Improvements on the mass reconstruction of single top quark associated Higgs production at ATLAS analyzing the decay channel $H \rightarrow \tau\tau$

Mathias Weiß

Masterarbeit in Physik
angefertigt im Physikalischen Institut

vorgelegt der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
Bonn

April 2023

I hereby declare that this thesis was formulated by myself and that no sources or tools other than those cited were used.

Bonn, 05.04.23
Date

Mathias Weiß
Signature

1. Gutachter: Prof. Dr. Ian Brock
2. Gutachter: Prof. Dr. Florian Bernlochner

Acknowledgements

The research in contemporary physics, escpecially in experimental elementary particle physics, is highly developed. Contributing to the field without collaborating with other researchers is very unlikely if not impossible. The accelerators and detectors needed for the experiments are so sophisticated that construction and operation of those are research fields on their own. Since accelerators and detectors are not only very sophisticated but also very expensive, the challenges in the field are not only of scientific and technichal but also of financial nature.

Without CERN, which has been founded in 1954 as an international collaboration in order to face these challenges, the analysis presented in this thesis would probably not have been possible. I have to acknowledge the effort of many people responsible in science as well as in poliltics who enabled the foundation of CERN and/or its subsequent projects, especially the LHC and ATLAS. My thesis is only a tiny piece in a huge international physics community. Moreover, my thanks are due to all people in the present who do not only concentrate on their own research but also invest efforts in order to keep the collaboration running.

Among the people I know in person, my thanks are first and foremost due to Ian Brock, who defined the topic of my thesis, and to Tanja Holm, who was my supervisor. She was always supportive and showed me thinking constructively anytime. I thank Florian Bernlochner for agreeing to be the second referee of my thesis and for his personal commitment in finding a date for my colloquium. I also thank all other members of the Brock group, the Physics Institute or the ATLAS collaboration who gave me feedback or support on various issues.

The personal circumstances of the work on this thesis have not always been easy. At this point I would like to thank Ian again for his initiative to support me and for his valuable ideas. Furthermore, there are many people who supported me and who I would like to thank at this point: my parents, my sister, my friends and my brothers and sisters in faith. I forego giving a list of names in fear of forgetting someone. Last but not least, my special thanks go to Michael Zeunert, Eric Bernhard and David Hickson for checking parts of my thesis for linguistic acurracy.

Contents

CHAPTER 1

Introduction

Elementary particle physics aims to reveal the fundamental matter constituents and fundamental forces. Due to many successes in both theoretical and experimental research in the last decades, it was possible to work out the Standard Model (SM) of particle physics, which arranges all known elementary particles in a symmetric pattern and describes their interactions with mathematical methods. The latest particle predicted by the Standard Model which lacked experimental evidence, was the Higgs boson. Although found in 2012, there is still research ongoing on its properties predicted by theory. One of these properties is the coupling to the top quark, another elementary particle.

The coupling is described by the coupling constant y_t, a complex number. Analyzing events of single Higgs bosons, produced in association with a pair of top quarks, called $t\bar{t}H$, the magnitude of y_t was found to be compatible with the SM prediction [1]. But as the analysis of this process does not reveal any information about the phase of y_t, a different process, called tHq, has to be analyzed. This process also includes a single Higgs boson, but not produced in association with a pair of top quarks but with a single top quark, which is much rarer. A significant deviation from the Standard Model prediction would be a further hint to physics beyond the Standard Model (BSM physics).

Both $t\bar{t}H$ and tHq, together with many, much more frequent background processes, are detected with the ATLAS detector at the Large Hadron Collider (LHC), a particle accelerator at CERN (the European Organization for Nuclear Research). As tHq is very rare, techniques for separating signal from background events have to be developed and tested on Monte Carlo simulated events, before applying them to data. The analysis reported in this thesis is exclusively applied to simulated events and is going to be applied to data events taken during Run 2 of the LHC.

The analysis of tHq has to consider the various decay channels of both the Higgs boson and the top quark. The analysis reported in this thesis deals with the very specific decay channel, called $2\,e/\mu + 1\tau_{\text{had}}$, where the Higgs boson decays into a pair of taus, of which one decays leptonically and the other one hadronically, while the top quark decays leptonically. The assigning of the two final-state light leptons to the Higgs boson and the top quark is not clear in all cases. There are four final-state neutrinos, which cannot be detected.

This thesis aims to contribute to the separation of signal to background by improving the mass reconstruction of the Higgs boson in tHq. It reports about two improvements. First: Based on correlation studies, components of the neutrino four-momentum originating from the top quark are predicted. This allows using methods for the mass reconstruction of the Higgs boson in tHq which are already known from pure $H \rightarrow \tau\tau$ events, to deal with the missing momentum from neutrinos of the subsequent tau

decays. These methods are the transverse mass method, the collinear approximation and the missing mass calculator (MMC). Second: Exploiting properties of some kinematic variables helps to improve the assignment of the light leptons.

The thesis is structured as follows. Chapter 2 introduces theoretical background and concepts, like special relativity and quantum field theory. It also introduces the particles of the Standard Model and presents the recent status of research about y_t. In chapter 3, the LHC and ATLAS are presented, with emphasis on the object reconstruction in ATLAS. The analysis setup of tHq events in the $2\,e/\mu + 1\tau_{\text{had}}$ channel, including the selection cuts and the general strategy of mass reconstruction, is explained in chapter 4. Chapter 5 reports about the correlation studies for predicting the four-momentum of the neutrino from the top quark and shows results of the improvements of the mass reconstruction. Results of the improvements in the assignment of the light leptons are shown in chapter 6. The thesis ends with a conclusion in chapter 7.

