a) : 100% ℓ_H

(b) : 90% ℓ_H



(c) : 70% ℓ_H

(d) : 50% ℓ_H



(e) : 30% ℓ_H

(f) : 0% ℓ_H

Figure 4.7: Comparison of the $H(\tau\tau)$ visible mass distribution using four different probabilities to use the correct lepton ℓ_H. The dashed lines indicate the range $(\mu - \sigma, \mu + \sigma)$ based on the signal distribution.

dressed before more sophisticated mass reconstruction can be employed. Due to the complexity of the selected final state a solution is not easily found. Different approaches to a possible solution are the main focus of the following studies.

CHAPTER 5

Simple variables

Naively one would expect that there are topological characteristics in $t\bar{t}H(\tau\tau)$ events, by which the Higgs decay system is set apart from that of the top quarks. The correct lepton, ℓ_H, would then fulfil the expected characteristic when combined with τ_H^{vis} while the wrong lepton, ℓ_{top}, shouldn't. In this chapter such simple topological variables and their potential to associate the leptons correctly are studied on the truth matched level.

5.1 p_{T} ordering of the leptons

The possibility to distinguish the two leptons by their p_{T} ordering is considered as a first simple attempt. In the top quark decay the lepton is directly produced by the W boson, while in the $H \to \tau\tau$ system there is one additional intermediate step in form of the τ lepton. On the other hand, more phase space is available in the $H \to \tau\tau$ decay than in the top quark decay due to the masses of the particles involved. To some degree these two effects could cancel each other, assuming that the decaying particles don't have very different p_{T} values to begin with. The p_{T} distributions of the true Higgs and leptonically decaying top quark shown in Fig. 5.1 confirm that the latter assumption is valid.

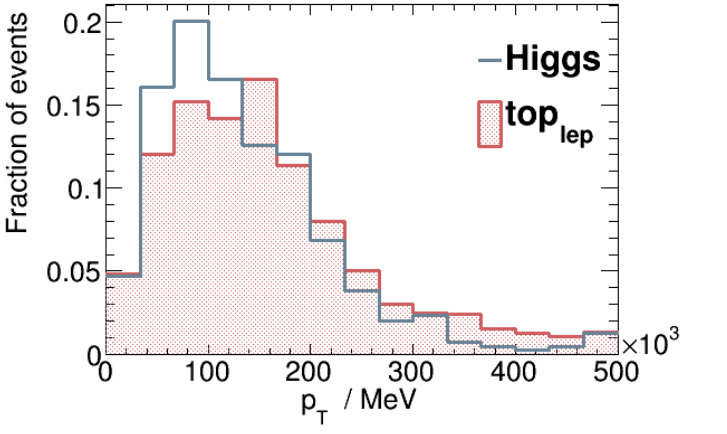


Figure 5.1: Comparison of the p_{T} distributions of the true Higgs and leptonically decaying top quark.

In fact the transverse momentum distribution for the lepton from the top quark decay only tends to slightly higher values than that of ℓ_H, as can be seen in Fig. 5.2. Quantifying this

effect, in (69.0 ± 4.4)% of the events the leading p_T lepton originates from the top quark and the one with lower p_T from the Higgs decay. Already this method to find ℓ_H appears far from ideal.

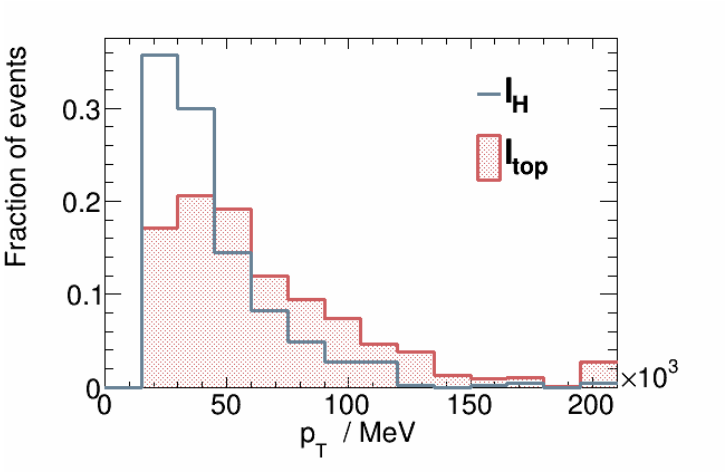


Figure 5.2: Comparison of the p_T distributions of leptons from the $H \to \tau\tau$ decay, ℓ_H, and of leptons from a top quark decay, ℓ_{top}.

Nonetheless reconstructing the invariant mass from the lower p_T lepton and the correct τ_H^{vis} results in a peak at lower values, with only few events yielding high invariant masses, as can be seen in Fig. 5.3. While with this method the mean value of the mass distribution at (76.8 ± 1.6) GeV is very similar to that of the truth matched case, the standard deviation is increased up to (31.0 ± 1.1) GeV. This emphasises that the mass resolution deteriorates, as the distribution broadens.

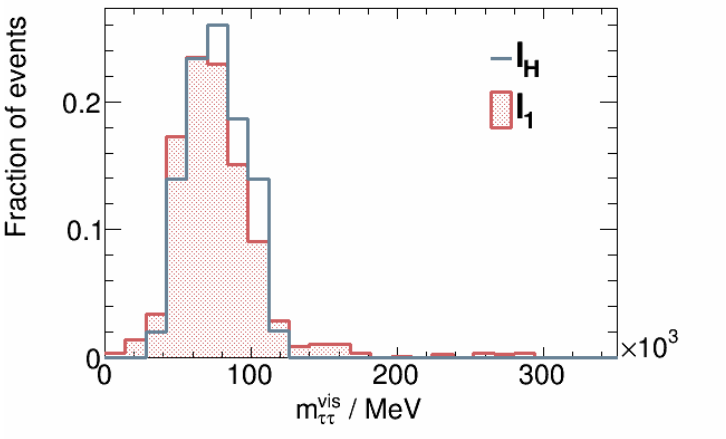


Figure 5.3: Comparison of the visible mass distribution for signal events when using the truth matched combination of ℓ_H and τ_H^{vis} and when using the lower p_T lepton, l_0 in combination with the true τ_H^{vis}.

Considering that the invariant mass is derived from the sum of the four vectors the shift to lower invariant masses is most likely an artificial effect: Always choosing a low p_T lepton

forces the invariant mass to take accordingly low values. To confirm this the effect of the method on the mass spectrum of $t\bar{t}$ pair production events is also studied. For comparison again the very specific selection of background events is used, in which the lepton and the τ-jet are required to originate from an actual top quark decay. Figure 5.4 shows the resulting mass distributions.

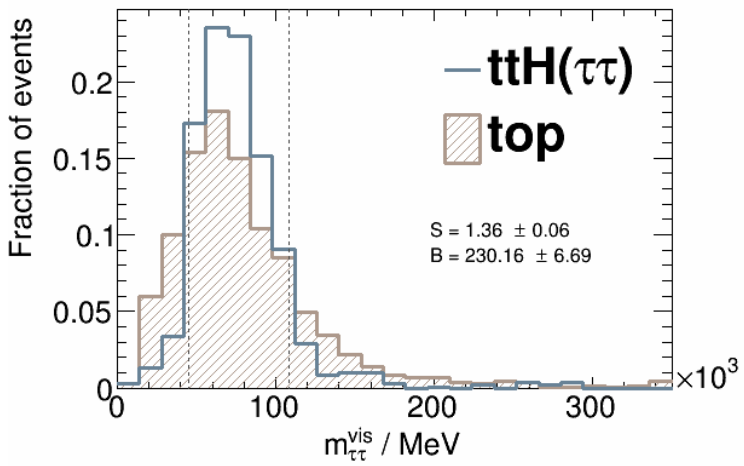


Figure 5.4: Comparison of the visible mass distribution using the lower p_{T} lepton for signal and $t\bar{t}$-pair production background events. The τ-jet is truth matched to a Higgs or top decay respectively.

Here the peak is also shifted to low values for the background events and most of the shape separation is lost this way. Due to this effect and also due to the increase in the standard deviation for the signal spectrum, more events - of both signal and background processes - are found within the mass interval defined as before. The signal to background ratio is then only moderate at $(5.91 \pm 0.31) \cdot 10^{-3}$.

To summarize, associating the lepton to the Higgs decay via its lower p_{T} value does not reach sufficient success. Moreover it also strongly shapes the background spectra in much the same way as the signal, since the invariant mass is correlated to the p_{T} of the particles considered. Therefore this method seems unsuitable for a practical application to data.

5.2 Distances between final state objects

5.2.1 Distance ΔR to the τ_H^{vis}

Looking back at Fig. 5.1, which shows the p_{T} distributions of the Higgs and the leptonically decaying top quark, it is evident that both often possess considerable transverse momenta. This means that in most cases their decays are boosted, resulting in small distances between the decay products. Thus it seems intuitive that the lepton from the Higgs decay should be closer in ΔR to the τ-jet than the lepton from a top quark decay.

Comparing the distributions of ΔR (l,τ_H^{vis}) as shown in Fig. 5.5, indeed the peak is situated

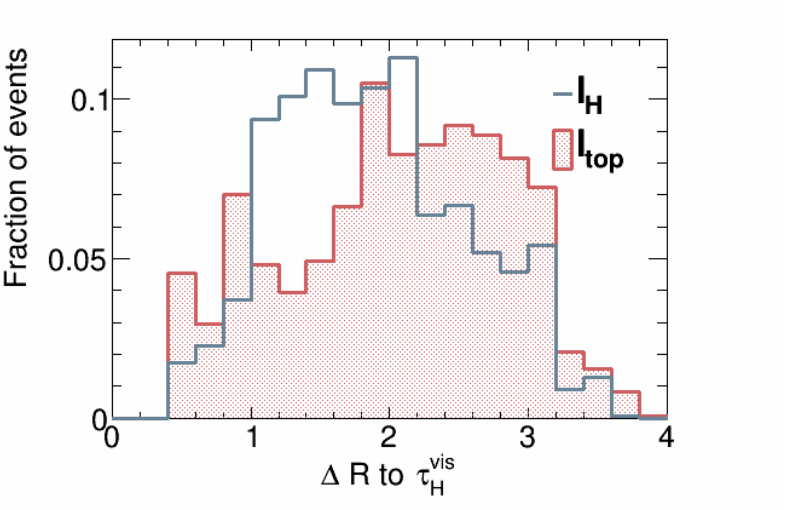


Figure 5.5: Angular separation ΔR between τ_H^{vis} to the lepton from the Higgs and top decay.

at higher values for the lepton from the top decay than for that of the Higgs decay [1]. Thus choosing the lepton that is closer in ΔR to τ_H^{vis} is also a simple variable to associate the lepton with the $H \to \tau\tau$ system. Doing so, the correct lepton ℓ_H is found in (56.3 $\pm$ 3.8)% of the events and the visible mass distribution shown in Fig. 5.6 is obtained.

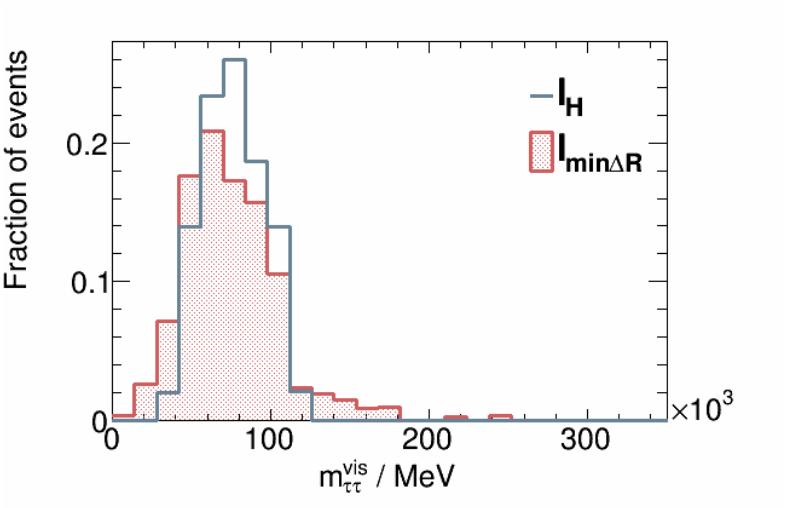


Figure 5.6: Comparison of the visible mass distribution for signal events when using the truth matched combination of ℓ_H and τ_H^{vis} and when using the lepton closer in ΔR to the tau, $l_{\min\Delta R}$, in combination with τ_H^{vis}.

Again the peak is slightly shifted to lower values but overall the distribution broadens. The mean value using the smaller ΔR criterion is at (75.2 $\pm$ 1.5) GeV and the standard deviation is (30.6 $\pm$ 1.1) GeV.

[1] Here and in the following ΔR distributions no small values of ΔR occur due to the overlap removal.

The angular distance ΔR between the particles taken into account is directly correlated to the invariant mass value reconstructed, since both of them are dependent on the angle between the particles. Therefore when using this method, the background distribution is also shaped in much the same way, as can be seen in Fig. 5.7. Here the signal to background ratio within the mass range around the mean of the signal spectrum is $(5.76 \pm 0.32) \cdot 10^{-3}$.

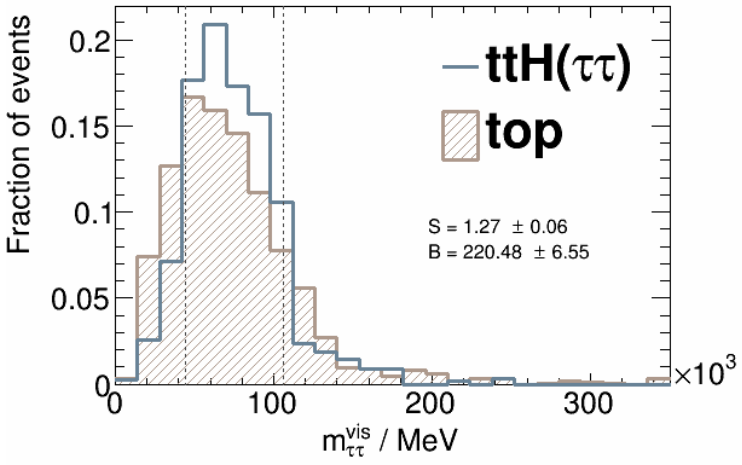


Figure 5.7: Comparison of the visible mass distribution using the lepton that is closer in ΔR to the τ-jet for signal and $t\bar{t}$-pair production background events. The τ-jet is truth matched to a Higgs or top decay respectively.

All in all finding ℓ_H by its smaller ΔR to the τ-jet isn't more successful than using the smaller p_{T} value. It suffers from much the same shortcoming, namely that it also strongly affects the mass spectrum for background events. Thus, this method is also likely not applicable under realistic analysis conditions.

The distance ΔR combines both the difference in pseudo-rapidity, $\Delta\eta$, and the azimuthal angle, $\Delta\phi$, into one variable. For the sake of completeness, the distributions of those two variables are shown separately as well in the appendix in Fig. A.2. Neither of them showcases a characteristic difference that could be exploited.

5.2.2 Distance ΔR to the closest true b-jet

Another angular separation that is worth considering is that to a b-tagged jet: the lepton from the top quark decay could naively be expected to have a smaller distance ΔR to a jet caused by a bottom quark. Hence in Fig. 5.8 the angular separation ΔR of both leptons to the closest true b-jet is compared. To exclude the influence of the inefficiencies of the b-tagging algorithm the true jet flavour is used. The shapes of the two distributions resemble each other closely, with only a slight tendency for the lepton ℓ_H to yield higher values of ΔR. Thus there appears to be no potential to base the decision which lepton belongs to the Higgs decay on the ΔR to the nearest b-jet alone. Again, the same is true for the separate distributions of $\Delta\eta$ and $\Delta\phi$, which are shown in Fig. A.3 in the appendix.

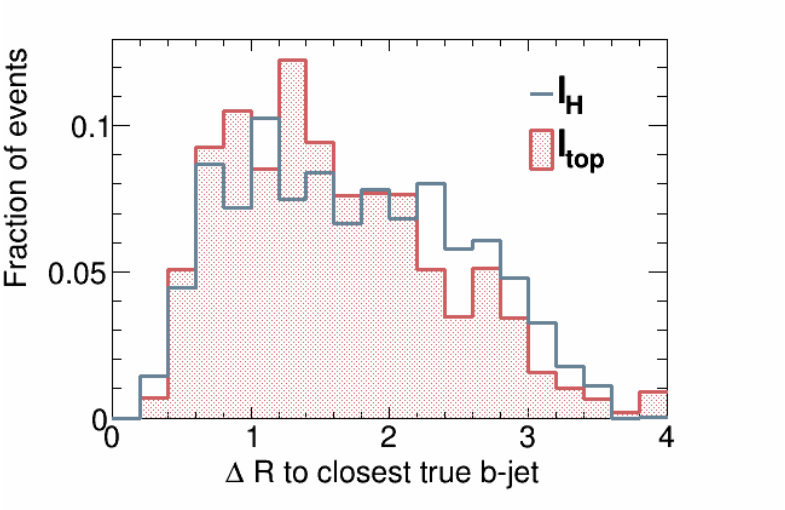


Figure 5.8: Angular separation ΔR between the leptons from the Higgs and top decay to the closest true b-jet respectively.

5.2.3 Distance ΔR to any closest jet

Similarly, excluding the τ-jet, all of the remaining jets are products of the $t\bar{t}$ decay system. Hence if there is any kind of spatial separation of the two decay systems in $t\bar{t}H(\tau\tau)$ events that remains throughout the multiple decay processes until the final state, the lepton from the $t\bar{t}$ system should more likely be closer in ΔR to any non-tau jet.

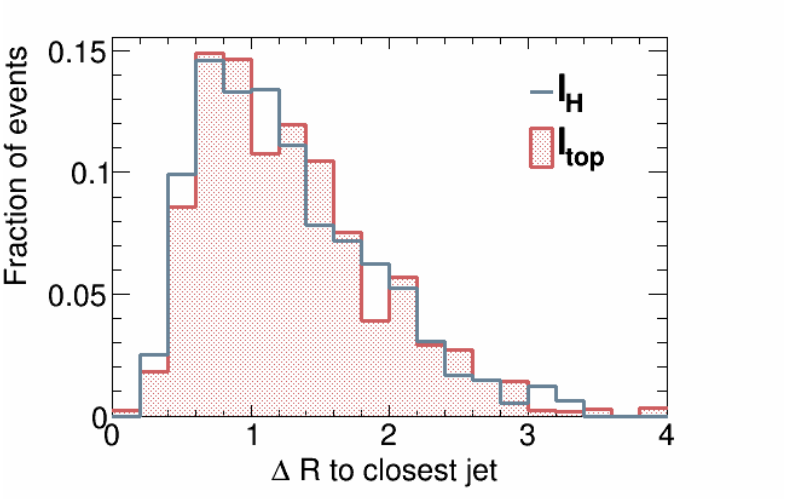


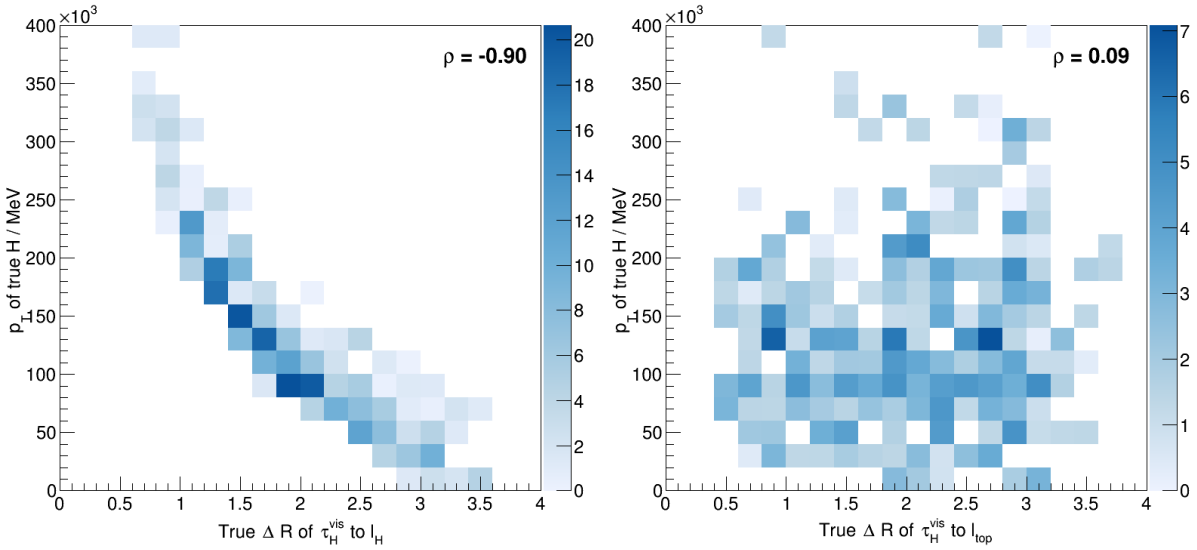
Figure 5.9: Angular separation ΔR between the leptons from the Higgs and top decay to the closest jet respectively.