CHAPTER 2

Theoretical background and concepts

This chapter is meant to introduce both physicists as well as non-physicists in the shortest possible way into the domain of the thesis. Therefore section 2.1 gives a short idea of the worldview of Particle Physics. Section 2.3 introduces the elementary particles according to the Standard Model (SM). The sections 2.2 and 2.4 introduce basic concepts of special relativity and quantum field theory, which are necessary to know in order to understand the thesis. These sections are not meant to give deep understanding but are to remind a physicist of these concepts and also to convey an idea thereof to non-physicists. This gives just enough to understand the motivation in section 2.5 about the work reported in this thesis.

2.1 Basic concepts about particle physics

This thesis belongs to the field of elementary particle physics, a branch of research, which explores the smallest pieces of matter and their interactions. The idea that matter consists of small, indivisible pieces, called atoms (from ancient Greek "atomos", which means "uncuttable"), is found already in ancient Greek philosophy [2]. When natural sciences had been elaborated enough to handle the question, what matter consists of, it became possible to divide matter into smaller and smaller pieces and to find particles, which were called atoms, relating to the ancient Greek idea. Meanwhile, it has been found out that atoms consist of other particles, of which some seem to be point-like and are called elementary particles. Although the atoms, however, are not point-like but composite particles, yet the physics of a single atom vastly differs from that describing the everyday life with large pieces of matter, composed of many ($\approx 10^{23}$ and more) atoms.

2.1.1 Quantum mechanics

The physical laws describing atoms and elementary particles, are not the classical mechanical laws, known for macroscopic objects and also experienced in everyday life, but quantum mechanics, a theory developed in the 1920s [3]. A quantum-mechanical state can be defined as a set of particles of a certain type and with certain momenta. Quantum mechanics describes the time evolution of those states and can give probabilities for a system to change its state. This description is not deterministic. Physical observables are expressed as expectation values, which are calculated using a complex-valued wave function, whose magnitude gives the probability density for the system to be in a certain state [4]. This limitation to a probabilistic description does not appear to be a shortcoming of the theory but a

property of nature [5]. Other limitations of quantum mechanics in its original formulation are that it is only applicable in a non-relativistic limit, i.e. for systems with speeds $v \ll c$ and for a fixed number of particles. Those limitations can be overcome by quantum field theory, an extension of quantum mechanics, which has developed over decades and proved itself up to date. This is why basic concepts of QFT are presented in section 2.4.

Spin

Quantum mechanics gave rise to a new physical quantity, named spin. It has some analogies to angular momentum, with the essential difference that spin can be assigned to elementary, i.e. point-like particles, which should not be able to have inner angular momentum in its original meaning. Furthermore, it turned out that all known particles can be divided into two types: bosons[1] are all particles with integer spin quantum number (0, 1, 2, ...) and fermions[2] have half-integer spin quantum number (½, ³⁄₂, ⁵⁄₂, ...).

2.1.2 Particle decays

Not all particles that have been discovered are stable. Many of them (both composite and point-like) decay into lighter particles following an exponential law $N(t) = N_0 \cdot \mathrm{e}^{-\frac{t}{\tau}}$, i.e. after passing a time of τ the number of particles has decreased by a factor of $1/\mathrm{e} \approx 0.37$. Often the decay can occur into several final state particles. In these cases, a branching ratio $\mathcal{B}$ is defined to indicate, which final state occurs how often relative to all decays of that particle.

2.1.3 Fundamental forces

Four fundamental forces of interaction are known: the gravitation, the electromagnetic (generalized description of both electric and magnetic phenomena), the weak and the strong interaction. In experimental setups of elementary particle physics, gravitation is much weaker than the other forces and is neglected. The three other interactions can be described by quantum field theory. The electromagnetic and the weak interaction can be described by a unified electroweak quantum field theory (short "electroweak theory" - EWT). Since the strong interaction applies to all particles with a special type of charge, called color-charge, the quantum field theory for the strong interaction is called quantum chromodynamics (QCD). The strong interaction is the only fundamental force that increases by increasing the distance between the participating particles. For this reason, strongly interacting particles are not found alone but form bound states very quickly. These bound states are particles called hadrons.

2.1.4 Basic quantities for observation of particles and their interactions

Observation of particles and their interactions

In a macroscopic understanding a common observation consists of two things: firstly, directing a light source onto the object and secondly, detecting the light, which has been reflected by the object. Although it is somehow sensible, it ignores (or at least neglects) the interaction of the light with the object. In the microscopic world, it turns out that the interaction due to observation cannot be ignored anymore.

[1] named after Indian physicist Satyendra Nath Bose (1894 - 1974)
[2] named after Italian physicist Enrico Fermi (1901 - 1954)

Light also consists of particles, namely photons, which interact with the particles to be observed and change their state through the act of observation. Observation without interaction is not possible. Being a limitation on the one hand side, it generalizes the concept of observation and opens the use of other particles as an observation medium. In the end, observation is directing a particle source (photons or particles of a different kind) onto a target (consisting of the particles to be observed) and detecting the particles after this interaction. The distinction between the medium of observation (light) and the object (particles) vanishes. Therefore, experimental particle physics focuses on accelerating and detecting particles. One way to do this is to shoot particles onto a fixed target (fixed target experiments). Another way is letting beams of particles collide (colliding beam experiment) [6]. The main advantage of a colliding beam experiment in contrast to a fixed-target experiment is that there is no energy loss due to recoil.

Basic quantities

The more particles are accelerated onto a target, the higher the rate of interactions with the target , $\dot{N}$, is expected to be. The amount of particles instigated for interaction per time is quantified by the luminosity $\mathcal{L}$, measured in $\mathrm{cm}^{-1}\mathrm{s}^{-1}$, which is proportional to the number of beam particles per geometrical beam cross section and time. The ratio of the measured value of $\dot{N}$ divided by the controlled value of $\mathcal{L}$ is denoted with σ and called the cross section of the interaction.[3] For its measurement, the unit "barn" has been introduced: $1\,\mathrm{b} = 1 \times 10^{-28}\,\mathrm{m}^2$. Equivalently, σ can be expressed as the number N of detected interactions over a certain amount of measurement time divided by the integral of the luminosity over the measurement time, which is simply called integrated luminosity, $\int \mathcal{L}\mathrm{d}t$.

$$\sigma = \frac{\dot{N}}{\mathcal{L}} = \frac{N}{\int \mathcal{L}\mathrm{d}t} \tag{2.1}$$

Cross section

If the interaction of the accelerated particles with the target was just a mechanical collision of macroscopic objects, this cross section of the interaction would be the geometrical cross section area of the target. For microscopic targets, however, this geometrical interpretation does not hold for two reasons. Firstly, for a microscopic target, which can be a point-like particle, its area is not well-defined. Secondly, often more interactions than just mechanical scattering are possible. For each of these possible interactions, a cross section of the respective interaction is assigned. So σ has to be understood in this analogy, not as an actual area. What can be deduced anyway is that the higher σ the more probable the interaction takes place. Quantum field theory follows the aim of calculating σ with respect to any kind of interaction of elementary particles.

[3] Actually it is just called "cross section". Here it is named "cross section of the interaction" to distinguish it from the geometrical cross section.

2.2 Basic concepts of special relativity

2.2.1 Mass-energy equivalence

Known from classical mechanics, the energy, E, and the momentum, p, are constants of motion. In contrast to classical mechanics, where the mass, m, is a constant of motion as well, special relativity allows the transformation of mass into energy and vice versa, described by the famous formula

$$E = mc^2, \tag{2.2}$$

where $c \approx 3 \times 10^8\,\mathrm{m/s}$ is the speed of light in vacuum. Equation 2.2 also allows the transformation of the energy of colliding particles into the formation of new particles. With higher collision energy, particles of higher mass can be produced.

2.2.2 Frame of reference

Measurements of energy and momentum depend on their frame of reference. A frame of reference in which the lab is at rest is called a lab frame. A frame in which the total momentum of all considered particles vanishes is called a center-of-momentum frame. The special case that the origin of the coordinate system also lies in the center-of-mass of the considered particles is called a center-of-mass frame. Special relativity states that all frames are equivalent (if used consistently). The speed of light, however, is constant and the maximum speed ever possible to reach. This affects the addition of velocities and the measurement of space and time, leading to effects like time dilation and length contraction, which name effects that time and length appear different in different frames of reference [7]. The physics is consistent in each frame of reference, but for this, all observables have to be measured in the same frame of reference. There is no absolute frame of reference that could be preferred above the others. This is why it is called "relativity", although it sets an absolute maximum of all speeds, the speed of light in vacuum.

2.2.3 Units used in particle physics

Because the speed of light in vacuum is constant, it is common practice to omit the factor c in formulas. The same is done for the Planck constant $\hbar$ and sometimes also for the Boltzmann constant k_B. In other words, a new system of units is defined, called "natural units", in which all these quantities are set to 1:

$$c = \hbar = k_\mathrm{B} = 1$$

Using formula 2.2, masses can be expressed in units of energy. In particle physics it is common practice to express particle masses in the unit $\mathrm{GeV} = \mathrm{GeV}/c^2 \approx 1.78 \times 10^{-27}\,\mathrm{kg}$.

2.2.4 Lorentz transformation

Transformations from one frame of reference into another, which only consist of a relative motion with uniform velocity, are called Lorentz boosts (short "boosts"). General Lorentz[4] transformations also allow

[4] named after the Dutch physicist Hendrik Antoon Lorentz (1853 – 1928), not to be confused with the Danish physicist Ludvig Valentin Lorenz (1829 – 1891)

rotations. As explained above, in general, physical quantities change under Lorentz transformations, i.e. they are not "Lorentz-invariant". But some quantities are, for example, all those scalar quantities which are calculated via the special Minkowski scalar product of four-vectors.[5] Those four-vectors must be elements of the same Minkowski space, i.e. they must be filled with observables measured in the same frame of reference.

2.2.5 Energy-momentum relation

A popular example of a four-vector consists of energy and the three Cartesian coordinates of momentum. Such a four-vector is also called four-momentum, in order to distinguish it from the classical three-momentum, $\vec{p}$. For distinction, four-momenta are denoted with the upper case letter P in this thesis. The scalar product calculated with any two four-momenta $P_a = (E_a, p_a^x, p_a^y, p_a^z)$ and $P_b = (E_b, p_b^x, p_b^y, p_b^z)$ is Lorentz-invariant and defined as follows:

$$P_a P_b = E_a E_b - \vec{p}_a \vec{p}_b = E_a E_b - p_a^x p_b^x - p_a^y p_b^y - p_a^z p_b^z \tag{2.3}$$

Of special interest is the scalar product of a four-momentum with itself, which gives the square of the mass of a physical system (of particles). It is also called invariant mass.

$$P_a P_a = E_a^2 - \vec{p}_a^2 = E_a E_a - p_a^x p_a^x - p_a^y p_a^y - p_a^z p_a^z = m_a^2 \tag{2.4}$$

Equation 2.4 is also called the energy-momentum relation.

2.2.6 Mandelstam variables

Other important quantities calculated using four-momenta are the Mandelstam variables[6] s, t and u, which are defined for an interaction with two particles before interaction (named 1 and 2) and two particles after interaction (named 3 and 4). They are calculated using the respective four-momenta P_1, P_2, P_3 and P_4:

$$\begin{aligned} s &= (P_1 + P_2)^2 = (P_3 + P_4)^2 \\ t &= (P_1 - P_3)^2 = (P_2 - P_4)^2 \\ u &= (P_1 - P_4)^2 = (P_2 - P_3)^2 \end{aligned} \tag{2.5}$$

$\sqrt{s}$ is identical to the total energy, which is conserved throughout the interaction. The Mandelstam variables are also referred to in section 2.4.2.

2.2.7 Rapidity and pseudorapidity

The rapidity is another Lorentz-invariant quantity. In general, it is defined by:

$$y = \frac{1}{2} \ln \left(\frac{E + |\vec{p}|}{E - |\vec{p}|} \right) . \tag{2.6}$$

[5] The Minkowski scalar product, named after Hermann Minkowski, German mathematician (1864 – 1909), is not a scalar product in a strict mathematical sense, because it is not positive semidefinite.