Yet, as Fig. 5.9 comparing the ΔR distributions for ℓ_H and ℓ_{top} shows, this is not the case. Again the same holds for the $\Delta\eta$ and $\Delta\phi$ distributions which can be found in Fig. A.4.

5.3 Exploiting the p_T (ΔR)-correlation

While the variables found unsuccessful in the above studies all focused on individual aspects of the decay kinematics of $t\bar{t}H(\tau\tau)$ events, combining multiple observables may be more powerful. One method to do this is to exploit the expected dependency of p_T on ΔR within the Higgs decay. On truth level, an inverse correlation between the p_T of a decaying particle and the ΔR between its decay products is expected. This correlation should only be valid when combining particles which originate from the same decay.

Figure 5.10 confirms this for $t\bar{t}H(\tau\tau)$ events on truth level. It shows the true Higgs p_T against the ΔR between the visible hadronic τ-jet - excluding all the neutrinos of that decay - and the lepton ℓ_H and ℓ_{top}, in (a) and (b) respectively. Here, the combination shown in (a) corresponds to the correct one and the one in (b) should be nonsensical. For ℓ_H a clear correlation can be seen: the higher the p_T, the smaller the ΔR. This is also reflected in the *correlation coefficient* ρ of -0.9 , confirming an inverse correlation. On the other hand no such effect can be seen for using ℓ_{top}, which yields a near zero correlation coefficient of 0.09.

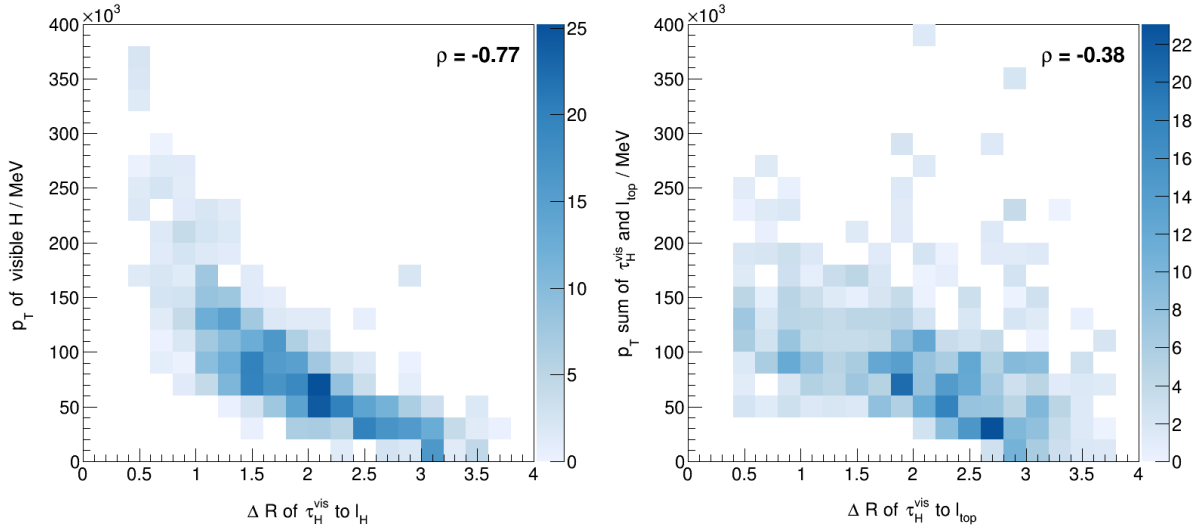


(a) (b)

Figure 5.10: True Higgs p_T against ΔR between τ_H^{vis} and ℓ_H, in (a), and ℓ_{top}, in (b), on truth level

In actual data the p_T of the Higgs is not directly accessible, but needs to be reconstructed. The $H \to \tau\tau$ decay is problematic to reconstruct due to the production of multiple neutrinos. While usually the missing transverse energy E_T^{miss} may be used to approximate the momenta of the escaping neutrinos, this is not possible in the case of $t\bar{t}H(\tau\tau)$ events, as the leptonic top quark decay also contributes one neutrino. To keep the considerations focused fully how to associate the leptons with their decay systems, without introducing further ambiguities by attempting to separate E_T^{miss} between them, only the lepton and τ-jet are used here to reconstruct the *visible Higgs*.

The p_T of the visible Higgs, determined from its truth matched visible decay products is



(a) (b)

Figure 5.11: Visible Higgs p_{T} as obtained from the vectorial sum of τ_H^{vis} with ℓ_H or ℓ_{top} against the ΔR between lepton and τ , in (a) or (b) respectively, on reconstruction level

plotted against the angular distance ΔR of τ_H^{vis} to the correct lepton in Fig. 5.11.(a). Comparing with what was shown on truth level before, a slight degradation can be observed as the distribution broadens. Reasons for this effect are to be found both in the neglected neutrinos as well as the limited detector resolution. Nonetheless, when picking the correct lepton the inverse correlation is still obvious at $\rho = -0.77$. In Fig. 5.11.(b), the p_{T} of the vectorial sum of τ_H^{vis} and ℓ_{top} against their distance ΔR is shown. A slight correlation is also noticeable in this case, as the magnitude of the vectorial sum depends on the angles between the vectors which are added. Nevertheless, in this wrong combination of lepton and τ-jet more events stray far from the clear correlation between p_{T} and ΔR, as also evidenced by the lower correlation coefficient of $\rho = -0.38$

In order to exploit this further the particular profile expected for the correct lepton can be fitted. In this way, for any given ΔR between lepton and τ-jet an expected value for the p_{T} obtained when combining the two is calculated. For each lepton it can then be determined how much the actual p_{T} value deviates from the expectation. The lepton for which the p_{T} expected by the fit and its actual value agree better is assumed to originate from the Higgs decay. For this method, simulated events have to be used in order to determine the fit function before it can be applied to data.

In practice the scatter plot Fig. 5.11.(a) has to be projected into a profile along the ΔR axis. This profile then gives an approximated dependency of p_{T} on ΔR and is fitted with a *Landau distribution* by minimizing the residual sum of squares, χ^2 [2]. Since there is no analytical form of the latter, a parametrised implementation is used. Here the parameters are: A is the nor-

[2] There is no physical reasoning for using a Landau function, it was only empirically found to describe the dependency.

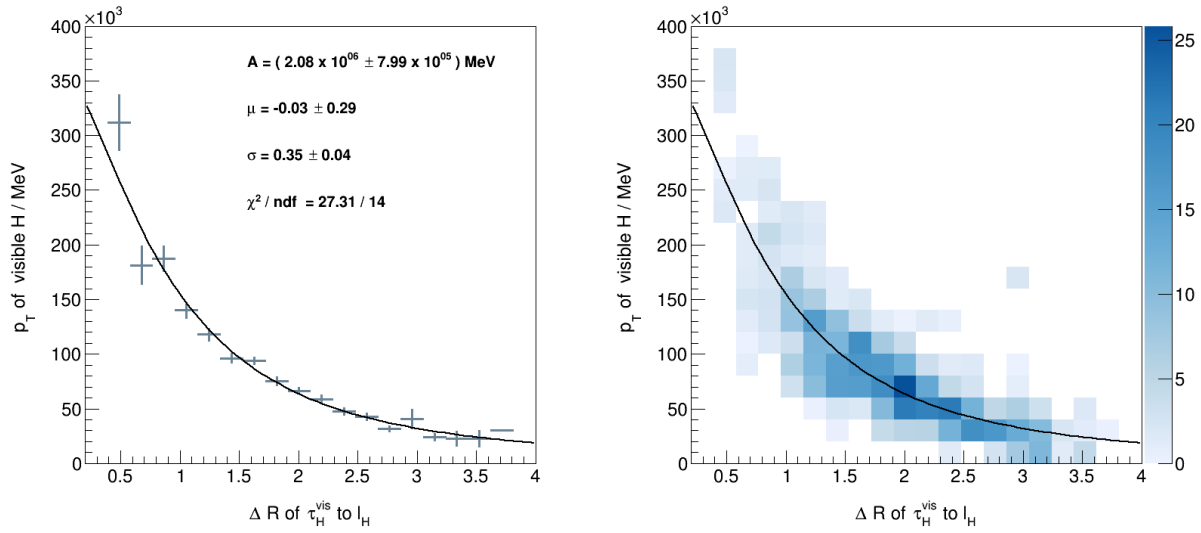


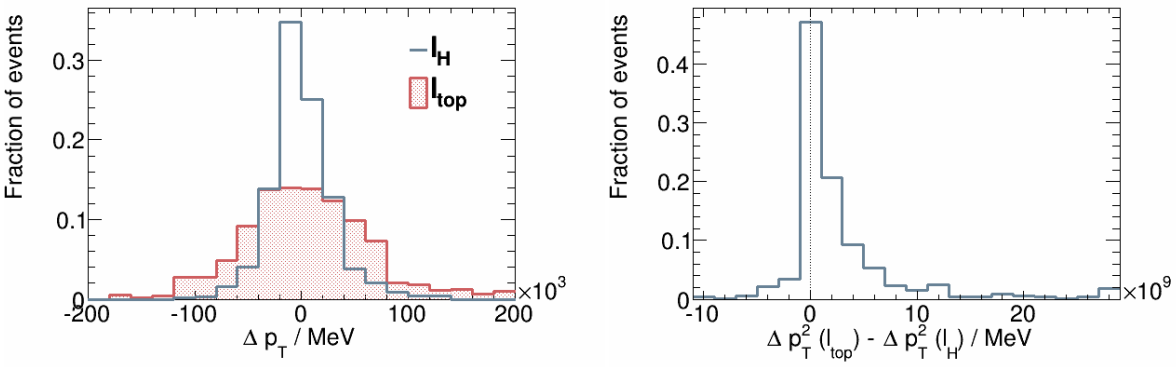
Figure 5.12: (a) Profile of the two dimensional distribution of visible Higgs p_T against ΔR of its constituents, Landau function to the profile as well as the parameters of the fit and (b) this fit function compared with the original two dimensional distribution.

malization constant, scaling the function to the correct height; μ is the most likely value for the distribution to take, hence describing the peak position and σ corresponds to a scale factor (but does not describe the distribution's width, as this is not defined). To quantify the goodness of fit here the χ^2 value can be used. It can be shown that its distribution takes a mean value which corresponds to the number of degrees of freedom (ndf) and a standard deviation of $\sqrt{2 \cdot \text{ndf}}$ [3].

Figure 5.12.(a) shows the profile and fit function as well as the resulting parameters of the fit. For comparison the fit function is also superimposed onto the original scatter plot in Fig. 5.12.(b). Looking at both of these illustrations and just judging by eye, the fit function appears to describe the distribution reasonably well. The χ^2 value of the fit deviates by approximately 2.51 standard deviations from its expected value. Thus the fit appears to be performed successfully within the bounds of statistical fluctuations. In fact it should be emphasized again that the severe shortage of statistics is detrimental to the conclusions which can be drawn.

For each lepton an expected p_T value is then determined using the fit function as well as the deviation of the reconstructed p_T sum from this expectation, given by $\Delta p_T = p_T(\vec{\tau} + \vec{\ell}) - p_T^{\text{fit}}(\Delta R_{\ell,\tau})$. Here the truth matched τ_H^{vis} is used with either of the leptons. As an intermediate validation of the approach, these deviations Δp_T are compared for the truth matched leptons ℓ_H and ℓ_{top}, shown in Fig. 5.13.(a). It is immediately apparent that when using the wrong lepton, ℓ_{top}, the Δp_T distribution is much broader and extends to much larger values than in the correct case. Nonetheless even for the correct lepton the p_T deviations can become quite

[3] See e.g. [40]

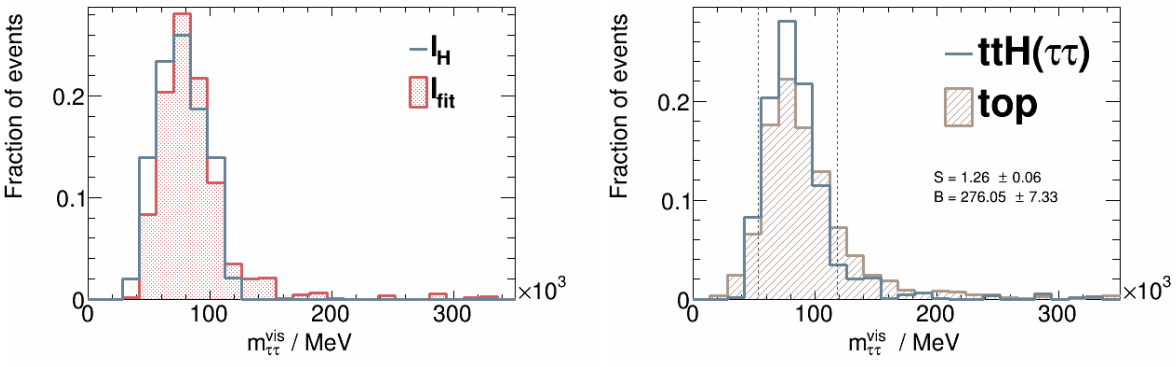


(a) (b)

Figure 5.13: (a) Deviations of the p_{T} value reconstructed from τ_H^{vis} and leptons ℓ_H and ℓ_{top} from the value expected by the p_{T} (ΔR) fit and (b) difference of the squared p_{T} deviations between the two leptons.

large. While these characteristics of the distributions are a first indicator that indeed picking the lepton based on its Δp_{T} from the fit is a sensible approach, a stronger case can be made when comparing the p_{T} deviations for each event. This is done by subtracting the Δp_{T} values determined for each of the leptons from each other. Since the direction of the deviation is irrelevant, the square of this deviation $(\Delta p_{\mathrm{T}})^2$ is used (which was also found empirically to lead to the best results). The resulting distribution is shown in Fig. 5.13.(b). In events for which the value of the difference is greater than 0 the wrong lepton ℓ_{top} yields a larger Δp_{T} than the correct one ℓ_H. Evidently the larger part of the events fulfils this, but in most cases the difference is very small.

If the p_{T} (ΔR)-fit approach is employed as a criterion to distinguish between the leptons, the correct lepton, ℓ_H, is found in (67.2 $\pm$ 5.8) % of the signal events. In this respect there apparently is no improvement compared to the simpler method based only on the p_{T} ordering.

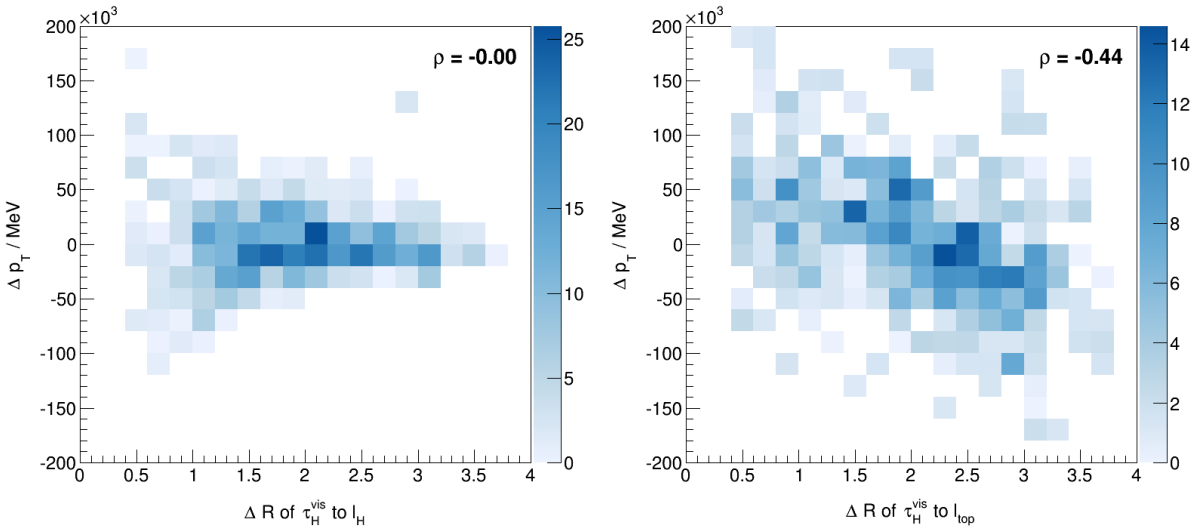


(a) (b)

Figure 5.14: Deviations of the p_{T} value reconstructed from τ_H^{vis} and lepton from the value expected by the p_{T} (ΔR) fit plotted against the ΔR between them, for lepton ℓ_H in (a) and ℓ_{top} in (b).

Nonetheless, the resulting mass distribution for the $t\bar{t}H(\tau\tau)$ $2\,\ell, 1\,\tau_{\text{had}}$ signal is compared to the ideal case as well as to the top quark pair production background in Fig. 5.14. Again the peak of the mass spectrum becomes broader, reflected in a mean value of (85.9 $\pm$ 1.6) GeV and a standard deviation of (32.5 $\pm$ 1.1) GeV. The signal to background ratio becomes $(4.56 \pm 0.25) \cdot 10^{-3}$, which is a decline from the simpler methods.

It is worth it to examine from what kind of events the large p_T deviations originate, as this may imply that in certain cases the approach is not valid. In this case it could be improved by introducing further restrictions. Figure 5.15 shows the p_T deviations against the ΔR to the τ-jet for both the correct and the wrong lepton.



(a) (b)

Figure 5.15: Deviations of the p_T value reconstructed from τ_H^{vis} and lepton from the value expected by the p_T (ΔR) fit plotted against the ΔR between them, for lepton ℓ_H in (a) and ℓ_{top} in (b).

In case the correct lepton is used, only a slight tendency for small ΔR to yield larger Δp_T is noticeable. For the wrong lepton large values of Δp_T are found at all ΔR. Introducing a minimum value for ΔR might help in excluding events for which the approach cannot be successful, but doesn't seem sensible given the limited statistics.

All things considered, the p_T (ΔR)-fit approach does not offer advantages over the simpler approaches presented before. One reason for its shortcoming may be due to neglecting the neutrinos. Possibly the approach could be improved by taking into account their transverse momenta which are so far disregarded in the reconstruction of the $H \to \tau\tau$ decay. However, this can only be done , if their contribution can be extracted from E_T^{miss}. A short study into this problem is presented in the following chapter.

5.4 Conclusion

In summary, the problem appears too complex to be solved by simple considerations. It seems that none of the intuitive expectations one could make for the characteristic event topologies is fulfilled. Before turning to more elaborate methods which aim at a full reconstruction of the event kinematics in Chapter 7, the topology of $t\bar{t}H(\tau\tau)$ events is therefore investigated on truth level.

Throughout the studies presented here multiple quantifiers to judge the invariant mass resolution were used. While it is important to consider how the mass spectrum is shaped for both signal and background events, it seems that only the rate of events in which the correct lepton is found already gives a good estimate of the quality of a method. This is important for some of the following studies employing a more complex approach, where the available statistics becomes too small for a sensible mass distribution.

CHAPTER 6

Topology of $t\bar{t}H(\tau\tau)$ $2\,\ell, 1\,\tau_{\text{had}}$ events

In the previous studies, it was found that the distances between the visible final state particles from the $H \to \tau\tau$ and $t\bar{t}$ decay are not very characteristic within each decay system. The stable particles seem to be equally spread out, resulting in events in which, for example, the lepton ℓ_H is closer to a b-jet than the lepton ℓ_{top}. There appears to be little isolation between the decay systems in the final state and hence it is difficult to distinguish between them. There are two main reasons which could cause such topologies. For one thing the $t\bar{t}$ and Higgs- systems could be produced in a such way, that there is hardly any separation between these original three particles to begin with. In this case their decay products would naturally overlap. Another possibility is that even if there is a clear separation between the top quarks and the Higgs boson, it is not conserved throughout the subsequent decays. Therefore the topologies of the different decay processes are investigated on truth level. If signal events exhibited a very characteristic topology, this might also be used for signal extraction. In the first part of this chapter the focus is on the visible final state objects.

Another important problem to address is how to handle the undetected neutrinos. In the signal channel, three of them are produced in the $H \to \tau\tau$ decay and one in the leptonic top quark decay. It is unlikely that it is still possible to assume that $E_{\text{T}}^{\text{miss}}$ can approximate the momenta of the neutrinos in the $H \to \tau\tau$ system. This, however, is necessary for most elaborate reconstruction methods of $m_{\tau\tau}$. Hence, the second part of this chapter focuses on the truth neutrinos in order to determine how feasible it is to separate the reconstructed $E_{\text{T}}^{\text{miss}}$ into the contributions from Higgs and top decay.

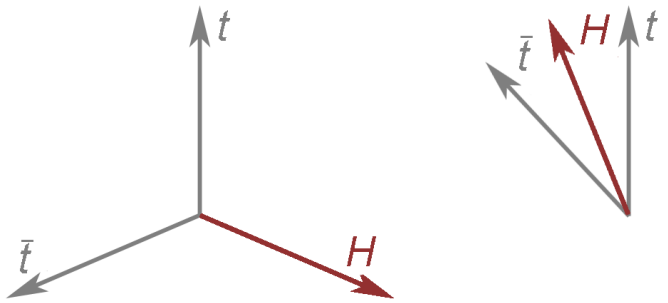
Since the studies presented in this chapter are performed on truth level, the basic event selection on reconstruction level as described in Chapter 4 is dropped. Instead a truth level selection of $t\bar{t}H(\tau\tau)$ $2\,\ell, 1\,\tau_{\text{had}}$ events is used. However it is important to note that due to the nature of the samples the requirement for two reconstructed and isolated leptons persists.