[6] named after the American physicist Stanley Mandelstam (1928 - 2016)

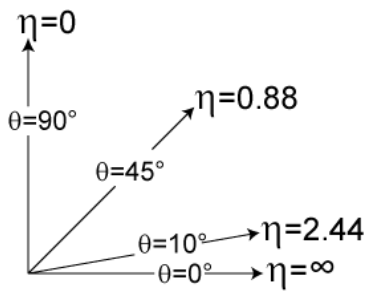


Figure 2.1: Relation between polar angle θ and pseudorapidity η, as calculated by formula 2.8. Taken from [8].

In particle collider physics, where the beam goes along the z-axis, it is common to use p^z, the z-component of the momentum, instead of $|\vec{p}|$, in the definition:

$$y = \frac{1}{2} \ln \left(\frac{E + p^z}{E - p^z} \right) \tag{2.7}$$

With the new definition, y is Lorentz-invariant under boosts along the beam axis. Starting from the latter definition, a similar quantity, called pseudorapidity, is defined as:

$$\eta = \frac{1}{2} \ln \left(\frac{|\vec{p}| + p^z}{|\vec{p}| - p^z} \right) = -\ln \left(\tan \frac{\theta}{2} \right). \tag{2.8}$$

η is Lorentz-invariant under boosts in z-direction in the approximation $|\vec{p}| \approx E \gg m$ and especially convenient because of the direct relation to the polar angle, θ (shown in figure 2.1). The polar angle is the angle between the total three-momentum-vector and the beam axis (in z-direction) and can be expressed as: $\cos\theta := \frac{p^z}{|\vec{p}|}$.

2.3 The Standard Model (SM) for elementary particles

Figure 2.2 shows name, mass, electric charge and spin of the elementary particles of the Standard Model. They can be arranged into different groups, which are discussed in the following subsections.

decay channel	branching ratio
$W \to e\nu_e$	$(10.71 \pm 0.16)\,\%$
$W \to \mu\nu_\mu$	$(10.63 \pm 0.15)\,\%$
$W \to \tau\nu_\tau$	$(11.38 \pm 0.21)\,\%$
$W \to$ hadrons	$(67.41 \pm 0.27)\,\%$

Table 2.1: Most prominent decay modes of the W boson [10].

2.3.1 Gauge bosons

In the formalism of quantum field theory, interactions between particles are transmitted by a special type of particles, which are called gauge bosons and marked in red in figure 2.2. Each gauge boson has spin 1

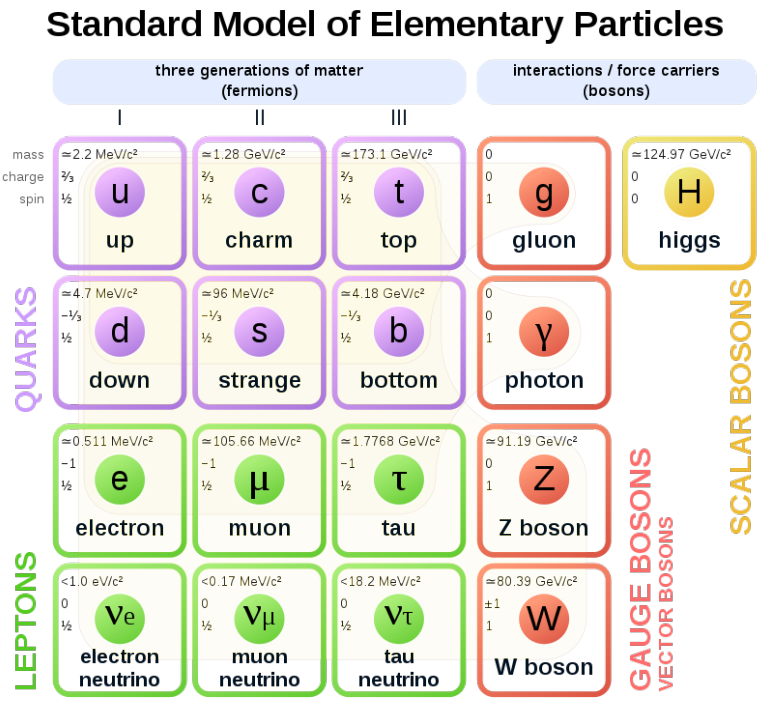


Figure 2.2: All elementary particles known to the Standard Model (SM) arranged in a typical pattern [9]. Antiparticles also belong to the SM but are not listed explicitly.

decay channel	branching ratio
$Z \rightarrow e^+e^-$	(3.3632 ± 0.0042) %
$Z \rightarrow \mu^+\mu^-$	(3.3662 ± 0.0066) %
$Z \rightarrow \tau^+\tau^-$	(3.3696 ± 0.0083) %
$Z \rightarrow$ neutrinos	(20.000 ± 0.055) %
$Z \rightarrow$ hadrons	(69.911 ± 0.056) %

Table 2.2: Most prominent decay modes of the Z boson [11].