6.1 Geometric configurations of the decays

6.1.1 Initial $t\bar{t}H$ system

First, it needs to be understood whether the original top quark pair and the Higgs boson are distributed isotropically or whether they are produced in a collimated way to begin with. The sketch shown in Fig. 6.1 illustrates these two geometric configurations in two (arbitrary) dimensions. In the isotropic case the momenta of the $t\bar{t}H$ system are balanced out. The second,

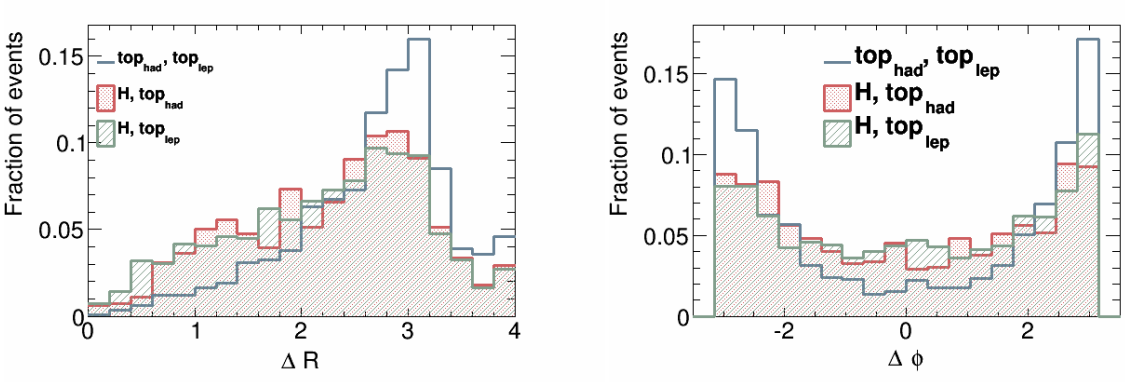
collimated case corresponds to special topologies. At the LHC, the longitudinal momentum components of the partons in the scattering reaction are not necessarily balanced out, as the partons involved may carry different fractions of the total proton momentum. The reaction may then be boosted in the forward direction along the beam axis. In a different topology, the $t\bar{t}H$ system could be boosted in a transverse direction due to the recoil of an additional jet of high momentum.



(a) (b)

Figure 6.1: Geometric properties of $t\bar{t}H$ production: In (a) the three particles are distributed isotropically, while in (b) they are collimated.

Figure 6.2.(a) compares the angular distances ΔR between the Higgs boson and the top quarks as well as top quarks to each other. It is apparent that the three distributions look very similar and all have a peak at a value around 3, which probably corresponds to the case in which the three particles are equally spaced, as in Fig. 6.1.(a). Apparently, small distances of $\Delta R < 1$ as well as large distances of $\Delta R > 3$ are less likely. This indicates that the very collimated case occurs less frequently. Figure 6.2.(b) depicts the distributions of the azimuthal angle $\Delta\phi$ between the three particles. Again, a large separation is also more likely in the transverse plane.



(a) (b)

Figure 6.2: Angular separation ΔR (a) and angle (b) between the directions of Higgs and top quarks in $t\bar{t}H(\tau\tau)$ events.

A more complex variable by which to distinguish the configurations more quantitatively is

the *sphericity*. This event shape variable is calculated as follows [41]: a momentum tensor $S^{\alpha\beta}$ is defined as

$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\vec{p}_i|^2} \tag{6.1}$$

where α and β are the three spatial momentum components x, y and z and the sum runs over all particles i. Determining the two smallest eigenvalues of the tensor $S^{\alpha\beta}$, λ_2 and λ_3, the sphericity can then be calculated as:

$$S = \frac{3}{2}(\lambda_2 + \lambda_3) \tag{6.2}$$

Hence the sphericity S can be understood as a measure of the transverse momentum sum in the event relative to an axis, which is given by the eigenvector exhibiting the largest eigenvalue and starting from the interaction point. The values of S are then in the range between 0 and 1. While a value of $S = 0$ is typical of a balanced out di-jet event, the sphericity approaches 1 for events with a very isotropic configuration. For the kind of events under investigation here, where three vectors are distributed in an equally spaced way within one plane the sphericity should be smaller than $\frac{3}{4}$. Hence, when taking into account only the top quarks and the Higgs boson to calculate the sphericity the range of values is actually only up to $\frac{3}{4}$.

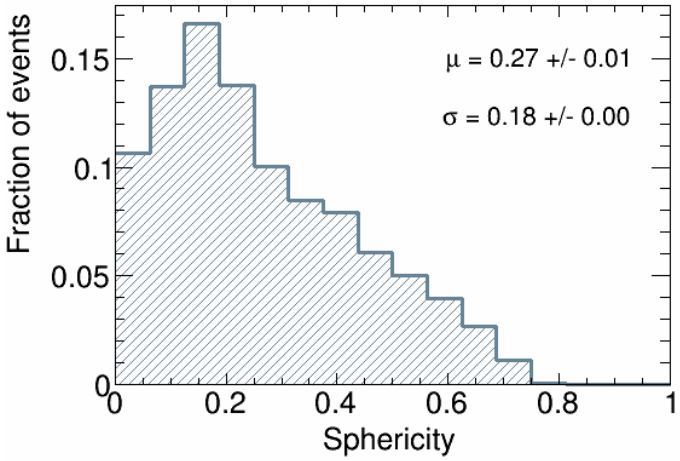
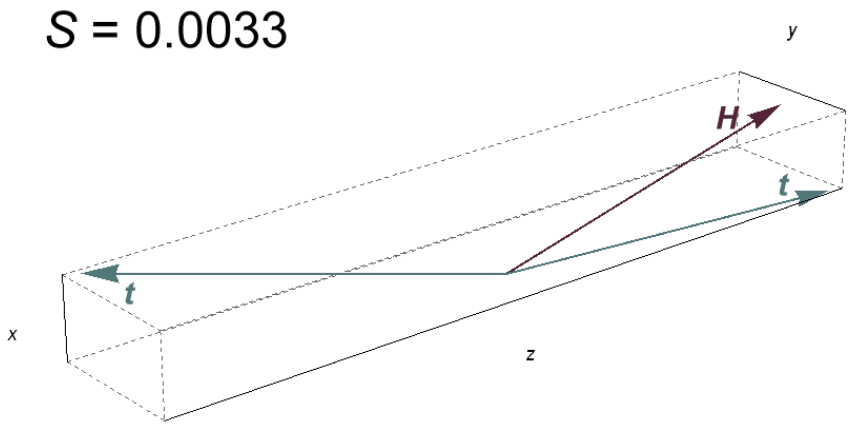
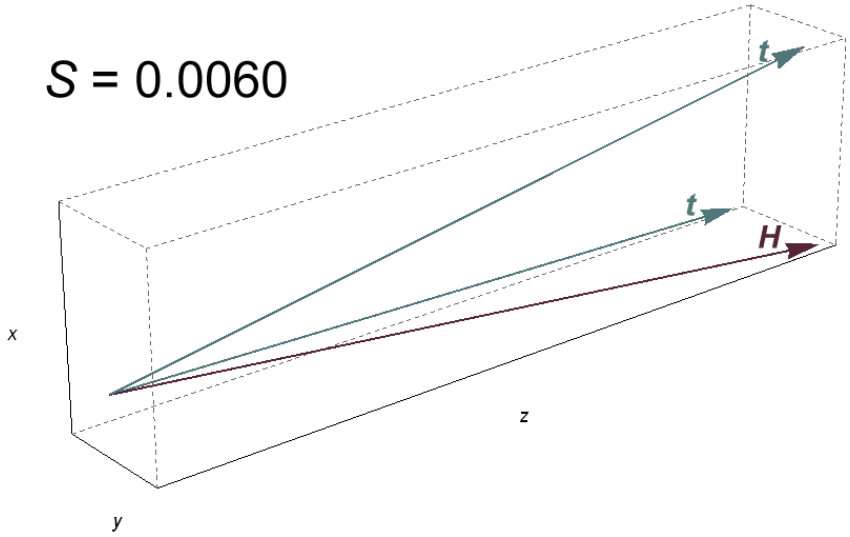


Figure 6.3: Sphericity a taking into account only the true Higgs boson and the top quarks

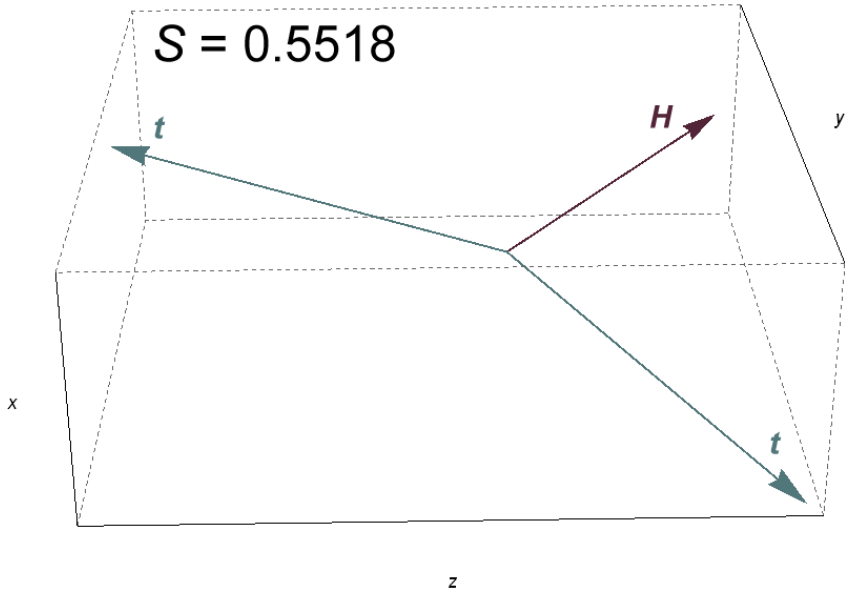
Figure 6.3 shows the resulting distribution. Evidently, most events yield a very small value of the sphericity. This is reflected by a mean value of $\mu = 0.27 \pm 0.01$ of the distribution. To take a closer look at the many events which yield sphericity values of $S \approx 0$ the particle momenta are sketched for exemplary events in Fig. 6.4. Two different different configurations appear to contribute. First, it is possible that two out of the three particles are collimated and balanced out in the longitudinal by the third one on the opposite side. This kind of event is illustrated in part (a). On the other hand, the case in which the directions of all three particles are relatively parallel also results in a sphericity value near zero. Such an event is shown in part (b), here the whole initial $t\bar{t}H$ state is boosted along the longitudinal axis.



(a)



(b)



(c)

Figure 6.4: Event shapes of the $t\bar{t}H$ initial state for different values of the sphericity

With increasing sphericity the separation between the Higgs boson and the top quarks becomes larger as the event becomes more spherically distributed. As a rough estimate based on the simple event displays, mostly events with $S \geq 0.4$ correspond to an idealized case of an equidistant particle distribution in the initial state. An example of such an event is shown in Fig. 6.4.(c). As can be seen from the sphericity distribution this is only the case for about one third of the events.

Hence, the assumption that the decay systems of top quarks and Higgs boson are sufficiently separated to begin with is not confirmed for a large fraction of events. However at least part of the events that do not fulfil this naive expectation most likely exhibit a higher jet multiplicity than those that do. At high enough statistics a cut on this property of events therefore may be used to enhance the selection of ideally shaped events.

6.1.2 Subsequent decays

Even if initially the Higgs is well separated from the top quarks, the subsequent decays of each particle may lead to overlapping stable particles of different origins. To investigate such effects, the distances between the final state particles and their respective parent particles are helpful. Since there is particular interest in the leptonic top quark decay for the studies presented so far, only this decay as well as the Higgs decay are examined in detail here, while the hadronically decaying top quark is disregarded.

Focussing on the top quark side of the $t\bar{t}H$ process to start with, the angular distance ΔR as well as angle $\Delta\phi$ between the lepton ℓ_{top} and the top quark it originates from, top_{lep}, is shown in Fig. 6.5.(a) and (b). While in both distributions a large part of events yields small values below 1, the distributions are very broad. This means that in many cases the lepton strays far away from the top quark from which it was produced. The same applies when considering the distance and angle of the lepton to its parent W_{lep}, as can be seen in parts (c) and (d) of the same figure. Looking at the separation and angle between the lepton and the bottom quark produced within the same top quark decay, as shown in (e) and (f), the tendency to large values becomes even more pronounced. Even on truth level, there seems to be no strong preference for the lepton ℓ_{top} to be close to a bottom quark and hence the jet reconstructed from the latter. However it is obvious that the regions of small separation in ΔR and in angle $\Delta\psi$ between the lepton and the bottom quark are depleted. This is probably due to the persisting requirement for events to exhibit two isolated leptons and the overlap removal practice used to select such leptons. In case a muon and jet overlap within the defined separation the muon is removed. While this is performed using reconstructed objects, small distances on truth level between lepton and bottom quark also mean small distances between this lepton and the corresponding jet. If then the muon is removed out of such events, the events do not fulfil the requirement for exactly two leptons any more and are therefore not considered here.

Turning to the $H \to \tau\tau$ system we are interested in both the leptonically and hadronically decaying tau, as the latter is also one of the selection criteria for this channel. Figure 6.6.(a) and (b) display the distance ΔR and angle $\Delta\phi$ between the lepton ℓ_H and the full, true τ lepton from which it originates. As opposed to the lepton from the top quark decay, the separation of intermediate and final state particle is very small. In the hadronic τ decay represented in part (c) and (d) of the figure the situation is similar. Here the visible τ-jet is found in close

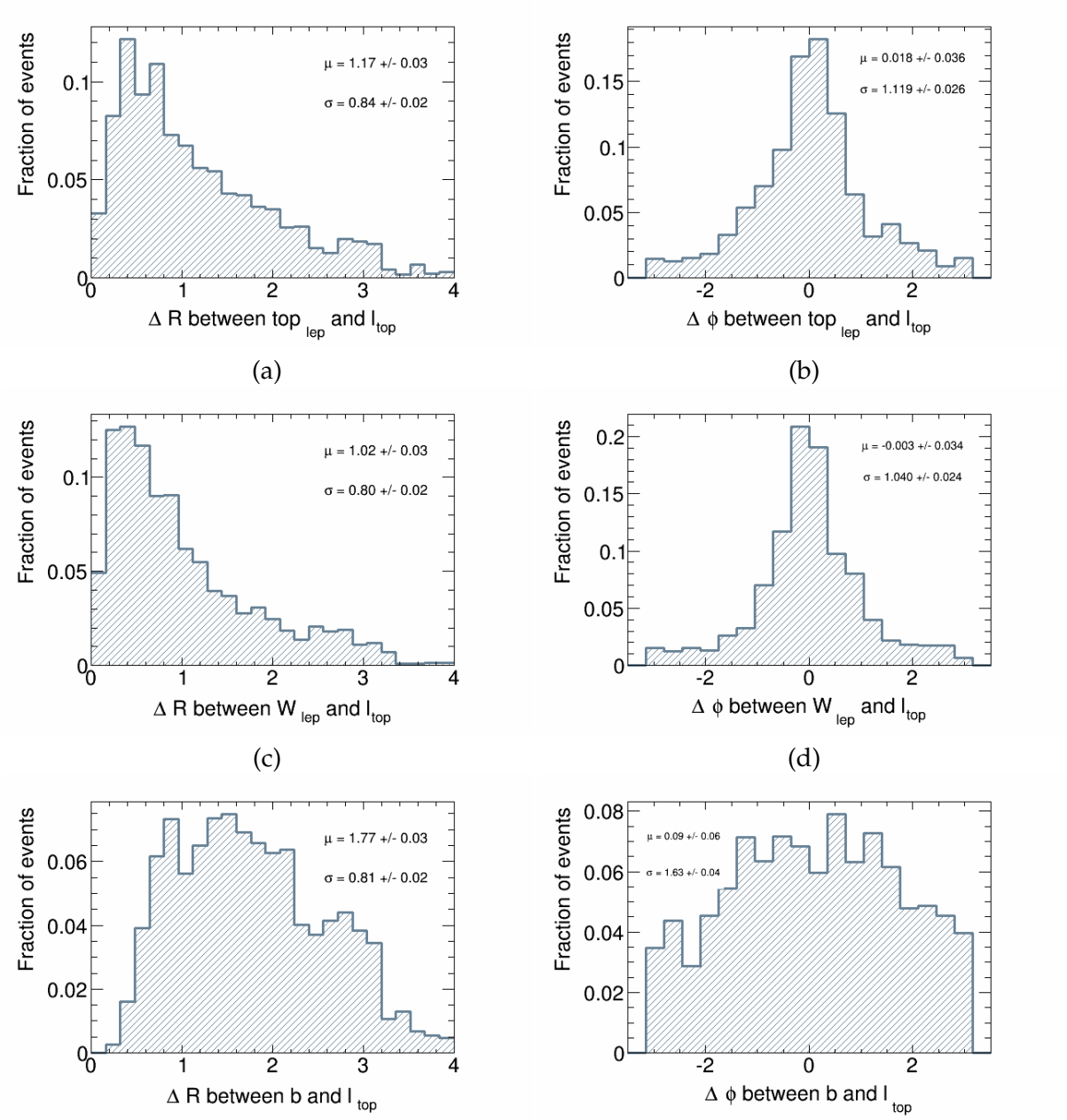


Figure 6.5: Left hand side: Angular separation ΔR of the lepton produced in a top decay, ℓ_{top}, to its parent top quark and W boson as well the bottom quark produced along with the W_{lep}. Right hand side: The angle $\Delta\phi$ between those particles.

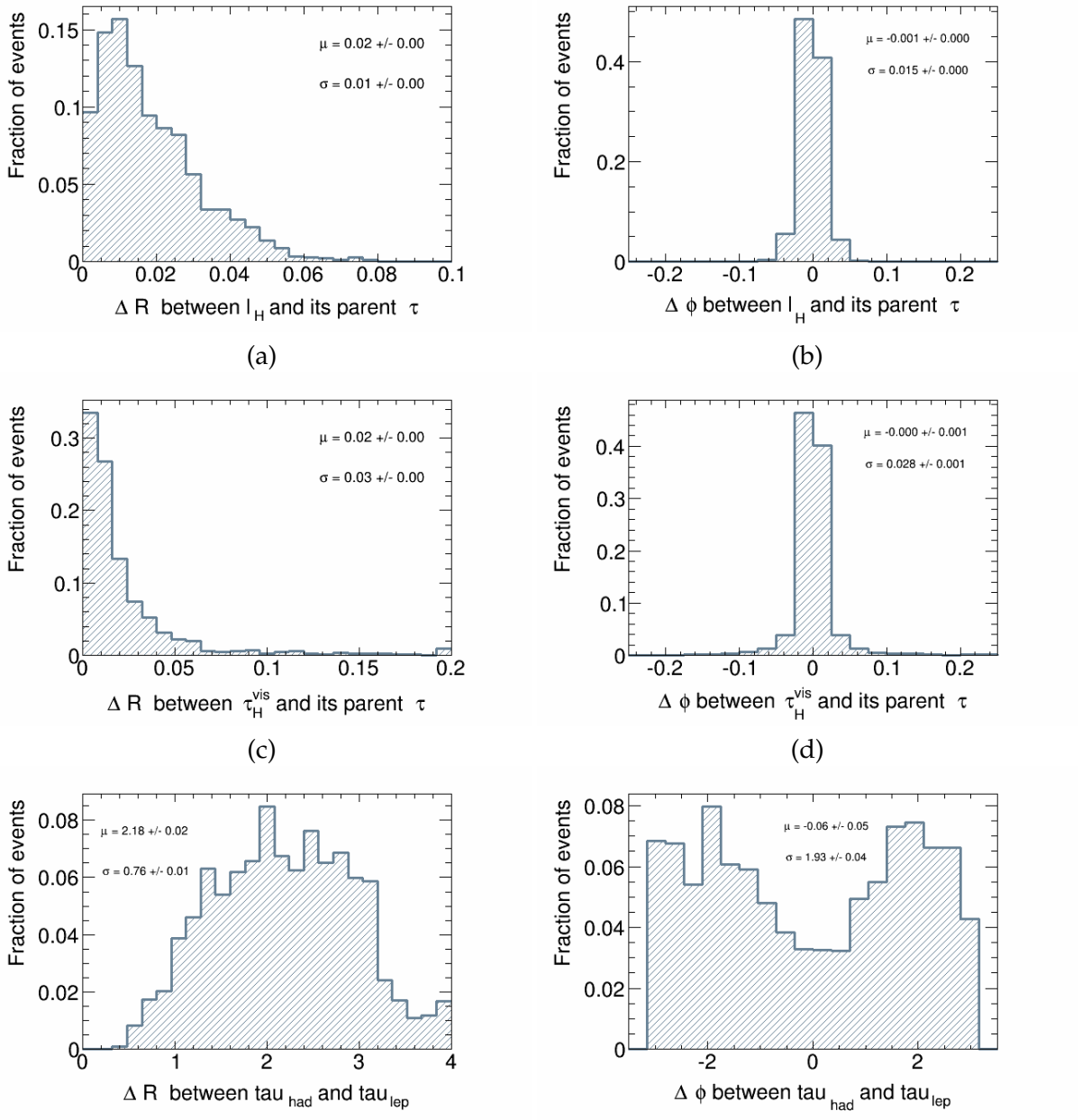


Figure 6.6: Left hand side: Angular separation ΔR of the lepton produced in the Higgs decay, ℓ_H, to its true parent τ as well as the same for the visible hadronic tau, τ_H^{vis}. Additionally the distance between the two τ leptons is shown. Right hand side: The difference in three dimensional angle $\Delta\phi$ between those particles.

proximity to its true parent. Therefore, it appears that the decaying τ leptons are never at rest, but boosted, so that their decay products are collinear to the original τ-direction. On the other hand, this is not the case for the $H \to \tau\tau$ decay as can be seen in parts (e) and (f) of Fig. 6.6. The separation and angle between the two true τ leptons is much larger, and therefore there also appears to be no preference for the final state particles ℓ_H and τ_H^{vis} to be close to each other.