(and is therefore called a vector boson) and transmits a certain type of interaction. The gluon, g, transmits the strong interaction and couples to all color-charged particles. These are the quarks and the gluon itself. The self-interaction of the gluon makes that, in general, calculations in quantum chromodynamics (QCD) are difficult to perform. The photon, γ, transmits the electromagnetic interaction, which is a unification of electric and magnetic interaction and is described by quantum electrodynamics (QED). The photon couples to all particles with electric charge (short "charge"). Figure 2.2 show that this applies to all fermions except for the neutrinos and - as the only gauge boson - also for the W boson. The Z boson and the W boson transmit the weak interaction, which applies to all fermions. They are the only massive and unstable gauge bosons. Branching ratios of their most prominent decay modes are shown in tables 2.1 and 2.2 respectively.

particle name	particle symbol	decay time	mass
proton	p	stable	938.3 MeV
neutron	n	8.8×10^{2} s	939.6 MeV
pion (charged)	$\pi^{\pm}$	2.6×10^{-8} s	139.6 MeV
pion (neutral)	π^{0}	8.5×10^{-17} s	135.0 MeV
kaon (charged)	$K^{\pm}$	1.2×10^{-8} s	493.7 MeV
B-meson (charged)	$B^{\pm}$	1.6×10^{-12} s	5.3 GeV
B-meson (neutral)	B^{0}, $\bar{B}^{0}$	1.5×10^{-12} s	5.3 GeV
electron	e	stable	0.51 MeV
muon	μ	2.2×10^{-6} s	105.7 MeV
tau	τ	2.9×10^{-13} s	1.78 GeV
top quark	t	$\approx 5 \times 10^{-25}$ s	173.0 GeV
Higgs boson	H	$\approx 1 \times 10^{-22}$ s	125.18 GeV
W boson	W	3.2×10^{-25} s	80.38 GeV
Z boson	Z	2.6×10^{-25} s	91.19 GeV

Table 2.3: Decay times and masses of some prominent particles without all decimal places and without uncertainty for simplicity [12].

2.3.2 Fermions in general

Over the years, many particles have been found, but most of them are composed of elementary particles and the number of different types of elementary particles is still manageable. They can be arranged in a certain pattern, which can be seen on the left hand side of figure 2.2. They are all fermions with spin ½ and come in three generations, which can be subdivided into quarks and leptons. Quarks carry both electric and color charge and are the only fermions interacting through the strong interaction. For each generation, there is one up-type quark (up, charm, top) with electric charge ⅔ e and one down-type quark (down, strange, bottom) with charge -⅓ e, where e is the elementary charge. These six different quark types are also called quark flavors. Leptons have no color charge. They can be subdivided into electrically charged leptons (electron, muon and tau) and electrically neutral neutrinos. Considering only the leptons, the three generations are called flavors as well, and the name of each flavor is identical to the name of the charged lepton of the corresponding generation. The electrically charged leptons interact electromagnetically and weakly, while the neutrinos only interact weakly. Every fermion has a corresponding antiparticle (antifermion) with opposite charges (both electric charge and color charge), but with the same mass. These antiparticles are not listed in the pattern explicitly.

2.3.3 Leptons

Charged leptons

The electron is the lightest and only stable charged lepton. The muon is heavier and decays into a muon-neutrino and a pair of electron and electron-antineutrino. The tau is the heaviest charged lepton. It can decay either leptonically, i.e. into leptons, or hadronically, i.e. into hadrons. In both cases, one decay product is a tau-neutrino. In the case of a leptonic decay, additional decay products are either an electron and an electron-neutrino or a muon and a muon-neutrino. In the case of a hadronic decay, pions

decay channel	branching ratio
leptonic	35.21 %
$\tau \rightarrow \mu \bar{\nu}_\mu \nu_\tau (+\gamma)$	(17.39 ± 0.04) %
$\tau \rightarrow e \bar{\nu}_e \nu_\tau (+\gamma)$	(17.82 ± 0.04) %
hadronic	63.08 %
$\tau \rightarrow \pi^- \nu_\tau$	(10.82 ± 0.05) %
$\tau \rightarrow K^- \nu_\tau$	(0.70 ± 0.01) %
$\tau \rightarrow h^- \nu_\tau$ + neutrals	(37.01 ± 0.09) %
$\tau \rightarrow h^- h^+ h^- \nu_\tau$ + neutrals	(14.55 ± 0.06) %

Table 2.4: Branching ratios of most prominent tau decay modes [13], excluding events with neutral Kaons. h^- stands for π^- or K^-, neutrals means γs and/or π^0s.

and/or kaons as decay products are possible. Branching ratios of the most prominent tau decay modes are listed in table 2.4.

Neutrinos

According to the SM, neutrinos are stable and massless. But at least a small mass is necessary to explain the phenomenon called neutrino oscillation. The name of this phenomenon refers to the observation that neutrinos can oscillate between different flavors during propagation. Though in conflict with the Standard Model, evidence for this phenomenon can not be dismissed and was rewarded by a Nobel prize in 2015 [14]. This is one of the hints, and maybe the most prominent one, that the SM cannot explain everything. On the other hand, the neutrino masses are so small that they can be neglected in any practical concern.

2.3.4 Quarks and hadronization

Since the force of the strong interaction increases with the distance of participating particles, quarks never appear alone but are always bound in special types of particles, called hadrons. The proton and the neutron are the most prominent representatives of hadrons, of which the proton is the only stable one.[7] Since atomic nuclei are formed out of those, they are also called nucleons. The next prominent representatives are pions and kaons. Decay times of some prominent particles are listed in table 2.3. The proton consists of two up quarks and one down quark. Vice versa, the neutron consists of two down quarks and one up quark, resulting in a total charge of +1 for the proton and 0 for the neutron. Because of the complicated structure of the strong interaction, allowing self-interaction of gluons, the inner structure of hadrons is much more complicated, containing additional gluons and a "quark-sea" of quark-antiquark-pairs also of other flavors. For this reason, the aforementioned three quarks responsible for the total electric charge of the proton are called valence quarks, in order to distinguish them from the "sea quarks". The general term "partons" addresses all constituents of a hadron, including all quarks and antiquarks and also the gluons. If, by a particle interaction, a quark is dissociated from a hadron, the field energy between these quarks increases to such an extent that a new quark-antiquark-pair emerges, which allows the formation of new hadrons. This process is called hadronization. It can also be initiated by a dissociated gluon which splits up into a quark-antiquark-pair.