To summarize, it has become apparent that the decay products of the leptonic top decay are typically not collimated with respect to the original top quark direction. Especially the b quark and lepton ℓ_{top} are often found with large separation between them. The same is true for the two τ leptons from the Higgs decay and therefore the lepton ℓ_H also is not necessarily found close to the τ-jet τ_H^{vis}. However, the taus themselves are sufficiently boosted, so that their own visible decay products do not stray far from the original tau direction. In fact, ℓ_H and τ_H^{vis} are much more collinear to the respective intermediate taus than ℓ_{top} is to its parent W boson. Considering the masses of the particles involved, it is not surprising that the τ leptons from the Higgs decay are produced with higher momenta than the W boson and bottom quark from a top decay. In the first case more phase space is available. The apparent collinearity of the τ decay products may be of value for applying more sophisticated mass reconstruction methods such as the collinear mass, which relies on this assumption. Nonetheless no strong case can be made for associating the different final state particles with their respective decay systems by only considering their geometric configuration. The final state particles are most likely not found within isolated, non-overlapping cones around the directions of the initial state particles.

6.1.3 Final state

As a last consideration, the event shape variables sphericity and aplanarity are also calculated for the true, visible final state and compared to the initial $t\bar{t}H$ system. The particles taken into account as the true final state are the visible decay products of the Higgs, i.e. the light lepton ℓ_H and the τ-jet τ_H^{vis}, and those of the top quark decays, i.e. the lepton ℓ_{top}, the two b quarks as well as the two light quarks from the hadronic top decay which exhibit the highest transverse momentum. This is done to exclude additional final state radiation, which is common in the hadronic top decay. The resulting sphericity distribution is depicted in Fig. 6.7.

When looking only at the initial $t\bar{t}H$ system, small sphericity values near 0 were very common. This is not the case here. Instead, the spatial distribution of final state particles is more likely to be spherical.

Lastly, the sphericity of the final state is plotted against the sphericity of only the $t\bar{t}H$ system in Fig. 6.8 to assess how detrimental the multiple subsequent decays are to a good initial separation. While especially for small values the sphericity on initial and final state level is often similar, the case that the latter is much larger also occurs frequently. This is exactly the situation in which the top quarks and Higgs boson are produced in an isotropic, easily separable way but this separation is then lost due to the decays which follow.

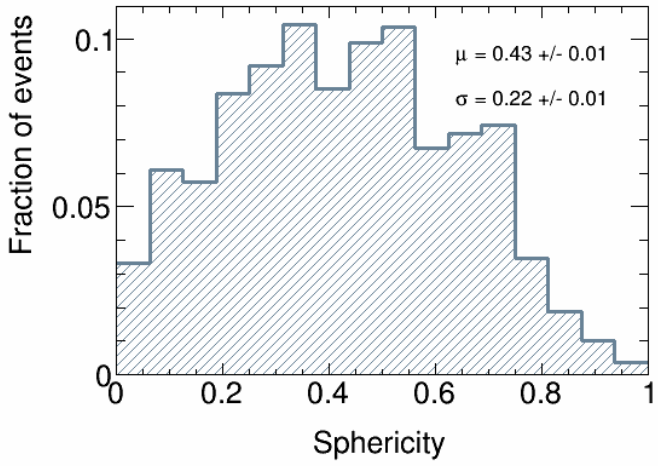


Figure 6.7: Sphericity taking into account all true final state particles of the $t\bar{t}H$ $2\,\ell, 1\,\tau_{\text{had}}$ channel.

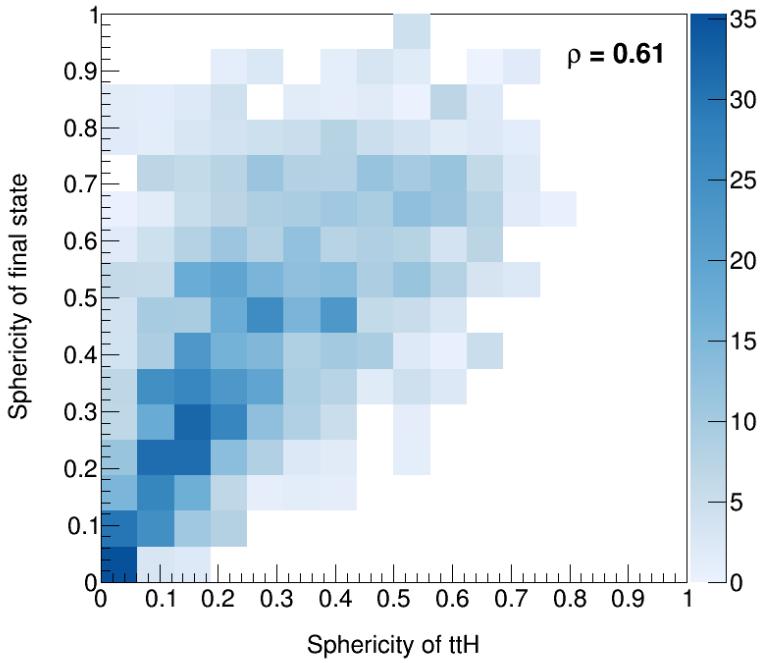


Figure 6.8: Distribution of the sphericity calculated from the visible final state particles plotted against the sphericity of the initial $t\bar{t}H$ system.

6.1.4 Summary

All in all, two effects contribute to event topologies, in which the final state particles cannot be associated to a Higgs or top quark decay by simple geometric considerations. Firstly, in many

cases the Higgs boson and top quarks are already produced in a way that either all three or two of them are parallel and will therefore overlap in their decays. This contradicts the intuitive expectation of a preferred production in balance. However the percentage of events that fulfil this expectation may very well be increased by limiting the jet multiplicity, as this would at least exclude events in which the $t\bar{t}H$ system is boosted in the transverse plane due the recoil of additional jets. On the other hand, both the $t\bar{t}$ and Higgs system undergo multiple decays which are not necessarily boosted enough for their decay products to stay close to the original particle. This is especially true for the top quark decay, in which the lepton and bottom quark can be found with large separation, as well as the intermediate $\tau\tau$ system from the Higgs decay. Only the final decay products of the two taus stay collinear to their parent particles. Nonetheless this leads to a spreading out of the final state particles, so that they cannot usually be found in narrow, non-overlapping cones around the initial $t\bar{t}H$ directions.

To illustrate these two situations again, an exemplary event for each of them is depicted in Fig. 6.9. On the left hand side, an overview of the full decay system is shown, while only the visible final state particles are displayed on the right hand side. In both events the visible part of the hadronically decaying τ is found closer to the lepton ℓ_{top} than to ℓ_H, which is produced in the same decay.

6.2 Composition of $E_{\text{T}}^{\text{miss}}$

Another big challenge of $t\bar{t}H(\tau\tau)$ events and especially the $2\,\ell, 1\,\tau_{\text{had}}$ channel is finding possibilities to use the reconstructed missing transverse energy $E_{\text{T}}^{\text{miss}}$ to approximate the neutrino momenta. As both the leptonic top quark decay as well as the τ decays involve neutrinos, $E_{\text{T}}^{\text{miss}}$ commonly reflects the vectorial transverse momentum sum of neutrinos from both systems. Turning to more complex methods which reconstruct the full event kinematics, this poses a significant problem as they rely on at least a rough approximation of the neutrino momenta being available. Basically, only knowing the full $E_{\text{T}}^{\text{miss}}$ is an ambiguity similar to trying to associate the two leptons into the Higgs or top system, as it is not known how much the neutrinos of each system contribute in terms of transverse momenta. Neither is information on their individual directions available. Moreover, a future application of more sophisticated methods to reconstruct the invariant Higgs mass is virtually impossible without knowing the $E_{\text{T}}^{\text{miss}}$ contribution of only the neutrinos from the Higgs decay. For example, the successful MMC algorithm is currently based on the assumption that $E_{\text{T}}^{\text{miss}}$ is caused solely by the $H \to \tau\tau$ neutrinos. An assessment of how the Higgs and top decays contribute to the reconstructed $E_{\text{T}}^{\text{miss}}$ is therefore necessary. For this purpose, the following studies focus on the truth level neutrinos. Since events with very low missing transverse energy are generally not well reconstructed, an additional cut of $E_{\text{T}}^{\text{miss}} > 15\,\text{GeV}$ is introduced. Again, events in which a top quark decays into light leptons via an intermediate τ are excluded, since they would only complicate the neutrino situation even further.

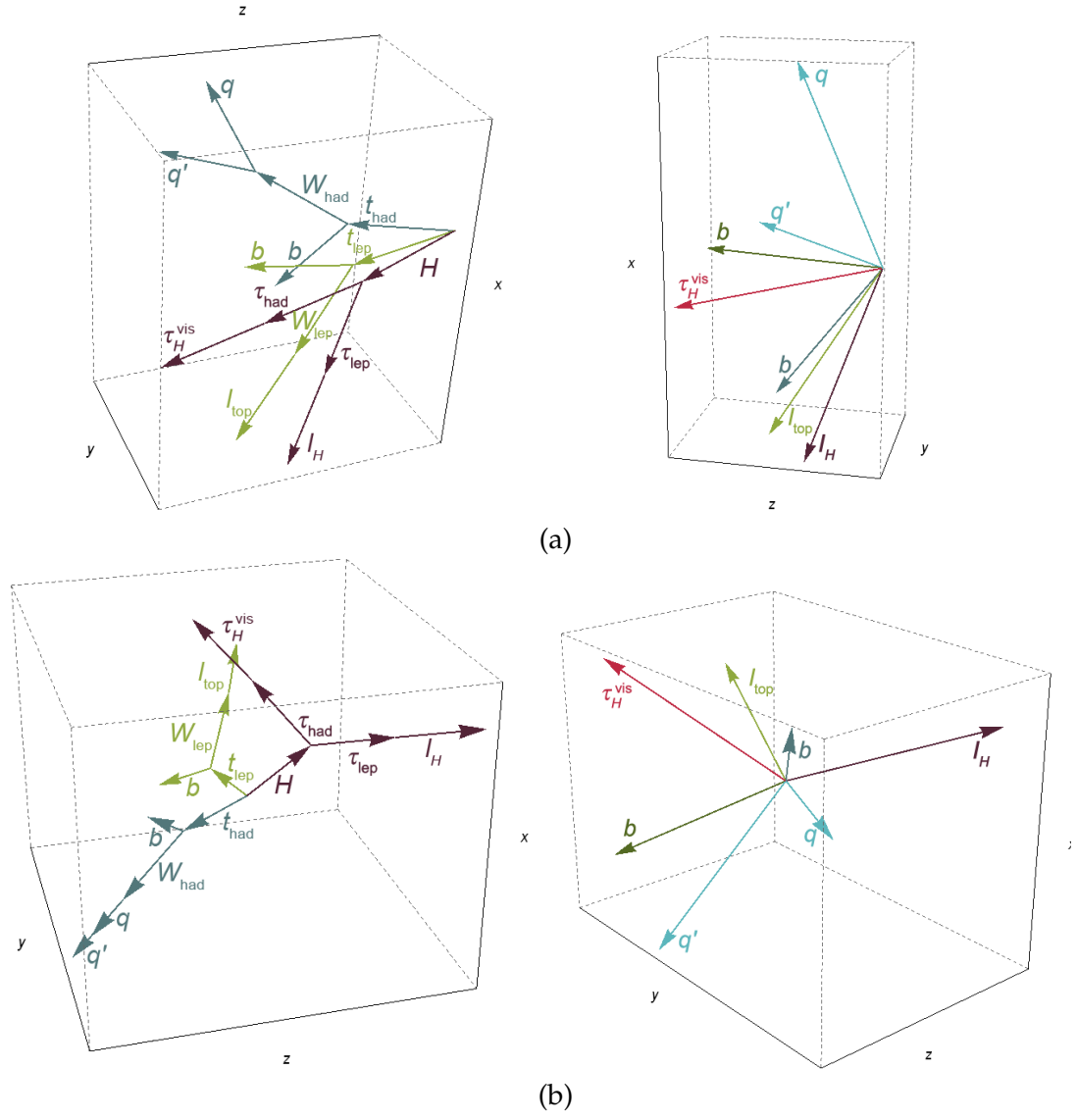


Figure 6.9: Full decay chain overview on the left and only final state particles on the right hand side for exemplary events. In the event shown in (a) the initial top quarks and Higgs-boson are produced close to each other, while in (b) they are well separated but this separation is lost due to the decays.

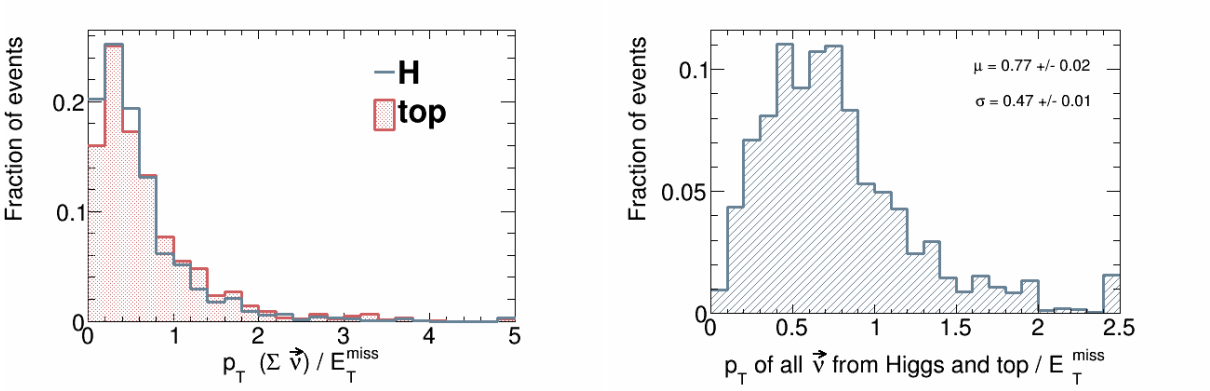
6.2.1 Relative contributions to $E_{\mathrm{T}}^{\mathrm{miss}}$ from Higgs and top quark decay

As a first step, the extent of the problem is investigated. For example, it may be possible that the one single neutrino from the top decay is usually of low transverse momentum and contributes little to $E_{\mathrm{T}}^{\mathrm{miss}}$. Then the full $E_{\mathrm{T}}^{\mathrm{miss}}$ could still be used to reconstruct only the $H \to \tau\tau$ decay. Unfortunately, this is not the case as can be seen in Fig. 6.10.(a). The distributions of the p_{T} of the vectorial sum of all neutrinos from Higgs and top system are depicted, and they have essentially the same shape. A reason for this may be the vectorial nature of the missing transverse energy. If the two taus are produced in opposing directions and the neutrinos stay collinear to the original *tau* direction, the momenta of the latter cancel each other at least

partially. This is then reflected in a smaller contribution to $E_{\mathrm{T}}^{\mathrm{miss}}$. For the same reason the relative contribution of a decay system can take values larger than 1.

As a check-up part (b) of Fig. 6.10 shows the p_{T} determined by summing up the momentum vectors of all four expected neutrinos divided by the value of the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$. It was found that employing a truth selection of events, the number of neutrinos per event is always four. No additional neutrinos from other processes are considered in these studies. Hence, if the reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$ was perfect, the fraction of p_{T} sum of neutrinos over $E_{\mathrm{T}}^{\mathrm{miss}}$ should peak exactly at the value of 1. However, it is shifted to a lower value due to the limited resolution.

In any case, it is not valid to assume that $E_{\mathrm{T}}^{\mathrm{miss}}$ mostly describes the neutrinos produced in the Higgs decay. The contribution from the neutrino produced in the lepton top quark decay is generally too large to be neglected.



(a) (b)

Figure 6.10: Comparison of the relative contributions of neutrinos from Higgs and top quark decay to reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ in (a): The momentum vectors of all neutrinos of each decay are summed up and divided by the value of $E_{\mathrm{T}}^{\mathrm{miss}}$. In (b) the p_{T} of the momentum vector sum of both Higgs and top neutrinos is divided by $E_{\mathrm{T}}^{\mathrm{miss}}$ as a check-up.

Nonetheless, only events of certain properties may yet fulfil this expectation, if the contribution of the neutrinos from the Higgs decay is enhanced in some way. To investigate whether such events exist, the relative contribution of the neutrinos from the Higgs decay to $E_{\mathrm{T}}^{\mathrm{miss}}$ is plotted against that of the top quark decay in Fig. 6.11.

Evidently for most events at small relative contributions below 1, the two values are similar. No clear region, in which the contribution from the Higgs system becomes large and that of the top system becomes small, can be found. Thus it seems that the problem cannot simply be solved by an appropriate selection of events. Instead methods to somehow extract the different contributions to $E_{\mathrm{T}}^{\mathrm{miss}}$ need to be developed.

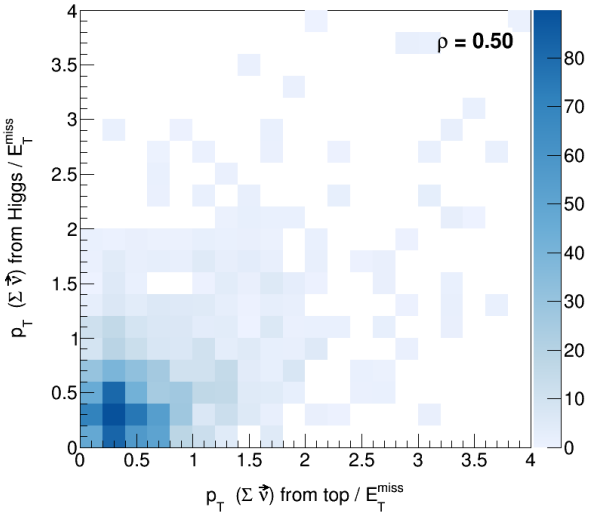


Figure 6.11: Correlation of relative contributions to $E_{\mathrm{T}}^{\mathrm{miss}}$ from Higgs and top quark decay

6.2.2 Separating $E_{\mathrm{T}}^{\mathrm{miss}}$ into Higgs and top quark contribution

It was seen in the previous studies that the visible decay products of the two τ decays from the Higgs system stay collinear to their parent particles. The three neutrinos of these decays are also collinear to the visible final state particles and therefore also to the two intermediate taus, as can be seen in Fig. 6.12. In the top row the angular distance ΔR and the difference in angle $\Delta\phi$ between the light lepton and the neutrino of the same flavour produced along with it from the leptonic τ decay is shown, while in the middle row the same is shown but for the τ neutrino from this decay. Finally, in the bottom row, the separations between the visible τ-jet and its τ neutrino are depicted. All of these distributions tend to very small values, near 0.

Therefore, it may be possible to extract the contribution of the three Higgs neutrinos to $E_{\mathrm{T}}^{\mathrm{miss}}$ by exploiting this. Their resulting momentum vector could be approximated by building the visible Higgs out of the momenta of ℓ_H and τ_H^{vis} and projecting the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ onto it. The approximated $E_{\mathrm{T}}^{\mathrm{miss},H}$ then takes the magnitude of the projected vector and the direction of the visible Higgs. The remaining vector is assumed to describe the neutrino from the leptonic top quark decay. The approach is sketched using the momenta vectors from an exemplary event in Fig. 6.13.