[7] Neutrons can only be stable when bound in an atomic nucleus. Alone, they decay with a decay time of ≈ 15 min.

decay channel	branching ratio
$t \to Wb$	
$t \to e\nu_e b$	13.3 %
$t \to \mu\nu_\mu b$	13.4 %
$t \to \tau\nu_\tau b$	7.1 %
$t \to q\bar{q}b$	66.5 %

Table 2.5: Branching ratios of top-quark decay modes relative to the main top-quark decay $t \to Wb$, which covers 95.7 % of all top-quark decays [16]. The branching ratio corresponding to $t \to \tau\nu_\tau b$ refers only to hadronically decaying taus. Leptonically decaying taus are included in the branching ratios of the decays $t \to e\nu_e b$ and $t \to \mu\nu_\mu b$.

Parton Distribution Functions (PDF)

Up to now, there is no theory predicting what fraction of the hadron's momentum is carried by which parton. However, there is a lot of data gained from various measurements, like scattering electrons or neutrinos off protons or neutrons and also high-energy collider data [6]. Based on those measurements, probability density functions, called parton distribution functions (PDF), are reconstructed, which indicate for each type of parton the probability to carry a certain fraction of momentum. For this purpose, all partons are assumed to be massless. Depending on the energy scale, at which the calculations are performed, this approximation is good enough for the four lightest quarks (4-flavor-scheme, short "4F-scheme"). This does also mean that the proton is assumed to not contain any b-quarks in the quark sea. A b-quark originating from the proton therefore must be assumed to originate from a gluon splitting into a $b\bar{b}$-pair. Alternatively, in the 5-flavor-scheme (short "5F-scheme"), the PDF also contains b-quarks (although its mass is clearly higher than the masses of the other quarks) [15].

The top quark

The top quark is the only quark that does not hadronize, because, with a decay time of $\approx 5 \times 10^{-25}$ s, it decays too fast. With a branching ratio of $\mathcal{B} = 0.957$, it almost always decays into a b-quark and a W boson. All subsequent decays of this decay are listed in table 2.5 with its branching ratios relative to the decay $t \to Wb$. Because of the short lifetime of the tau, subsequent tau decays into an electron or a muon are also assigned to the direct decays $t \to e\nu_e b$ and $t \to \mu\nu_\mu b$ respectively. Thus, the branching ratio corresponding to $t \to \tau\nu_\tau b$ refers only to hadronic tau decays. This is the main reason why the branching ratios of the three leptonic decays of the top quark are not equal. Assigning all $t \to \tau\nu_\tau b$-decays to the tau, irrespective of the subsequent tau decay, would lead to very similar branching ratios of all three lepton flavors of ≈ 11 %. Contrary to the top quark, the b-quark does hadronize. The most prominent b-hadrons, the light B-mesons, have a lifetime of $\approx 1.5 \times 10^{-12}$ s (neutral B-mesons) and $\approx 1.6 \times 10^{-12}$ s (charged B-mesons). They all have a similar mass of ≈ 5.3 GeV.

2.3.5 Higgs boson

The Higgs boson has spin 0, is therefore called a scalar boson and does not transmit any force. Its existence is postulated by the Higgs mechanism. The Higgs mechanism interprets the masses of the other Standard Model particles as couplings to the Higgs boson, which allows the interpretation of the other SM particles to be intrinsically massless. This interpretation is needed because massive particles

decay channel	branching ratio
$H \to b\bar{b}$	57.1 %
$H \to W^+W^-$	22.0 %
$H \to gg$	8.53 %
$H \to \tau^+\tau^-$	6.26 %
$H \to c\bar{c}$	2.88 %
$H \to ZZ$	2.73 %
$H \to \gamma\gamma$	0.23 %

Table 2.6: SM predictions for the branching ratios of most dominant Higgs-boson decay modes for a Higgs mass of $m_H = 125.36\,\text{GeV}$ [19], obtained by linear interpolation of m_H-dependent branching ratios reported in [20].

would violate gauge invariance. The principle of gauge invariance is the basis of the description of interactions through gauge bosons. Without gauge invariance, this description and thus the whole theoretical foundation of the SM would be inconsistent.[6] That it was found in 2012 [17] as the latest of all SM particles was very important therefore. SM predictions for the branching ratios of most dominant Higgs-boson decay modes are shown in table 2.6 for a Higgs mass of $m_H = 125.36\,\text{GeV}$. The current global fit according to [18] gives $m_H = (125.25 \pm 0.17)\,\text{GeV}$.

2.4 Basic concepts of quantum field theory

As known from quantum mechanics, still in quantum field theory, one or more particles build a quantum mechanical system. The interaction between those particles changes the state of this system. From now on, the transition of a quantum mechanical system from initial state $|i\rangle$ to final state $\langle f|$ is called a process. It includes all possible ways, how this transition can happen. A specific process is therefore identified with its initial and final state. For the distinction between various processes with the same initial state, they are often named only after their final state. With quantum field theory, the probability for a process to happen can be calculated.

2.4.1 The matrix element

A state of a quantum mechanical system can mathematically be described as a vector in an abstract vector space.[8] Quantum field theory (QFT) is a perturbative theory, i.e. transitions from one state to another are treated as perturbations of the initial state. These transitions can be described by a matrix $\mathcal{M}$. Choosing the initial state $|i\rangle$ and final state $\langle f|$ of the interaction as basis vectors, the probability of this transition is encoded in one specific matrix element $\mathcal{M}_{fi} = \langle f|\,\mathcal{M}\,|i\rangle$, which is a complex number: $\mathcal{M}_{fi} = |\mathcal{M}_{fi}|e^{i\phi}$, i.e. it consists of a magnitude $|\mathcal{M}_{fi}|$ and a phase ϕ.

Calculating the matrix element

The perturbations, which describe the transition from one state to another, have no analytically exact solution but are described by a series expansion. Therefore, solutions for the matrix element are approximate corresponding to different orders in a series expansion. They are called leading order (LO),

[8] Hilbert space is used for fixed numbers of particles and Fock space for variable numbers of particles

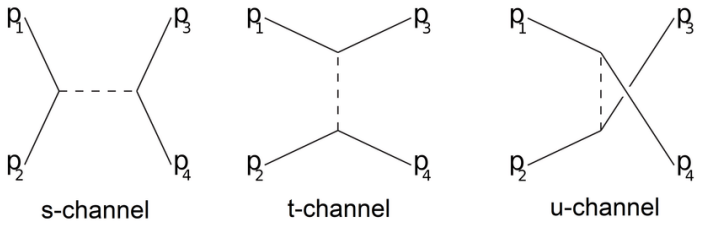


Figure 2.3: Simplified Feynman diagrams showing the characteristic shape of s-channel [21], t-channel [22] and u-channel [23]. p_1 and p_2 are the four-momenta of the initial state particles, p_3 and p_4 of the final state particles. The four-momentum of the corresponding virtual particle (dashed line) is calculated according to equation 2.5.

next-to-leading order (NLO), next-to-next-to-leading order (NNLO) and so on (and for each order you get one matrix element which is added up to the matrix element of the previous order). The series is expanded on the vertex factors (explained in section 2.4.2), which are proportional to the coupling strength of the interaction. Fortunately, it is often sufficient to consider just the leading order of this expansion. Even then, there is more than one way how the transition between the initial and the final state could have happened. In that case, the matrix elements of all possible transitions have to be calculated, and the magnitude of the sum of all those. The relative phases of each matrix element thus play a role in the final magnitude. In analogy to the superposition of waves, where the relative phase between the waves determines the result of the superposition, this is called interference.

2.4.2 Feynman diagrams

In order to make it more clear which possible transitions between an initial and a final state are considered in a certain calculation for a process, Feynman[9] diagrams are used. Using these Feynman diagrams, it has become practice to speak of which diagrams contribute to a certain process. In a Feynman diagram, the horizontal axis, going from left to right, indicates the direction of time, and the vertical axis, going from bottom to top indicates space.[10] So the initial state is on the left hand side and the final state is on the right hand side. A particle propagating through space and time is depicted by a line, fermions with solid lines (with an arrow going in positive time-direction for particles and into negative time-direction for antiparticles), gauge bosons of the electroweak theory with wiggled lines (γ, Z, W), the gluon as the gauge boson of quantum chromodynamics with curly lines and scalar bosons (in the Standard Model only the Higgs boson) with dashed lines. The name of the corresponding particle can be written next to the line as a mathematical symbol. A point where those lines meet is called a vertex and is indicating that these particles couple to each other. A symbol for the coupling strength can be drawn next to such a vertex. Two possible Feynman diagrams of the same process (i.e. with the same initial and final state) are shown in figure 2.6.

Virtual particles Particles, which only appear between two vertices, but not in the initial and final state are called "virtual" particles. As there is four-momentum conservation at each vertex, the four-momentum of the virtual particles is represented by one of the Mandelstam variables. Depending on

[9] named after American physicist Richard Feynman (1918 - 1988)

[10] As the axis represents all three spatial dimensions, there is no scale of measure on this axis in the classical meaning. But the arrangement of the particles along the axis is important for the correct naming of the four-momenta used for calculating the Mandelstam variables.

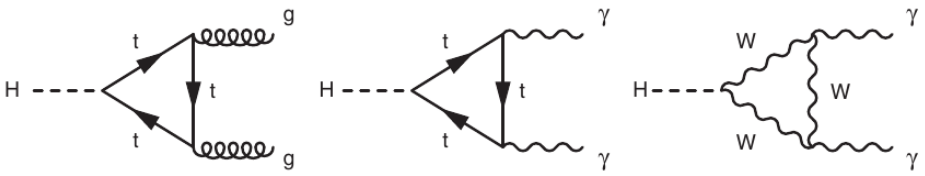


Figure 2.4: Decay of the Higgs boson into massless gauge bosons. Taken from [6].

which Mandelstam variable assigns the correct four-momentum to the virtual particle, the process is called s-channel, t-channel or u-channel. This can be seen in figure 2.3.

Energy-momentum relation for virtual particles While four-momentum conservation also holds for virtual particles, the energy-momentum relation does not necessarily apply to them. If it does, the virtual particles are called "on-mass-shell" (short "on-shell") and if not, they are called "off-mass-shell" (short "off-shell"). However, if the conservation of energy and momentum at the vertices allows, virtual particles are assumed to be "on-shell", because the matrix element, which depends on the energy and the momentum of the virtual particle, has a sharp peak in that case and thus probability is maximized.

2.4.3 Phase space

Since the initial and the final state of an interaction are - among others - also defined by their momenta, the matrix element describing the transition between those states depends on the momenta as well. Often the momenta of the final state are not of interest. Therefore, the transition from a state into one with arbitrary momenta is considered. As the components of all involved momenta are linearly independent they span a vector space, called phase space ρ. The volume of this phase space is restricted by conservation laws, like the conservation of momentum and energy. The higher the volume the more final states are allowed and the more likely the transition will happen. In order to calculate a cross section for the interaction for any momenta in the final state to happen, the magnitude of this matrix element has to be integrated over the whole allowed phase-space-volume:

$$\sigma = \int |\mathcal{M}_{fi}|^2 \, \mathrm{d}\rho \tag{2.9}$$

2.5 Production of a Higgs boson in association with top quarks

2.5.1 Higgs couplings to gauge bosons and fermions

For the gauge bosons $V \in \{ W, Z \}$, the relation between their mass, m_V, and their coupling to the Higgs, g_{HVV}, is

$$g_{HVV} = 2 \frac{m_V^2}{v}, \tag{2.10}$$

with the vacuum expectation value of the Higgs field, $v = 246\,\mathrm{GeV}$. In accordance with this formula, the massless gauge bosons g and γ do not couple to the Higgs, but an indirect coupling is possible through loops with the top quark, or the W boson, as shown in figure 2.4. The relation of the couplings of the