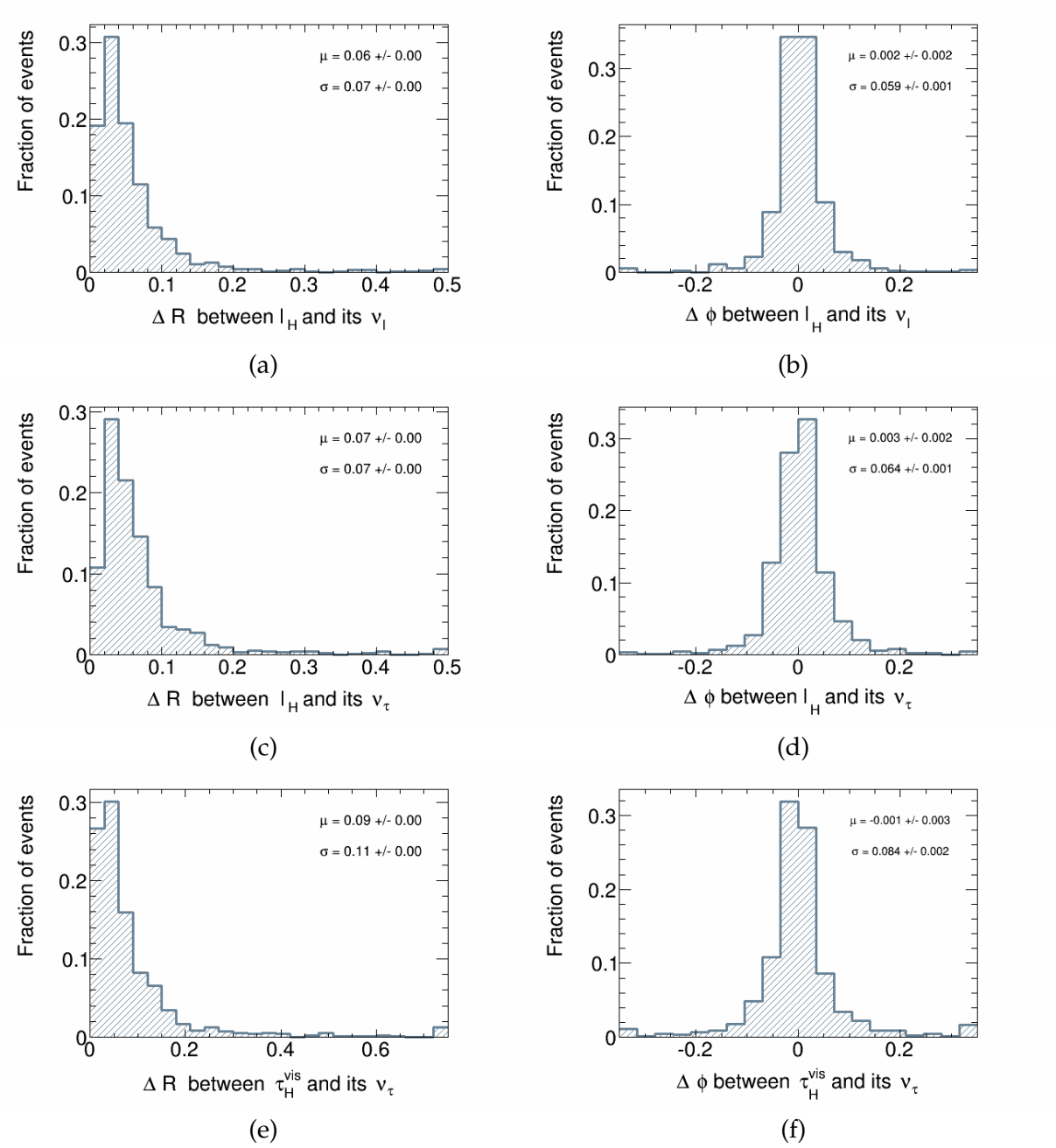


Figure 6.12: Angular distance ΔR and as well as angle $\Delta\phi$ between the neutrinos of the Higgs system and the visible final state particles along with which they are produced: The top row, (a) and (b), depicts the distance between the light lepton and its neutrino of the same flavour while (c) and (e) show the distance between this lepton and the τ neutrino of this leptonic τ decay. In (e) and (f) the distance between visible τ-jet and τ neutrino in the hadronic τ decay is shown.

The resulting distribution of the magnitude of $E_{\mathrm{T}}^{\mathrm{miss},H}$ is compared to the true value calculated from the neutrino momenta in Fig. 6.14.(a) . Part (b) of the figure depicts the angle between the true direction of $E_{\mathrm{T}}^{\mathrm{miss},H}$ and the direction obtained from the projection method.

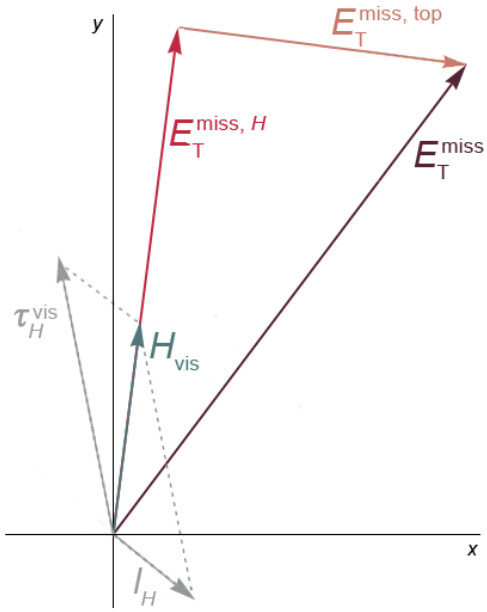


Figure 6.13: Sketch of the projection method for E_T^{miss}: The contribution of the neutrinos from the Higgs decay, $E_T^{miss,H}$, is estimated by projecting E_T^{miss} onto the visible Higgs direction, composed of ℓ_H and τ_H^{vis}. The remainder is then assumed to be $E_T^{miss,\,top}$

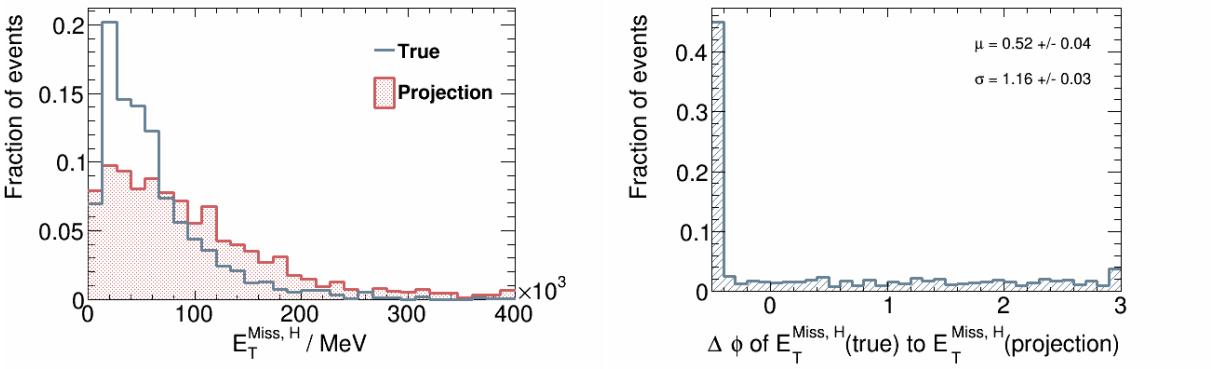


Figure 6.14: Comparison of the value and direction of $E_T^{miss,H}$ to the true value as well as the angle .

Obviously the shape of the distribution of $E_T^{miss,H}$ values is not very well reconstructed by the projection method. When calculated from the truth neutrinos it shows a characteristic peak at low values, which is not so pronounced in the projection. Moreover, the latter distribution is slightly shifted to higher values in general. Thus using this method the $E_T^{miss,H}$ value is likely often overestimated. Yet how detrimental this is for the quality of analysis results, may depend on the practical application. While the angle $\Delta\phi$ in between the true and estimated direction of $E_T^{miss,H}$ is mostly near 0, its distribution extends to high values as well. Therefore, it is not necessarily well estimated using this approach.

6.3 Conclusion

To summarize, it was found that the full topology of $t\bar{t}H(\tau\tau)$ processes is very complex. Firstly, a good geometric separation between the decay systems of the Higgs boson and top quarks is not guaranteed. This is both due to the assumption of a balanced out production of the $t\bar{t}H$ at rest not being fulfilled, as well as caused by the multiple subsequent decays. Moreover, the undetected neutrinos produced in the $H \to \tau\tau$ and the leptonic top quark decay contribute equally much to the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$. It is unlikely that a simple geometric solution is sufficient to disentangle them. However, it may be possible to solve the topological complications by restricting the considerations. For example, if the decay systems are boosted in the transverse directions, their decay products move closer together and the simple geometric expectations should be true for a larger fraction of events. Such special topologies may for example be selected by their higher jet multiplicity. However, the overlap of the systems caused by a boosted topology in the forward direction would be much harder to account for. In any case, the data samples used for this thesis did not provide enough statistics to put such considerations to test.

CHAPTER 7

Kinematic event reconstruction

It was found that the topology of $t\bar{t}H(\tau\tau)$ events is too complex for simple considerations to be successful at associating the final state particles to their decay systems. Hence, a more elaborate approach, namely fully reconstructing the $t\bar{t}$ kinematics, is tested in this chapter. This is provided by KLFitter. As a brief outlook, the possibility of a new method, more specialized for the $t\bar{t}H(\tau\tau)$ application is investigated.

7.1 The KLFitter approach

The **K**inematic **L**ikelihood **Fitter** is a framework, which implements a 'likelihood-based reconstruction algorithm for arbitrary event topologies' [42]. Its objective is to model the physics process of the hard-scattering reaction in the event as accurately as possible by associating the reconstructed objects correctly with the true particles expected in the final state. For this purpose, it makes use of the kinematic information of the event by exploiting any available constraints in this model - such as the mass of a decaying particle. Additionally detector effects are taken into account. This elaborate approach facilitates a more precise analysis due to the suppression of combinatorial backgrounds. It has been well tested as well as employed successfully in the analyses of top-quark pair production (e.g. [43]).

Addressing the problem of finding the charged lepton which belongs to the Higgs decay, the possibilities and limitations of an application of the KLFitter approach for the single lepton top quark pair decay to the selected final state are studied in this chapter. As an introduction a summary of this approach and its results on top quark pair Monte Carlo samples is given (cf. [42]).

7.2 Likelihood for semileptonic $t\bar{t}$ decays

The production of a $t\bar{t}$ pair which subsequently decays semileptonically is depicted in Fig. 7.1.

Since in this case only one charged lepton is present, the only ambiguity in reconstructing such final states lies in the association of the reconstructed jets to the true particle jets, which are created by the hadronization of the partons. Ideally, in the leading-order picture, a single lepton $t\bar{t}$-event would therefore contain exactly four reconstructed jets: one due to the bottom quark from the leptonically decaying top quark, a second one due to the bottom quark from the hadronically decaying top quark and two more from the hadronically decaying W boson. Additionally, in the ideal case the former two jets would be b-tagged. There are, however,

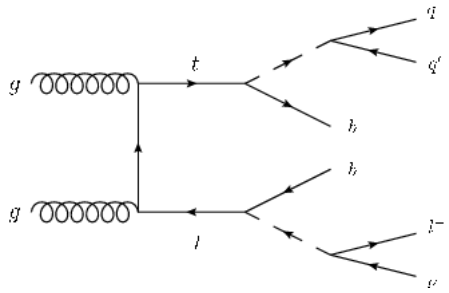


Figure 7.1: Feynman diagram describing the production and subsequent decay into a charged lepton, neutrino and jets of a top-quark pair

multiple reasons why an event might not fulfil these expectations and more or less jets than those four are detected. The latter case may occur due to jet merging in which a pair of jets that exhibit little separation from each other (e.g. due to a boosted topology or a very high jet multiplicity, for more details see e.g. [44]) is wrongly reconstructed as one single jet. Apart from this, the experimental conditions cannot grant an ideal reconstruction of all the expected objects: the acceptance of the detector is limited, making it possible for jets to escape detection. Moreover jets can also be misidentified as hadronically decaying τ leptons (or vice versa). On the other hand, the number of jets detected can be increased due to additional radiation in either the initial or final state. Additionally, extra jets can be added to the actual scattering process due to pile-up effects. Finally, another factor to take into account is the limited accuracy and efficiency of the jet reconstruction algorithm employed.

Here, in principle the number of possibilities to match jets to truth partons, referred to as *permutations*, is given by $N_{\text{perm}} = 4! = 24$, when taking into account exactly four jets. However the KLFitter approach is invariant under a commutation of the two jets in the hadronically decaying W boson, and hence this number is reduced by a factor $\frac{1}{2}$ to $N_{\text{perm}} = 12$. For events with $N_{\text{jet}} > 4$ of course extra permutations are necessary - all possible subsets of exactly four jets can then be considered in the fit. Making use of the fact that jets which are b-tagged most likely don't originate from the hadronically decaying W boson any permutation with such an association can be excluded. This reduces the number of possible permutations to $N_{\text{perm}}^{1\text{ btag}} = 6$ and $N_{\text{perm}}^{2\text{ btag}} = 2$ in the case of one or two b-tagged jets respectively. Events with more than two b-tagged jets are rejected as not consistent with the model.

Introducing the notation that $t_{\text{had}} \rightarrow W_{\text{had}} + q_3 \rightarrow q_1 + q_2 + q_3$ and $t_{\text{lep}} \rightarrow W_{\text{lep}} + q_4 \rightarrow l + \nu + q_4$ the likelihood is parametrized in the following way:

$$\begin{aligned} \mathcal{L} &= B\left(m_{q_1q_2q_3}|m_{top},\Gamma_{top}\right)\cdot\exp\left(-4\cdot\ln 2\cdot\frac{\left(m_{q_1q_2}-m_W\right)^2}{\Gamma_W^2}\right) \\ &\cdot\ B\left(m_{q_4l\nu}|m_{top},\Gamma_{top}\right)\cdot B\left(m_{l\nu}|m_W,\Gamma_W\right) \\ &\cdot\ \prod_{i=1}^{4} W_{\text{jet}}\left(E_{\text{jet},i}^{\text{meas}}|E_{\text{jet},i}^{\text{true}}\right)\cdot W_l\left(E_l^{\text{meas}}|E_l^{\text{true}}\right)\cdot W_{\text{miss}}\left(E_{T,x}^{\text{miss}}|p_x^{\nu}\right)\cdot W_{\text{miss}}\left(E_{T,y}^{\text{miss}}|p_y^{\nu}\right) \end{aligned} \tag{7.1}$$

Here, Breit-Wigner distributions, B, are used to describe the resonances of both the hadronic and leptonic top-quark decay as well as for the W boson to charged lepton and neutrino. In this way the invariant masses of the combined objects $m_{q_1q_2q_3}$ and $m_{q_4l\nu}$ as well as $m_{l\nu}$ are constrained to the mass of the top-quark $m_{top} = 172.5\,\text{GeV}$ and that of the W-boson $m_W = 80.4\,\text{GeV}$ using the respective decay widths of $\Gamma_{top} = 1.5\,\text{GeV}$ and $\Gamma_W = 2.1\,\text{GeV}$. The further away the invariant mass of a particular combination of objects is from the peak of the given distribution the less likely it is the true combination. The hadronic decay of the W boson to two (light) quarks q_1 and q_2 is approximated by a Gaussian as given in the second factor of Eq. 7.1. This choice can be motivated by the finite resolution of jet energies: especially compared to measurements of lepton energy, jet energies are measured with less precision. Hence, especially when combining only two jets to the hadronically decaying W boson, it is important to prevent the acceptance of fluctuations to very large invariant masses. For this purpose a Gaussian is better suited than a Breit-Wigner distribution, as the latter exhibits a tail to large values. It was found empirically that this combination of Gaussian and Breit-Wigner distributions yields the best results in reconstruction efficiency of the event [42].

Considering particle resonances as Breit-Wigner functions is only valid in when neglecting the finite detector resolution. To take these effects into account so-called *transfer functions* W_i are used to map the measured and true quantities onto each other. For example, the transfer function $W_l\left(E_l^{\text{meas}}|E_l^{\text{true}}\right)$ gives the probability to measure the energy E_l^{meas} for a lepton whose true energy is E_l^{true} (same for the jets). In the case of muons the transfer function depends on their transverse momentum instead of energy. In order to make use of the relation between the true neutrino momentum p^ν and the missing transverse energy $E_{\text{T}}^{\text{miss}}$ special transfer functions W_{miss} are used in the transverse components x and y. All of the transferfunctions are derived by fitting the difference between true and measured value with either a single or double Gauss-function from Monte Carlo simulation samples in different regions of η. More details can be found in [45].

The procedure of kinematic fitting maximizes the (negative logarithm of the) likelihood given above for each permutation by varying the jet energies, the electron energy or the muon momentum as well the transverse neutrino momentum within the realm of the detector resolution as modelled by the set of transfer functions. By doing so, it is ensured that slightly mismeasured particles can still be associated correctly. Since no measurement of the longitudinal momentum component of the neutrino p_z^ν is available, it is left as a free parameter in the fit. It is important to note that within this approach the angles of the reconstructed objects are assumed to be measured precisely with negligible uncertainties and are therefore left unchanged. The permutation which yields the highest likelihood is then chosen for the event reconstruction.

In [42] performance studies applying this ansatz on MC simulated top quark pair production events at a center of mass energy $\sqrt{s} = 7\,\text{TeV}$ are presented. It is important to note that the efficiencies cited in the following are achieved using only *truth matched events* - in these kind of events all four jets expected on truth level are present as reconstructed jets and can be matched unambiguously to their truth counterparts within $\Delta R < 0.3$. Defining the *reconstruction efficiency* as the percentage of events, for which all four jets have been identified correctly, and using KLFitter as described above the reconstruction efficiency is shown to reach $(83.3 \pm 0.3)\%$. Hence in conclusion for this application the likelihood based kinematic fit implemented in

KLFitter has been proven successful in its suppression of combinatorial backgrounds.

7.2.1 Applying KLFitter to $t\bar{t}H(\tau\tau)$ events

Concept

Since it works so well for reconstructing top quark topologies, the idea is to apply this more complex, multivariate concept to $t\bar{t}H(\tau\tau)$- $2\,\ell, 1\,\tau_{\text{had}}$ events in order to use its strength at reconstructing the decay topology of top quark pairs to identify which lepton belongs into the leptonic top decay. Obviously, this final state contains more objects than just the $t\bar{t}$ decay considered above, which ideally would all be considered in a likelihood modelling the full $t\bar{t}H(\tau\tau)$ decay. However to do so, it would be necessary to consider the Higgs mass as a constraint and since this study investigates the potential of KLFitter as a tool to improve the reconstruction of the invariant $H(\tau\tau)$ mass proceeding in this way would be redundant. Noting that the principles of the fit are based on the energies and transverse momenta of the particles and not their directions, it seems natural to assume that the additional presence of the Higgs boson in $t\bar{t}H(\tau\tau)$ should have no impact on the top quark pair reconstruction in this respect since a potential overlapping of the $t\bar{t}$ and H decay products should not cause problems. Moreover, the KLFitter method has been proven to work for top quark pair events with additional jets, which imbalance the $t\bar{t}$ decay in the transverse plane much like the presence of the Higgs would. Thus, sufficient motivation to apply the KLFitter likelihood for the semileptonic top quark pair decay as given in Eq. 7.1 to this final state and simply excluding the Higgs system is given. Employing this likelihood the kinematic fit can be performed for each lepton separately - introducing an additional permutation to each event - and the lepton yielding the greater likelihood is chosen to belong to the $t\bar{t}$- decay, whilst the other one can then be identified as ℓ_H by exclusion.

Estimating the contribution to $E_{\text{T}}^{\text{miss}}$ by the neutrino of the leptonic top decay There is, however, another problem with this approach that needs to be addressed beforehand: in the likelihood Eq. 7.1 it is assumed that $E_{\text{T}}^{\text{miss}}$ is (within the detector resolution) entirely caused by the neutrino from the leptonically decaying top quark, which escapes detection. But looking at the $2\,\ell, 1\,\tau_{\text{had}}$ channel in $t\bar{t}H(\tau\tau)$ there is also a contribution to $E_{\text{T}}^{\text{miss}}$ by the three neutrinos in the Higgs decay and therefore the full reconstructed $E_{\text{T}}^{\text{miss}}$ needs to be separated into two parts beforehand. Here this is done by employing the projection of $E_{\text{T}}^{\text{miss}}$ onto the visible Higgs momentum vector, which was motivated and studied before in Chapter 6. In order to study the applicability of a KLFitter approach on data events this is used as a starting point. The distributions of the momentum components for the top system calculated in this way are compared to their true value in Fig. 7.2 and their similarity motivates the choice to base the following study on this approach.

It is important to note that the width of the transferfunctions for the $E_{\text{T}}^{\text{miss}}$ in the KLFitter likelihood is dependent on the scalar $\sum E_{\text{T}}$ of the event. Here, the $\sum E_{\text{T}}$ is approximated by the scalar difference of reconstructed $\sum E_{\text{T}}$ and the magnitude of the remaining vector after projecting $E_{\text{T}}^{\text{miss}}$ onto the visible Higgs.

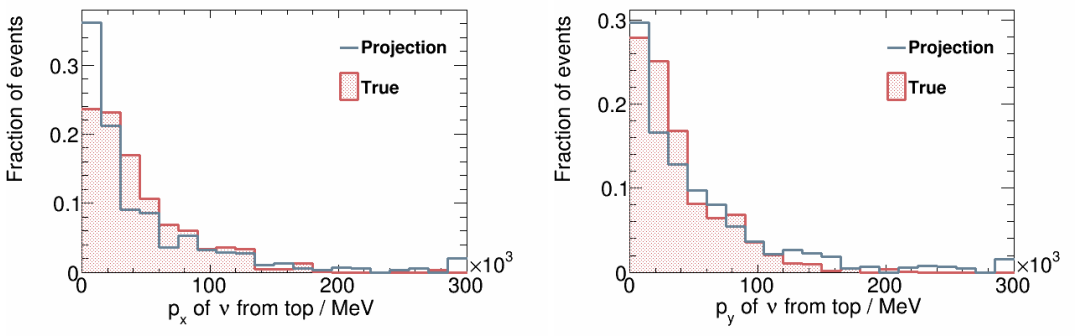


Figure 7.2: Comparison of the true transverse momentum components of the neutrino from the leptonic decay to the $E_{\mathrm{T}}^{\mathrm{miss}}$ projection method.

Performance

Using the method for splitting the missing transverse energy and the extra permutation of the two leptons as described above kinematic fits employing the likelihood Eq. 7.1 of $t\bar{t}H(\tau\tau)$ $2\,\ell, 1\,\tau_{\mathrm{had}}$ events have been performed with KLFitter. Due to technical reasons this method only works for events where both leptons are of the same flavour. For simplicity and due to their more precise measurement the following studies are restricted to $\mu\mu$-events. Since this significantly reduces the (already low) number of events available, the cut on the transverse momenta of the leptons has been lowered to $p_{\mathrm{T}}^{\ell_1,\ell_2} > 10\mathrm{GeV}$. Otherwise the event selection remains unchanged. Taking into account that the KLFitter-model requires four or more jets, an additional cut of $N_{\mathrm{jet}} \geq 4$ is imposed. Here, up to six of the jets leading in p_{T} are passed on to the fitter. Moreover a truth match via minimal ΔR has been performed for the leptons, ensuring that the lepton from the Higgs and top decay are indeed the two selected leptons. Since charged leptons are generally well measured and only signal events which definitely contain two leptons on truth level are analysed, no cut on the ΔR between true and reconstructed lepton is applied here. In this way about 16 % of the events are selected. The kinematic fit is performed successfully, without any technical problems, for 97% of these events. Out of those, the lepton ℓ_H is associated correctly with a reconstruction efficiency of (27.9 +/- 5.8)%. This is of course a surprisingly low percentage, especially for such a sophisticated and well motivated approach, indicating that this specific method is in fact not suitable for this application – hence, no distribution of the invariant $\tau\tau$-mass is shown here. It is nonetheless essential to analyse its problems and to understand whether in principle the theoretical concept is sound, and if not, if and how it could possible be improved. At first glance, one could interpret this very low reconstruction efficiency as refuting the assumption that this approach should work in the first place, since even randomly picking one of the two leptons should yield the correct one in about half of the events (provided the sample size is large enough). However, using KLFitter the situation is slightly different since not only the two leptons, but up to six jets are taken into account and sorted into the respective decays. The inefficiency of this part of the matching between particles on truth and reconstruction level probably also contributes to the inefficiency of identifying the lepton, as purely by chance the wrong lepton may fit better into the top quark decay than the correct one when combined with the wrong jets. Therefore the

following studies investigate the kinematic fit under more idealized conditions.

Truth matched events

As mentioned above it is crucial for the success of the kinematic fit that all of the objects expected by the model are detected and considered in the fit. However, due to the experimental limitations this ideal case is not given for all events. Especially due to the limited detector acceptance not all of the expected objects may be available. While in the above study only the leptons were truth matched, here a matching is also performed for the jets. To prevent possible ambiguities the event selection is refined by excluding events with additional b-jets on truth level and such where the hadronically decaying W boson produces more than two true partons [1]. While the first cut reduces the number of events by only a little more than 2% , the second one leaves only 49% of that. For each of the truth level partons the closest jet in ΔR is then determined and if this distance is below the cut-off of $\Delta R = 0.3$ the two objects are considered matched. Events in which multiple partons are matched to the same jet are rejected due to their obvious ambiguity – for example, this could occur due to the aforementioned jet merging. In order to remove another uncertainty, the b-tag of the two jets matched to the b quarks is set to 1, while that of the other two is set to 0. Hence for each event exactly four jets, out of which two are b-tagged, are used in the kinematic fit. This fixes the number of permutations of jets to 2 in all cases. Only 37% of the events can be truth matched unambiguously in this way, with the biggest loss occurring when finding the jets of the W_{had}.

All of these events can be fitted in KLFitter and here the reconstruction efficiency of the charged lepton is found to be only slightly better than before for all events at (32.2 +/- 14.8) %. As a caveat it is important to mention that at this point only 20 events are selected, increasing the statistical uncertainty accordingly. Since this reconstruction efficiency is still very low there seems to be another, possibly conceptional, problem in this method.

Using truth information and particles

True neutrino p_{T} One likely source of inefficiency is the method of estimating the missing transverse energy of the top neutrino, which may be too inaccurate for the kinematic likelihood fit. Therefore in the following the problem of approximating the momenta of the neutrinos in the Higgs and top decay will be neglected and instead the truth neutrinos will be used. To do so, a slight change of the KLFitter likelihood is necessary: As the resolution of the missing transverse energy is generally low, the transferfunctions correlating $E_{\mathrm{T}}^{\mathrm{miss}}$ and the neutrino p_{T} are very broad, leaving much freedom to adjust the values of p_x^ν and p_y^ν within the fit. Additionally, their width is proportional to the scalar $\sum E_{\mathrm{T}}$ of all objects in the event. This is done to take into account the fact that the more objects are measured in a final state, the more likely a mismeasurement or the presence of additional neutrinos occurs, decreasing the resolution on $E_{\mathrm{T}}^{\mathrm{miss}}$ even further. The true value of p_{T}^ν is of course unambiguous and in theory the fit should have no freedom to change the respective parameters. Ideally, the transferfunction

[1] In principle the latter case may be accounted for by merging the pair of partons with the smallest distance between them until only two resulting decay products of the W_{had} are left. Since this would introduce an additional inefficiency to the truth match and hence the lepton identification this is not done here.

W_{miss} then takes the shape of a delta-distribution, which is approximated here by a narrow Gauss-function with a width of $\sigma = 5\,\text{GeV}$ and the dependence on $\sum E_{\text{T}}$ is removed. In this way, using the same truth matched events as above, a reconstruction efficiency of the lepton of $(58.9 \pm 21.8)\%$ is achieved. On the one hand, comparing this result to above illustrates what a large impact an inaccurate approximation of the transverse momentum of the neutrino has on the reconstruction efficiency, but on the other hand the latter is still too low for the desired application.

True leptons and jets In order to eliminate inefficiencies due to the limited detector resolution and possible misidentification of objects the fit is performed using true jets and leptons. As the truth match requirement for the reconstructed objects can be dropped then, this also increases the event count. Obviously this means that the following results are based on events for which the approach could never be employed successfully when analysing reconstructed data, but they only serve to get an insight into the nature of the problems with this method. In this way the reconstruction efficiency of the lepton is improved up to $(73.4 \pm 14.1)\%$, again documenting a significant influence of imprecise measurements on the success of this method. However, naively, one would expect to reach a value closer to 100% in such idealized conditions, provided the theoretical concept is valid.

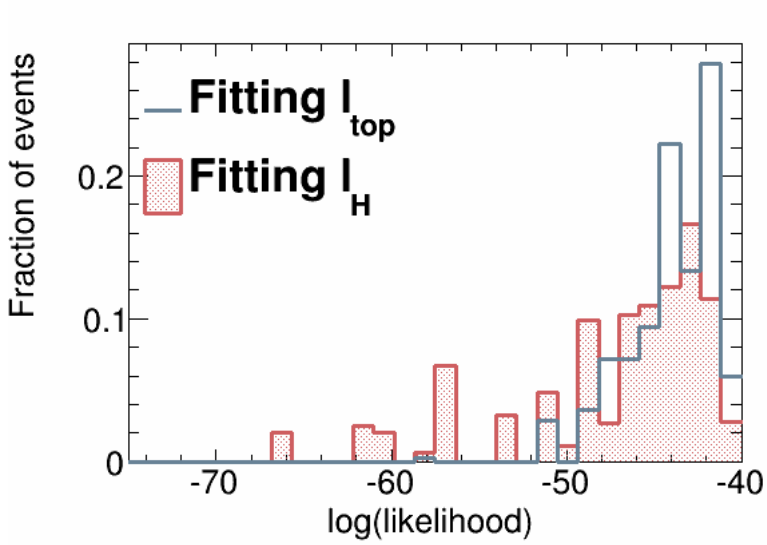


Figure 7.3: Logarithm of the KLFitter likelihood in the best permutation when fitting the correct lepton belonging to the top (black) and when fitting the wrong lepton (red) using full truth objects.

Comparing the logarithmic likelihood of the best permutation when fitting the correct lepton to fitting the wrong one, as shown in Fig. 7.3, a larger tail to lower values is observed in the fit of the wrong lepton. This behaviour is as expected: The peak around the maximally achieved value (in this case around -45) occurs for events for which all particles have been associated correctly. Each mismodelled particle decreases the likelihood, as wrong combinations are disfavoured. Therefore, the likelihood for the wrong lepton should tend to lower values than that of the correct lepton. Yet, this discrepancy is not pronounced enough to make an accurate

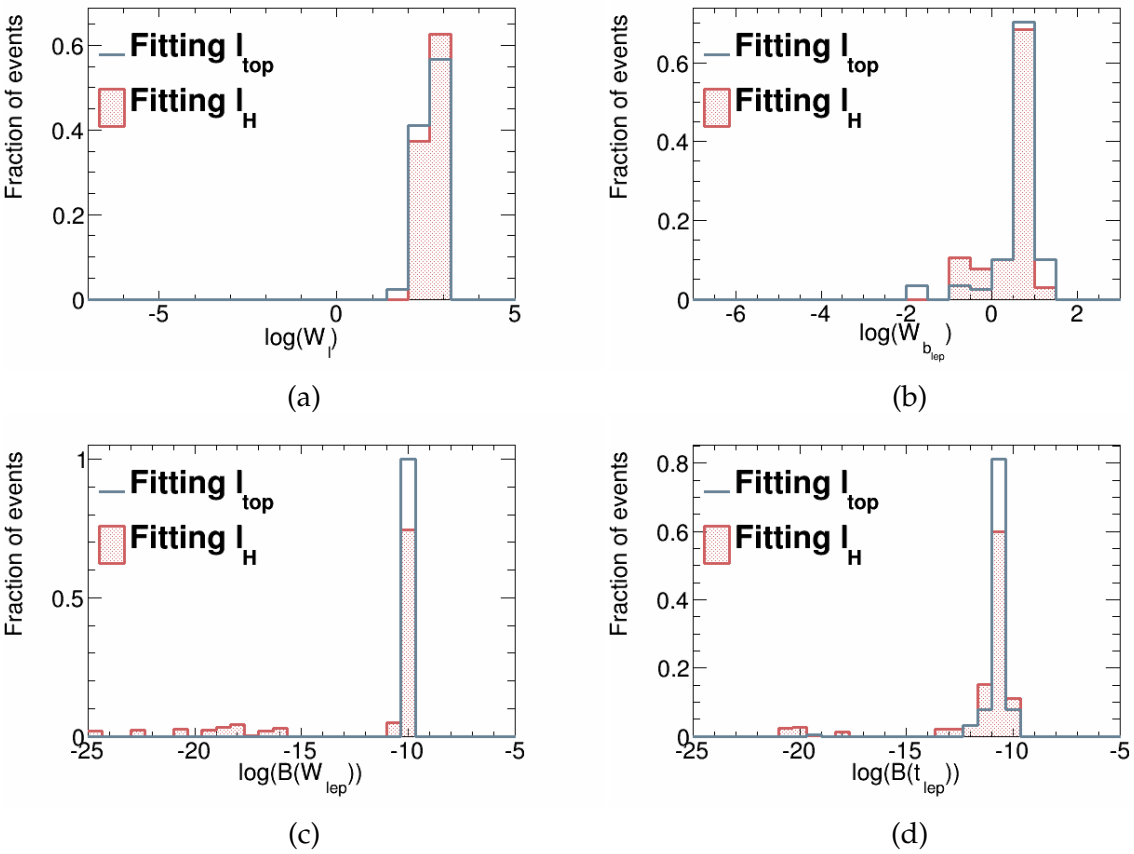


Figure 7.4: KLFitter likelihood components for the leptonic top decay using truth events: Transfer-function for the lepton (a) and the corresponding b-jet (b) as well as the Breit-Wigner terms for the leptonically decaying W boson (c) and top quark (d)

choice between the two leptons in many cases, resulting in the reconstruction efficiency above. To further investigate the possible shortcomings of the method, it is also helpful to regard the different components of the likelihood, since they are indicative of the quality of the object reconstruction.

Figure 7.4 gives an overview of the likelihood components for the leptonic decay - on the hadronic side the potential effects of mismodelling are simply reflected. The distributions of the transfer functions for the lepton (Fig. 7.4 (a)) and the b-jet which is produced alongside the leptonically decaying W boson (Fig. 7.4 (b)) are very similar when fitting the different leptons. This implies that in both cases the degree of freedom that they provide is used similarly to reconstruct the event and should not cause problems. Considering the contribution by the Breit-Wigner term describing the leptonic W decay (Fig. 7.4 (c)) it is startling that both distributions peak in much the same way, with only few events at low values for the wrong lepton. Since the Breit-Wigner distributions are very narrow around the resonant mass the wrong combination of lepton and neutrino should be strongly disfavoured, unless this effect could be compensated by a different contribution from the transfer functions for each lepton,

which is apparently not the case here. There seems to be another, underlying problem with this method. When combining the lepton and neutrino with the respective b-jet into the Breit-Wigner term for the leptonically decaying top quark (Fig. 7.4 (d)) the number of events which exhibit a strong rejection of the wrong combination is decreased even further. This indicates that the freedom in the association of the b-jets may partially compensate for the wrong lepton. To rule this out, only events where for both leptons the best permutation of the jets corresponds to the correct one can be regarded. In these kind of events the correct lepton is found with an efficiency of (76.1 +/- 18.8)% . As this is still an inefficiency larger than one would expect in such an ideal situation, there appears to be a conceptional problem, which is illustrated by the following study.

Dependence on the modelled neutrino p_z

One important fact not taken into account so far is that while a transferfunction to constrain the values of the neutrino p_T is employed, no such constraint exists for its longitudinal momentum component p_z. This parameter can thus be freely varied over a broad range and up into unrealistic values without negatively affecting the likelihood. Figure 7.5 exemplifies the situation: the vectors of the true lepton and neutrino from the leptonic top decay as well of the modelled objects are shown. While the reconstruction of the neutrino is inaccurate when fitting the correct lepton [2], it is moved even further from its true direction - especially in the z-component - when combined with the wrong lepton. In this way, the wrong lepton can be artificially engineered to yield a good estimation of the W boson mass. This wrong case then results in a higher likelihood than the correct one. Unfortunately, this happens for a sizeable fraction of events, in consequence severely diminishing the possibility to identify which lepton actually belongs to this decay.

In order to confirm this assessment the value of the neutrino p_z can be constrained by introducing an additional transfer function to the likelihood. To use the true neutrino p_z the same narrow Gaussian shape as for the transverse true neutrino momentum components is employed. With this set-up a reconstruction efficiency of the lepton of (97.4 +/- 21.2)% on full truth level is reached, highlighting that in the studies before the ideal case was not achieved due to this degree freedom. The comparison of logarithmic likelihoods shown in Fig. 7.6 supports this further: fitting the wrong lepton now results in no clear peak at maximum value but a smeared out distribution shifted to lower values than in the correct case. Hence, the shape of the likelihood is now suitable to discriminate between the two leptons successfully.

7.2.2 Conclusion

In conclusion, the likelihood based kinematic fit as provided by the KLFitter framework was found unsuitable for the desired application. It could be established that due to the ambiguities of the $2\,\ell, 1\,\tau_\mathrm{had}$ final state the event kinematics are under constrained. The freedom of the the neutrino p_z appears to be the crux of the matter. Ideally, this could be solved by the *neutrino weighting* approach as developed for precision measurements of the top quark mass in dileptonic decay channels [46]. This is based on assuming a neutrino rapidity distribution

[2] This is a known shortcoming of the KLFitter approach, cf. [45]

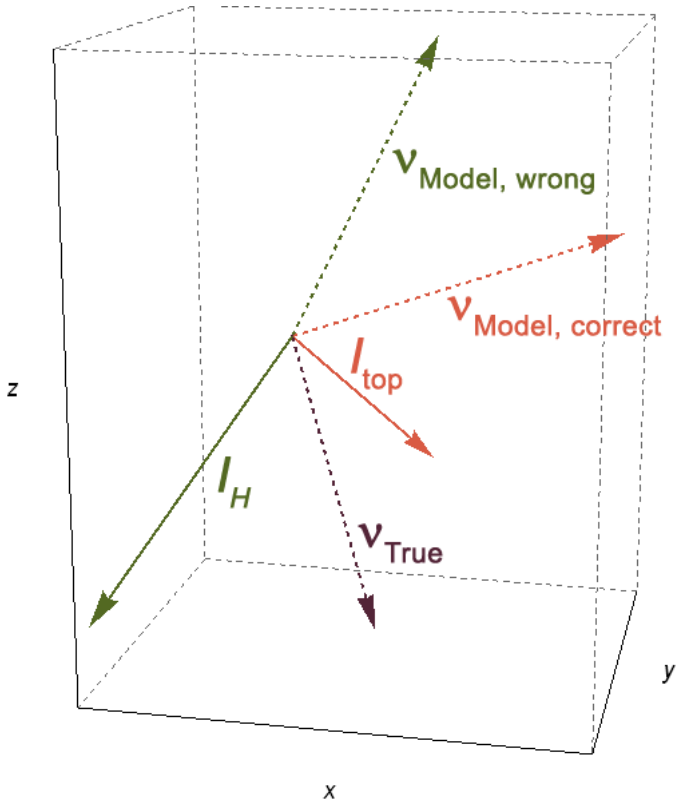


Figure 7.5: Vectors of the true and modelled neutrino and lepton when fitting the correct and wrong combination in KLFitter.

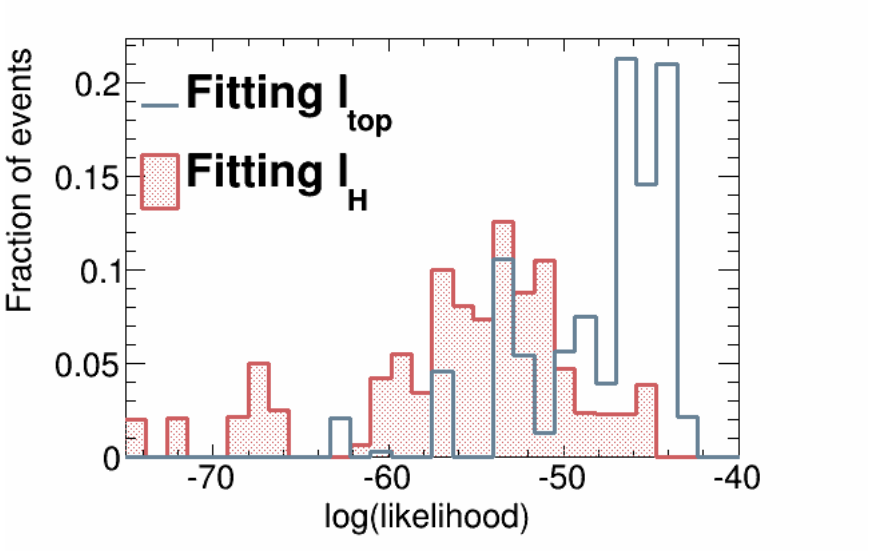


Figure 7.6: Logarithm of the KLFitter likelihood in the best permutation when fitting the correct lepton belonging to the top (black) and when fitting the wrong lepton (red) using full truth objects and a Gaussian-shaped transferfunction to additionally constrain the neutrino p_z .

from Standard Model expectations and then determining the most likely neutrino configuration for each event from them. However, due to the limited statistics this could not be done here. Instead an attempt at using a simpler, but more robust likelihood based approach is made in the following.

7.3 Outlook: Defining a simplified likelihood

Since it was found in the previous study that a full kinematic likelihood fit of the $t\bar{t}$ decay - as provided by the KLFitter framework - includes too many, not sufficiently constrained free parameters, simply calculating and comparing a likelihood, for each lepton to belong to the top quark decay, poses a good prospect. As a short outlook a simplified likelihood describing the top system of the $t\bar{t}H(\tau\tau)\ 2\,\ell, 1\,\tau_{\text{had}}$ process is defined and tested in this chapter. Since it only serves as a study of how successful such an approach could ideally be, truth matched leptons and jets are used in combination with the true p_{T} of the neutrino from the leptonic top decay.

No kinematic fit is performed and the likelihood contains no free parameters, which makes the latter conceptually different from the KLFitter likelihood. In the KLFitter approach certain parameters are allowed to be varied within their transfer function in order to take into account the finite detector resolution. Then the particle decays themselves can be modelled by Breit-Wigner distributions, as would be expected by theory. Omitting the kinematic fit and the transferfunctions of these parameters therefore means that the particle decays have to be modelled differently. Ideally they would each be accurately described by a convolution of the respective Breit-Wigner and resolution distribution. However, as the latter is Gaussian shaped and the dominant effect, to simplify the matter, the decays can be modelled by Gaussian distributions instead.

Additionally, this new approach provides opportunity to address another problem. Evidently the biggest difficulty in defining an accurate likelihood is the neutrino information and therefore the modelling of the leptonic top quark decay. Even if there was a sufficiently efficient method by which to separate the missing transverse energy $E_{\text{T}}^{\text{miss}}$ into the contributions of neutrinos from the Higgs and top quark decay, the longitudinal momentum component p_z can not easily be found from this, especially if it is not clear which lepton originates from the top quark decay. A simple solution to this problem may be restricting the likelihood components modelling the leptonic top quark decay to the transverse plane. To do so the *transverse mass*, which has found successful application in the measurement of the W boson mass (see e.g. [47]), can be employed. For the decay of a W boson into a charged lepton ℓ and neutrino ν it can be obtained as:

$$\begin{aligned} (m_{\text{W}}^{\text{T}})^2 &= \left(E_\ell^{\text{T}} + E_\nu^{\text{T}}\right)^2 + \left(\vec{p}_\ell^{\,\text{T}} + \vec{p}_\nu^{\,\text{T}}\right)^2 && (7.2) \\ &= m_\ell^2 + m_\nu^2 + 2 \cdot \left(E_\ell^{\text{T}} E_\nu^{\text{T}} + \vec{p}_\ell^{\,\text{T}} \vec{p}_\nu^{\,\text{T}}\right) && (7.3) \end{aligned}$$

where the transverse energy of a particle is given by the difference of its rest mass and transverse momentum, meaning $(E^{\text{T}})^2 = m^2 - (\vec{p}^{\,\text{T}})^2$. Neglecting the rest masses the transverse mass then equals

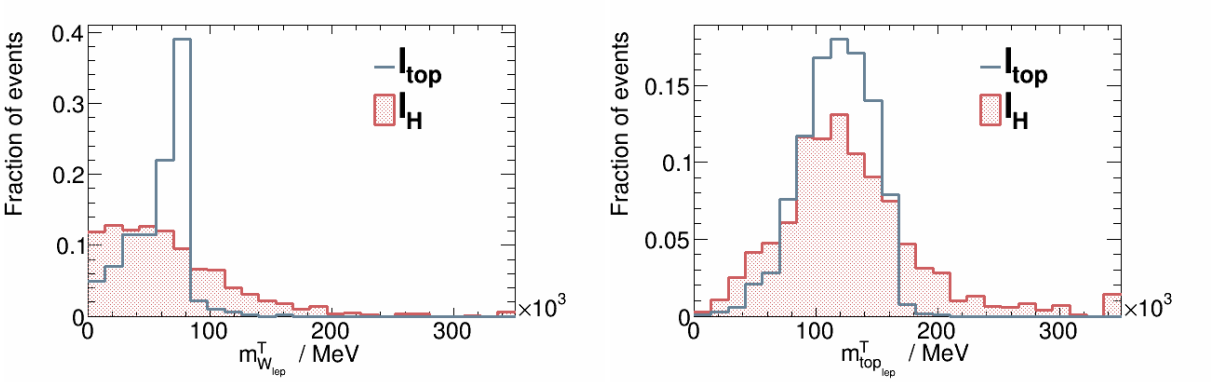
$$m_{\text{W}}^{\text{T}} \approx \sqrt{2 p_\ell^{\text{T}} p_\nu^{\text{T}} \left(1 - \cos\Delta\phi_{\ell,\nu}\right)} \tag{7.4}$$

, where $\Delta\phi_{l,\nu}$ describes the angle between the transverse momenta of lepton and neutrino.

Turning to reconstruct the top quark from its stable decay products - meaning the bottom quark, lepton and neutrino - the transverse mass has to be calculated for a three particle system. This is done in the same way as above and yields the following:

$$\begin{aligned}(m_{top}^{\mathrm{T}})^2 &= \left(E_b^{\mathrm{T}} + E_\ell^{\mathrm{T}} + E_\nu^{\mathrm{T}}\right)^2 + \left(\vec{p}_b^{\,\mathrm{T}} + \vec{p}_\ell^{\,\mathrm{T}} + \vec{p}_\nu^{\,\mathrm{T}}\right)^2 \\ &\approx 2 \cdot \left(p_b^{\mathrm{T}} p_\ell^{\mathrm{T}} \left(1 - \cos\Delta\phi_{b,\ell}\right) + p_b^{\mathrm{T}} p_\nu^{\mathrm{T}} \left(1 - \cos\Delta\phi_{b,\nu}\right) + p_\ell^{\mathrm{T}} p_\nu^{\mathrm{T}} \left(1 - \cos\Delta\phi_{\ell,\nu}\right)\right) \\ &= (m_{b,\ell}^{\mathrm{T}})^2 + (m_{b,\nu}^{\mathrm{T}})^2 + (m_W^{\mathrm{T}})^2\end{aligned} \tag{7.5}$$

The transverse mass is always smaller than the full invariant mass but the endpoint of its distribution should correspond to the true rest mass. The resulting distributions of the transverse mass of the leptonically decaying top quark and the W boson produced by it are compared for taking into account the correct lepton, ℓ_{top}, and the wrong lepton, ℓ_H, in Fig. 7.7 using truth matched events. While for the leptonic W boson decay the two distributions take very different shapes, part of the separation power is lost when also taking into account the bottom quark to reconstruct the top. However modelling the leptonic top decay with the transverse mass may be sufficient to distinguish between the two leptons.



(a) (b)

Figure 7.7: Transverse mass spectra of leptonic top quark and W boson decay determined using truth events.

Clearly, the transverse mass spectra are not symmetric, but rather exhibit a very sharp edge at high values. Nonetheless, as a first, rough attempt at modelling all truth matched mass spectra are fitted with a single Gaussian each. The resulting fit functions as well as their parameters - the mean μ and width σ - are depicted in Fig. 7.8. It is important to note that the scaling factor here is irrelevant, since the Gaussian functions used in the likelihood are normalized to 1. This corresponds to the most likely case being found in the peak of the distribution. An accurate modelling of the position of the latter is therefore more important. While the distributions of the hadronic top and W boson decay as well as the leptonic top decay are well described by the fit in that respect, the shape of the transverse mass of the leptonically decaying W boson is not

accurately represented. A more complex fit function would probably be needed to describe its sharp edge.

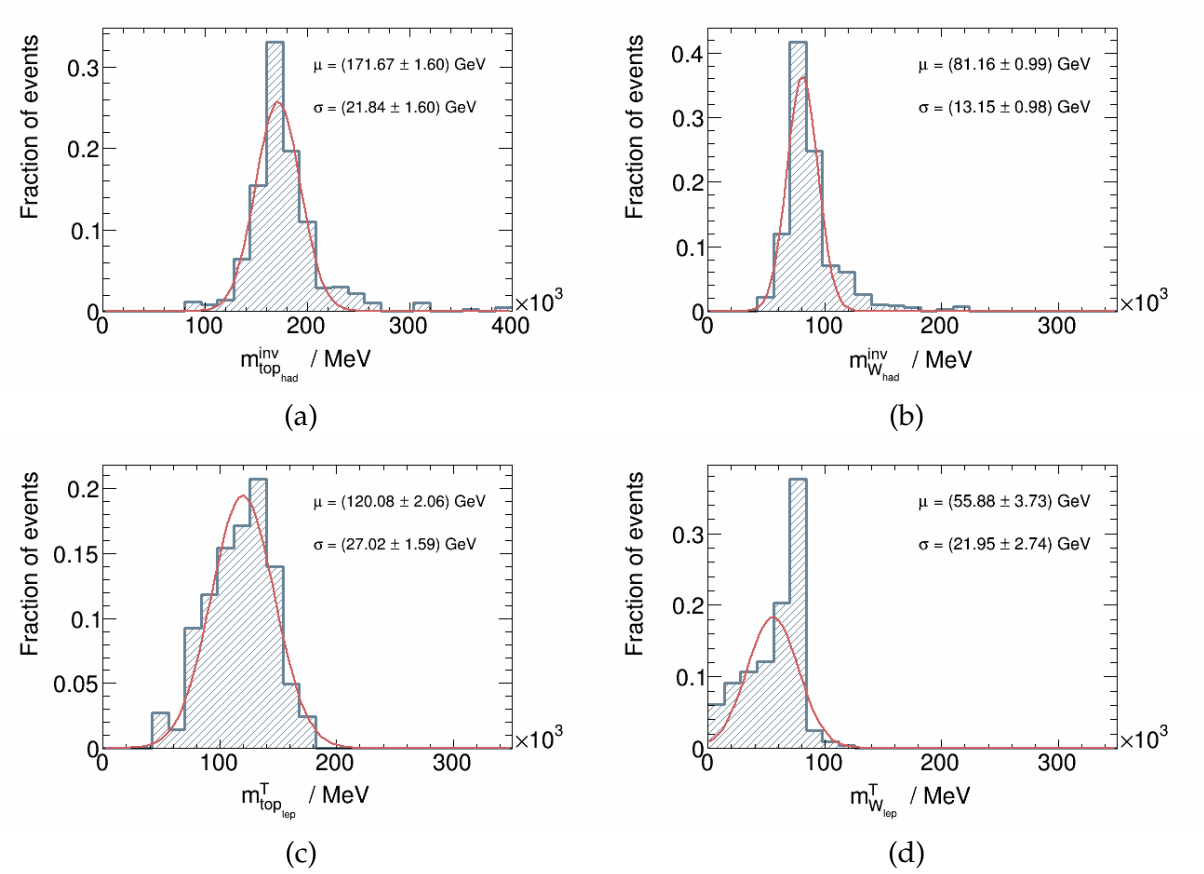


Figure 7.8: Gaussian fit function for the truth matched mass spectra entering the likelihood.

The full likelihood is then defined as follows

$$\mathcal{L} = G\left(m_{q_1q_2q_3}|\mu_{\text{top}_{\text{had}}}, \sigma_{\text{top}_{\text{had}}}\right) \cdot G\left(m_{q_1q_2}|\mu_{W_{\text{had}}}, \sigma_{W_{\text{had}}}\right) \tag{7.6}$$

$$\cdot \; G\left(m^{\text{T}}_{q_4lv}|\mu^{\text{T}}_{\text{top}_{\text{lep}}}, \sigma^{\text{T}}_{\text{top}_{\text{lep}}}\right) \cdot G\left(m^{\text{T}}_{lv}|\mu^{\text{T}}_{W_{\text{lep}}}, \sigma^{\text{T}}_{W_{\text{lep}}}\right) \tag{7.7}$$

where the same notation as used before applies. Again all possible permutations of jets are considered and for each lepton separately the permutation yielding the maximum value is determined. The resulting distributions of the likelihood value in the best permutation for the two leptons ℓ_H and ℓ_{top} are compared in Fig. 7.9. As can be seen, they take a very similar shape. Even when considering the correct lepton ℓ_{top} the likelihood often takes small values near 0. Quantifying the success of this method at associating the leptons with their respective decay systems, the correct lepton yields a higher likelihood in $(68.7 \pm 6.1)\%$ of events. While this is better than KLFitter performed using truth matched objects and true neutrino information, it is not an improvement over the simple variables. At least partially this may be due the very rough modelling of the mass distributions.

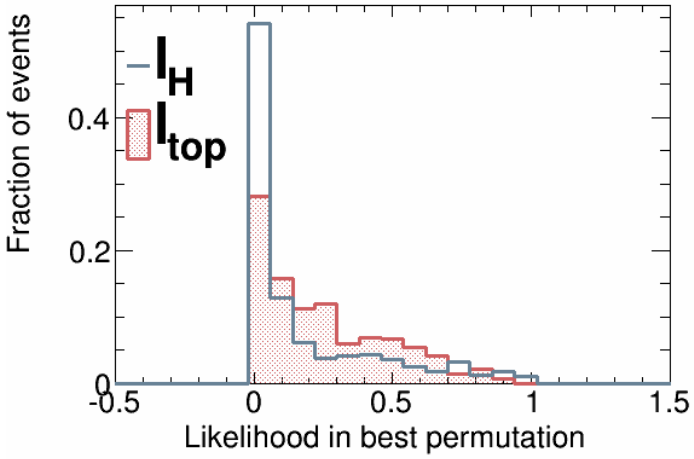


Figure 7.9: Simplified likelihood describing the $t\bar{t}$ process for the correct lepton ℓ_{top} and wrong lepton ℓ_H.

To investigate whether the concept is sound, again the different components of the likelihood can be compared for the two leptons. This is shown in Fig. 7.10. As can be seen, especially the Gaussian term describing the leptonic W boson decay provides good separation between the leptons. When combined with the b-jet into the leptonically decaying top this separation is lost partially. Moreover, the distributions of the components describing the hadronic decays look very similar. In this way the full likelihood probably loses its distinguishing power. Therefore it may be more successful to neglect the hadronically decaying top quark and only consider the likelihood for the leptonic decay.

To summarize, while this method was not found to be more successful than other ap-

proaches at this point, it exhibits much potential for improvements. For one thing, functions that model the invariant mass spectra considered in the likelihood more accurately are necessary. Additionally, the likelihood may be further simplified to include only the terms which actually contribute to the distinguishing of the leptons. Taking a different direction, extending the likelihood to describe the full $t\bar{t}H(\tau\tau)$ process and employing it directly for signal extraction is also a possibility.

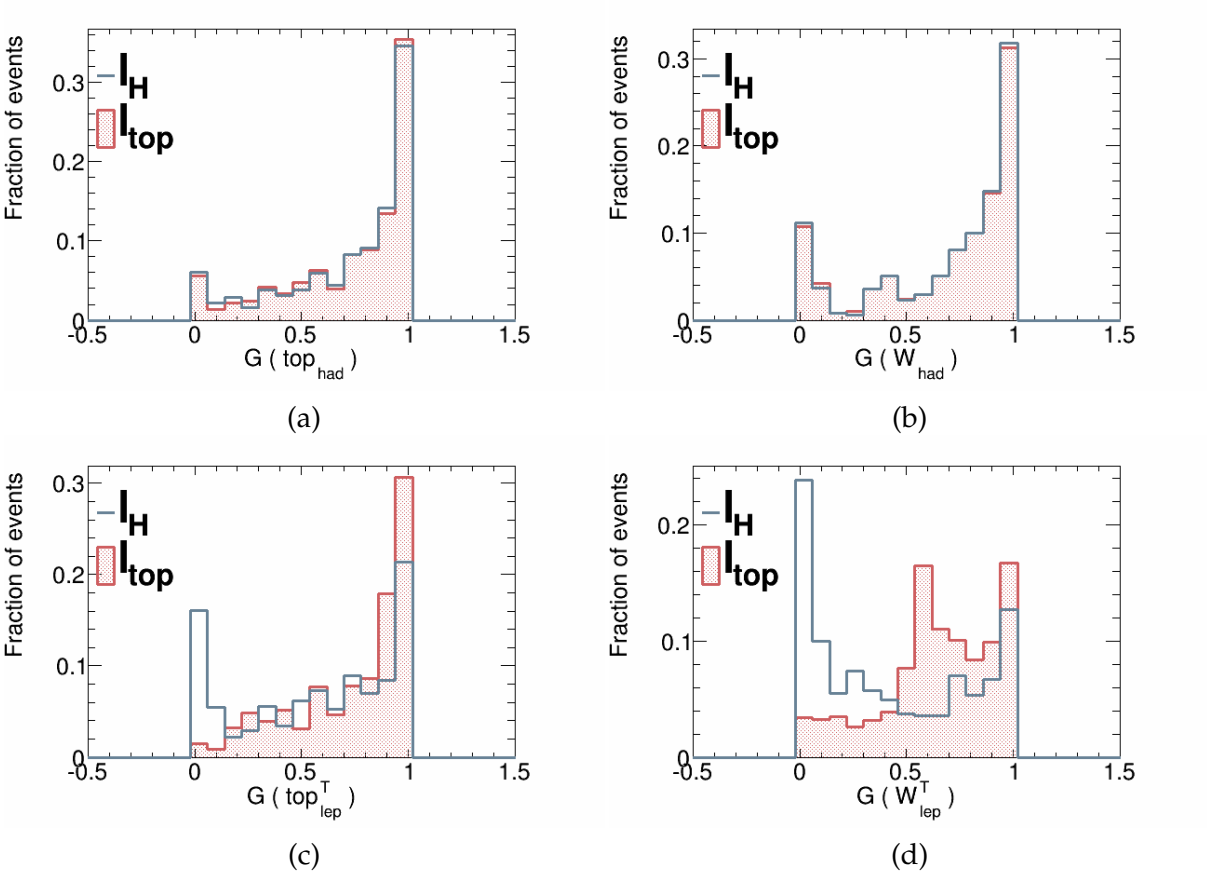


Figure 7.10: Components of the simplified likelihood for the leptons ℓ_H and ℓ_{top}.

CHAPTER 8

Conclusion and outlook

In this thesis the topologies and kinematics of $t\bar{t}H(\tau\tau)$ events have been studied in order to investigate the possibility of applying more complex methods in the course of the search for $t\bar{t}H$ processes. A strong incentive to improve the latter is given, as with the start of Run 2 of the LHC a good prospect of available statistics for this channel is given. Therefore, it is fitting to investigate more elaborate analysis going beyond the cut-and-count approach employed in the ATLAS analysis of Run 1 data. In particular, the idea was to use the shape of the invariant mass distribution of the $H \to \tau\tau$ decay for signal extraction. The studies were focused on the $2\,\ell, 1\,\tau_{\text{had}}$ channel, in which both the $\tau\tau$ and $t\bar{t}$ pair decay semileptonically. It was found that ideally the invariant $\tau\tau$ mass grants a promising shape separation against $t\bar{t}$ background topologies which include fake leptons. Such $t\bar{t}$ topologies are not well suppressed by the requirement for the two leptons to be of the same charge, as applied in the current analysis procedure. While it is efficient at excluding most other $t\bar{t}$ background topologies, this cut is of very low signal efficiency. Thus, employing an invariant mass reconstruction of the $H \to \tau\tau$ is especially promising. It could either supplement or replace the same charge requirement.

However, within the signal process, only one of the two selected leptons originates from the $H \to \tau\tau$ decay, while the other one is created in a leptonic top decay. The combinatorial background to the mass reconstruction due to this ambiguity is not easy to suppress. It was found that this ambiguity could not easily be solved by simple considerations, such as the distances between final state particles. This is due to the fact that the final state topology is very complex. The intuitive assumption that the products of the Higgs and top quark decays are found in narrow, non-overlapping cones around the directions of the original particles could not be confirmed.

Additionally, the production of neutrinos in both decay systems requires the development of methods to make the reconstructed $E_{\text{T}}^{\text{miss}}$ usable. It would need to be separated into the contributions from each decay system to be employed sensibly in the event reconstruction. This problem could not be solved in a satisfactory manner.

Nonetheless a likelihood-based kinematic fit of the $t\bar{t}$ system using KLFitter was investigated. It could be established that this approach was not conceptually sound insofar as it provides too much freedom of the fit parameters, especially the longitudinal momentum component p_z of the neutrino. In this way the event kinematics are underconstrained and the lepton from the $t\bar{t}$ does not necessarily yield a higher KLFitter likelihood. In principle, this problem might be solved by the neutrino weighting approach, as employed for example in the KLFitter likelihood for dileptonic $t\bar{t}$ topologies. This could not be done here, as statistics

were too limited. Instead a simplified likelihood was defined and tested, which was found to have more potential. However it still leaves much room for improvement.

With the prospect of higher statistics due to the increase in center of mass energy in Run 2 of the LHC, ample opportunity for follow-up studies is given. For one thing, instead of using a likelihood based approach as a tool for the mass reconstruction it could be used for signal extraction directly. For this purpose a new likelihood including the $H \to \tau\tau$ decay could be defined, to describe the full event kinematics. To do so the Higgs mass would need to be included as a constraint. The latter could be calculated using more complex mass estimators. Especially the collinear mass seems promising in that respect, as it was found that in $t\bar{t}H$ events the visible and invisible decay products are collinear to the original τ-directions. Yet, before this can be implemented the $E_{\mathrm{T}}^{\mathrm{miss}}$ ambiguity would need to be solved. Using an extended KLFitter likelihood including the neutrino weighting feature is therefore worth considering, as it would address this problem by definition. Taking a different approach, the simple topological variables investigated here may be combined as input to a multivariate analysis, i.e. employing a boosted decision tree.

Moreover, indications that the final state topology becomes easier to separate into $t\bar{t}$ and Higgs system given a certain boost of the Higgs boson were found. Therefore, if provided with enough statistics, the matter may be simplified by restricting the analysis to boosted topologies. Extending the horizon to consider also other possible decay modes is also a possibility. For example, a $t\bar{t}H(\tau\tau)$ process in which the $H \to \tau\tau$ system occurs fully hadronic, while one top quark decays leptonically would result in a final state containing only one light lepton and two τ-jets. While this would provide similar advantages as the $2\,\ell, 1\,\tau_{\mathrm{had}}$ channel in terms of branching fraction and clearly recognizable signature, it would include less ambiguities as well as neutrinos.

APPENDIX A

Useful information

A.1 Finding τ_H^{vis}

Leading p_{T} τ-jet Since for this channel a cut of requiring more than one hadronic τ-decay is applied, in principle there is also an ambiguity in the selection of the τ_H^{vis}. Even when excluding a decay into another τ-lepton of the $t\bar{t}$-system additional hadronic τ-decays could be present - either as fakes or part of additional radiation. It is both simple and intuitive to employ the p_{T} ordering of the detected tau-jets to establish which τ-jet belongs to the Higgs decay.

Taking into account all possible semi-leptonic decays of the $t\bar{t}$-system (hence not excluding those involving τ-leptons) in principle one of those τ-jets could originate from a top-quark decay. Yet in this case, its p_{T} should be less than that of the one originating from the $H \rightarrow \tau\tau$ decay. For one, in the decay of a top-quark into a bottom-quark, a τ-jet and neutrinos one more decay occurs when comparing to the Higgs decay system : In $H \rightarrow \tau\tau$ the Higgs decays directly into a τ-pair, one of which creates the hadronic τ-jet via a decay to tau neutrino and W boson. In contrast the top-quark first decays into a b-jet and W boson, which itself subsequently decays into a τ-neutrino and -lepton and here only the latter creates the hadronic τ-decay via another W boson. Additionally the mass of the Higgs-boson is larger than that of a W boson, hence more energy is available in the first decay.

Apart from this extra τ-jets could be fakes and/or due to additional radiation in the event, which typically also means low p_{T} objects. Thus it is likely that the τ-jet with the highest p_{T} value originates from the the $H \rightarrow \tau\tau$ decay.

Using this expectation the invariant $m_{\tau\tau}^{\text{vis}}$ distribution when reconstructing from the correct, truth matched combination of ℓ_H and τ_H^{vis} can be compared to that achieved when replacing the τ-jet with the one leading in p_{T}, as shown in the top of figure A.1. Evidently the distributions are very similar in shape and allowing for possible misidentifications of the hadronic τ-jet only causes a negligible tendency to large values of $m_{\tau\tau}^{\text{vis}}$. This is confirmed by considering the small difference between the distributions shown in the bottom of A.1. To quantify this effect the truth matching of the hadronic τ-decays can be employed in order to calculate the percentage of events in which the leading p_{T} τ-jet is a product of the Higgs-decay. In fact this is the case in $(88.6 \pm 4.4)\%$ of the signal events . One reason why for a small part of the events this criterion does not identify the correct τ-jet could be that in those events the true τ-jet is not detected and reconstructed, for example because its trajectory is outside of the detector acceptance or the hadronic τ-decay reconstruction algorithm did not work. Apart from this

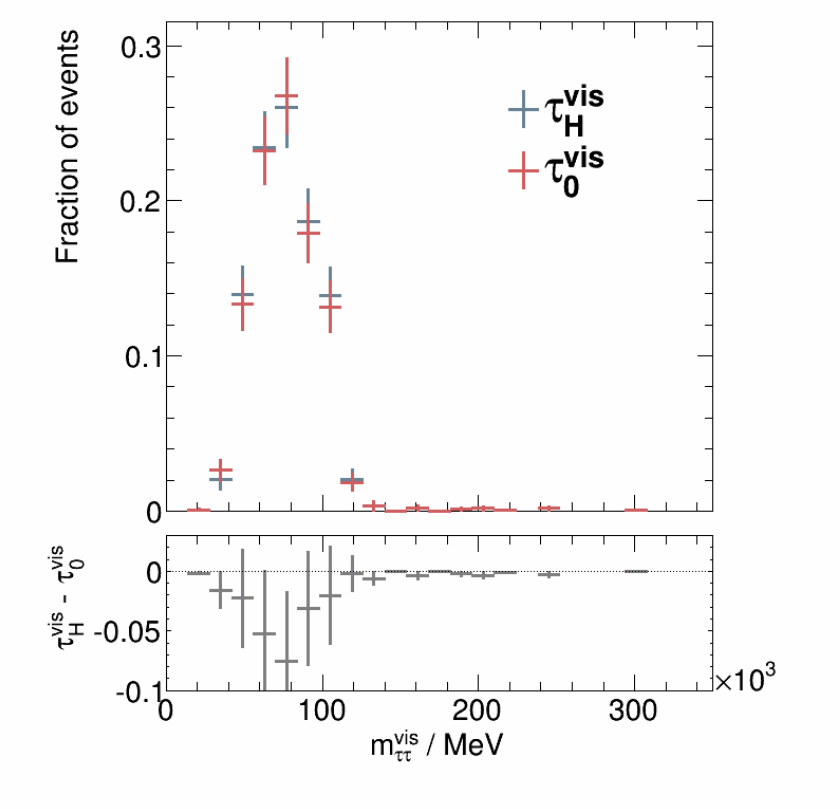


Figure A.1: Comparison of the visible mass distribution for signal events when using the truth matched combination of ℓ_H and τ_H^{vis} and when using ℓ_H together with the leading p_{T} τ_0^{vis}. In the bottom the difference of these distributions is shown.

of course, the transverse momenta of the decaying parent particles also play a role, so that in some events the arguments described above do not hold.

To summarize, this efficiency to find the τ-jet belonging to the Higgs-system correctly by exploiting the p_{T} ordering appears near ideal and sufficient to preserve the distinct narrow peaked shape of the invariant mass distribution, which holds good potential for signal and background discrimination. Thus for the further studies of this thesis the problem of choosing the right hadronically decaying τ is considered solved in a satisfactory way and only the more complex identification of the lepton ℓ_H will be investigated. Therefore to study the effect of problems linked with the latter in an isolated, idealized situation the truth matched τ_H^{vis} is used in combination with the various identification methods for ℓ_H.

A.2 Additional plots of simple variables for lepton association

A.2.1 Separation between lepton and tau

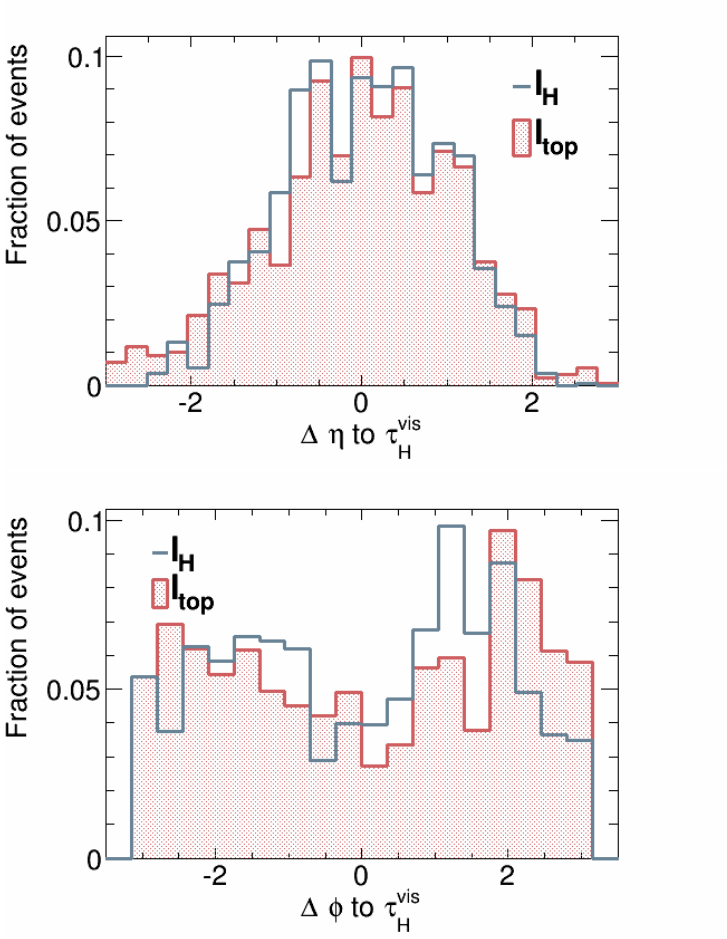


Figure A.2: Separation in pseudo-rapidity, $\Delta\eta$, and azimuthal angle, $\Delta\phi$, between the τ-jet from $t\bar{t}H(\tau\tau)$ to the lepton from the Higgs- and top-decay.

A.2.2 Separation between leptons and closest true b-jet

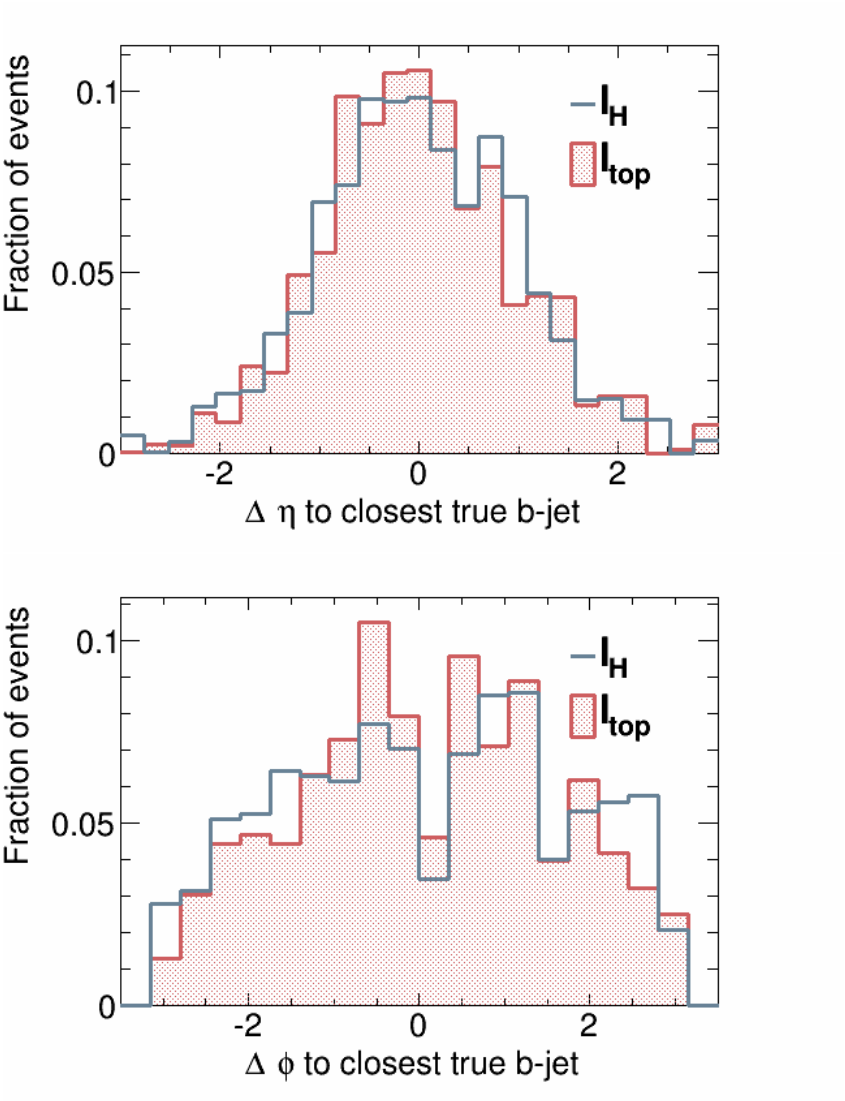


Figure A.3: Separation in pseudo-rapidity, $\Delta\eta$, and azimuthal angle, $\Delta\phi$, between the leptons from the Higgs- and top-decay and the closest true b-jet respectively.

A.2.3 Separation between leptons and closest jet

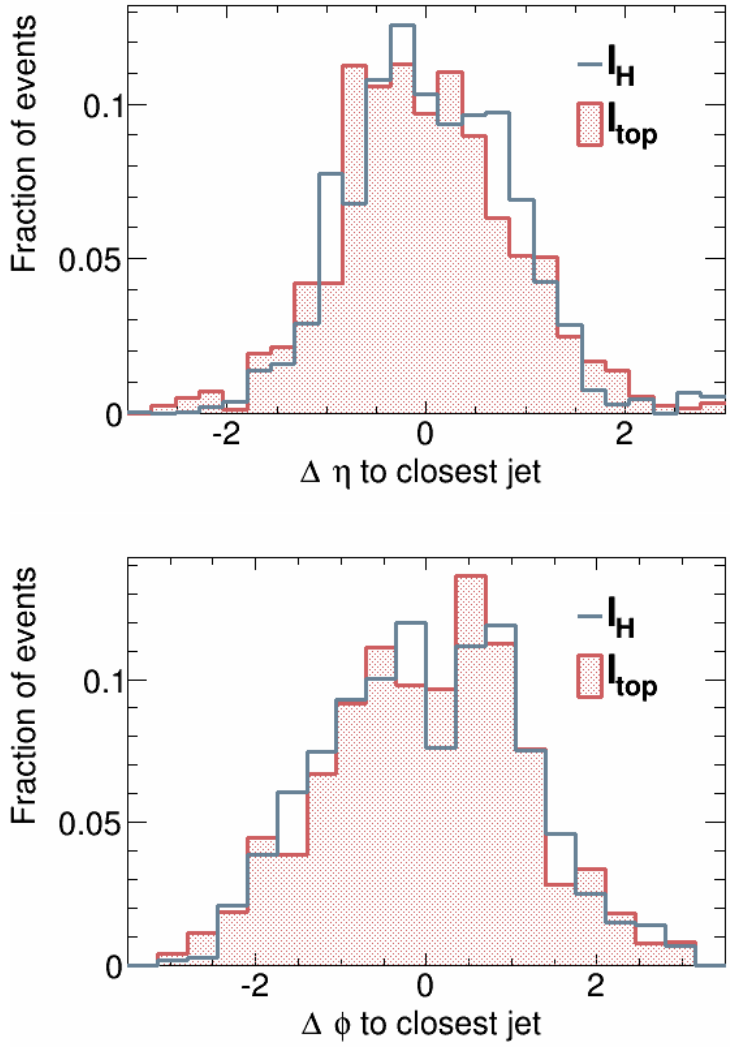


Figure A.4: Separation in pseudo-rapidity, $\Delta\eta$, and azimuthal angle, $\Delta\phi$, between the leptons from the Higgs- and top-decay and the closest jet respectively.

APPENDIX A

Bibliography

[1] P. Higgs, *Broken symmetries, massless particles and gauge fields*, Physics Letters **12**.2 (1964) 132–133, ISSN: 0031-9163.

[2] G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys.Lett. **B716** (2012) 1–29, arXiv: 1207.7214 [hep-ex].

[3] S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30–61, arXiv: 1207.7235 [hep-ex].

[4] G. Aad et al., *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (19 2015) 191803.

[5] K. Olive and P. D. Group, *Review of Particle Physics*, Chinese Physics C **38**.9 (2014) 090001.

[6] D. Griffiths, *Introduction to Elementary Particles*, Physics textbook, Wiley, 2008, ISBN: 9783527406012.

[7] M. Srednicki, *Quantum Field Theory*, Cambridge University Press, 2007, ISBN: 9781139462761.

[8] S. Dawson, *Introduction to Electroweak Symmetry Breaking*, AIP Conf.Proc. **1116** (2009) 11–34, arXiv: 0812.2190 [hep-ph].

[9] L. Reina, *TASI 2004 lecture notes on Higgs boson physics* (2005) 261–356, arXiv: hep-ph/0512377 [hep-ph].

[10] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (9 1964) 321–323.

[11] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (20 1964) 585–587.

[12] G. Aad et al., *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, JHEP **1504** (2015) 117, arXiv: 1501.04943 [hep-ex].

[13] *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s}$ = 7 and 8 TeV in the ATLAS experiment*, tech. rep. ATLAS-CONF-2015-007, CERN, 2015.

[14] F. Demartin et al., *Higgs production in association with a single top quark at the LHC* (2015), arXiv: 1504.00611 [hep-ph].

[15] M. Beneke et al., *Top quark physics* (2000), arXiv: hep-ph/0003033 [hep-ph].

[16] Gross, Eilam and Živković, Lidija, *$t\bar{t}H \to t\bar{t}\tau^+\tau^-$ —toward the measurement of the top-Yukawa coupling*, Eur. Phys. J. C **59**.3 (2009) 731–754.

[17] A. Belyaev and L. Reina, *pp —> t anti-t H, H —> tau+ tau-: Toward a model independent determination of the Higgs boson couplings at the LHC*, JHEP **0208** (2002) 041, arXiv: hep-ph/0205270 [hep-ph].

[18] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3**.08 (2008) S08001.

[19] Z. M. for the ATLAS Collaboration, *Simulation of Pile-up in the ATLAS Experiment*, Journal of Physics: Conference Series **513**.2 (2014) 022024.

[20] M. Lamont, *Status of the LHC*, Journal of Physics: Conference Series **455**.1 (2013) 012001.

[21] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3**.08 (2008) S08003.

[22] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19, CERN, 2010.

[23] ATLAS Outreach, *ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC*, http://www.atlas.ch/fact-sheets-1-view.html, 2010.

[24] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **0804** (2008) 063, arXiv: 0802.1189 [hep-ph].

[25] *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, tech. rep. ATLAS-CONF-2014-046, CERN, 2014.

[26] *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, tech. rep. ATLAS-CONF-2014-032, CERN, 2014.

[27] G. Aad et al., *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, Eur.Phys.J. **C74**.11 (2014) 3130, arXiv: 1407.3935 [hep-ex].

[28] The ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV* (2014), arXiv: 1412.7086 [hep-ex].

[29] *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV*, tech. rep. ATLAS-CONF-2013-082, CERN, 2013.

[30] The ATLAS collaboration, *Search for the associated production of the Higgs boson with a top quark pair in multi-lepton final states with the ATLAS detector* (2015).

[31] G. Aad et al., *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector* (2015), arXiv: 1503.05066 [hep-ex].

[32] *Higgs cross sections for HL-LHC and HE-LHC*, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/HiggsEuropeanStrategy.

[33] *ATLAS-CMS recommended predictions for top-quark-pair cross sections using the Top++v2.0 program*, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO.

[34] J. M. Campbell and R. K. Ellis, *$t\bar{t}W^{+-}$ production and decay at NLO*, JHEP **1207** (2012) 052, arXiv: 1204.5678 [hep-ph].

[35] R. Röntsch and M. Schulze, *Constraining couplings of top quarks to the Z boson in $t\bar{t}$ + Z production at the LHC*, JHEP **1407** (2014) 091, arXiv: 1404.1005 [hep-ph].

[36] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **0411** (2004) 040, arXiv: hep-ph/0409146 [hep-ph].

[37] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to {PYTHIA} 8.1*, Computer Physics Communications **178**.11 (2008) 852–867, ISSN: 0010-4655.

[38] *The CMS Offline WorkBook*, https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookHowToFit.

[39] A. Elagin et al., *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, Nucl.Instrum.Meth. **A654** (2011) 481–489, arXiv: 1012.4686 [hep-ex].

[40] A. Frodesen, O. Skjeggestad and H. Tofte, *Probability and Statistics in Particle Physics*, Universitetsforlaget, 1979.

[41] The ATLAS Collaboration, *Measurement of charged-particle event shape variables in inclusive $\sqrt{(s)}$=7 TeV proton-proton interactions with the ATLAS detector*, tech. rep., 2013 032004.

[42] J. Erdmann et al., *A likelihood-based reconstruction algorithm for top-quark pairs and the {KLFitter} framework*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **748** (2014) 18–25, ISSN: 0168-9002.

[43] G. Aad et al., *Measurement of the top quark pair production charge asymmetry in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, Journal of High Energy Physics **2014**.2, 107 (2014).

[44] A. Altheimer and G. Brooijmans, *Matrix Elements, Parton Showers and Jet Merging: Jet Substructure and New Physics at the LHC* (2011), arXiv: 1109.0030 [hep-ex].

[45] B. Lemmer, *Measurement of Spin Correlations in $t\bar{t}$ Events from pp Collisions at $\sqrt{s}$ = 7 TeV in the Lepton + Jets Final State with the ATLAS Detector* (2014), arXiv: 1410.1791 [hep-ex].

[46] B. Abbott et al., *Measurement of the Top Quark Mass Using Dilepton Events*, Phys. Rev. Lett. **80** (10 1998) 2063–2068.

[47] B. Abbott et al., *Measurement of the W boson mass*, Phys. Rev. D **58** (9 1998) 092003.

List of Figures

List of Tables

Acknowledgements

Firstly, I would like to thank Norbert Wermes for giving me the opportunity to work on an interesting topic at the forefront of high energy particle physics research in his group and for providing me with new ideas and helpful suggestions during the course of it. Besides that, I owe gratitude to Jochen Dingfelder, for kindly agreeing to be the second referee for my thesis. A big thank you also to my supervisor Jürgen Kroseberg who provided me with continuous and patient support, without which this thesis would not haven been possible. I am also grateful to Thomas Schwindt and David Hohn for assisting me with problems of technical nature as well as their readiness to answer all my questions. In this respect I also owe gratitude to Jessica Liebal for discussing both physics and the aesthetics of plots with me. Moreover, I would like to thank the developers of KLFitter, especially Boris Lemmer, Kevin Kröninger and Johannes Erdmann for their input as well as pratical help.