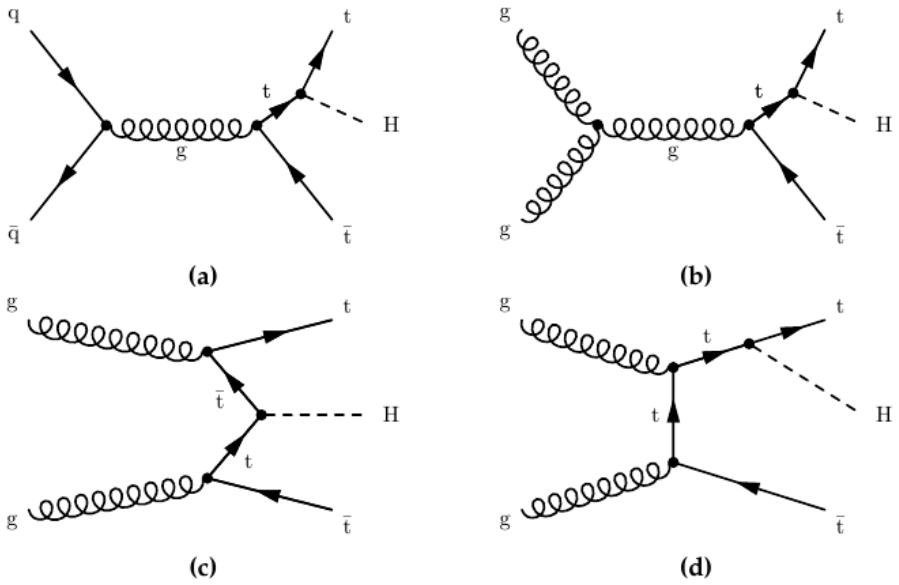


Figure 2.5: Feynman diagrams for $t\bar{t}$ production. Taken from [25].

fermions, y_f, to their masses, m_f, is given by

$$y_f = \sqrt{2}\frac{m_f}{v}. \tag{2.11}$$

According to the SM, the couplings g_{HVV} and y_f enter into the Feynman diagram with opposite sign [24]. In order to test the Higgs mechanism for its consistency, it is of course also necessary to investigate all these couplings, i.e. search for processes in which a Higgs boson is produced in association with one of those particles or decays into them.

It attracts some attention that for the heaviest fermion, the top quark, this formula gives

$$y_t = \sqrt{2}\frac{173\,\text{GeV}}{246\,\text{GeV}} = 0.995, \tag{2.12}$$

a value that is very close to 1. Investigating y_t could answer the question if this is a coincidence. In order to express experimental deviation from this theoretical prediction, coupling modifiers

$$\kappa_t = \frac{y_t}{y_t^{\text{SM}}} \qquad \text{and} \qquad \kappa_V = \frac{g_{HVV}}{g_{HVV}^{\text{SM}}} \tag{2.13}$$

are introduced [1]. Since the top quark is heavier than the Higgs boson, this coupling can not be studied through decay but only through associated production.

2.5.2 Higgs boson in association with a $t\bar{t}$-pair

The top quark associated Higgs boson production with the highest cross section is the Higgs production associated with a $t\bar{t}$-pair. It was found in 2017 by both the CMS [26] and the ATLAS [27] collaboration.

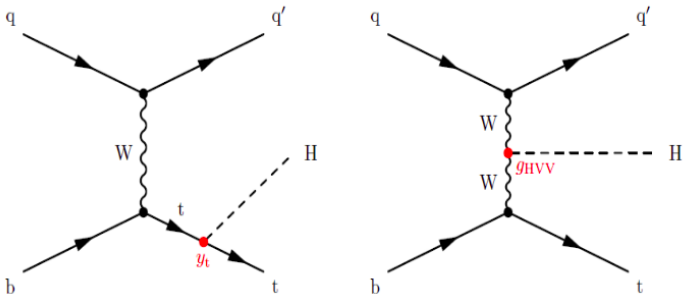


Figure 2.6: Two leading-order Feynman diagrams of tHq. Taken from [1].

Assuming SM branching ratios, the cross section has been determined by the ATLAS collaboration to be

$$\sigma_{t\bar{t}H} = 670 \pm 90\,(\text{stat.})\,^{+110}_{-100}\,(\text{sys.})\,\text{fb}$$

at 13 TeV in agreement with the Standard Model predictions [28], calculated up to NLO. Possible leading order Feynman diagrams are shown in figure 2.5. Since every diagram includes the coupling y_t exactly once, there are no interference effects and the cross section is [20]:

$$\sigma_{t\bar{t}H} = \kappa_t^2 \sigma_{t\bar{t}H}^{\text{SM}} \tag{2.14}$$

As can be seen from equation 2.14 the sign of the coupling y_t can not be determined by investigating this process. Single top quark associated Higgs boson production is needed for that.

2.5.3 Higgs boson in association with a single top quark

The production of a single top quark associated Higgs boson must - like the production of a single top quark- include a tWb-vertex. The Higgs boson is either radiated off the top quark or radiated off the W boson. Since the top quark must be in the final state, possible s-channel production modes have either an additional b-quark or an additional W boson in the final state. The latter production mode is referred to as tWH. The single top quark associated Higgs boson production with the highest cross section, however, is a t-channel production with an additional quark q of irrelevant flavor in the final state. It is therefore called tHq. Still, the cross section for tHq is much lower than for $t\bar{t}$. Calculated in a 4F-scheme on an energy scale, which is $\mu_0^s = (m_H + m_t)/4$ constantly for all events ("static scale"), the SM predicts at NLO [29]

$$\sigma_{tHq} = 69.43 \pm 0.07\,(\text{numeric})\,\text{fb}.$$

Combining 4F- and 5F-schemes on the same scale and combining tHq with $\bar{t}Hq$, the cross section up to NLO is

$$\sigma_{tHq} = 72.55 \pm 0.10\,(\text{numeric})\,\text{fb}.$$

Leading order Feynman diagrams for tHq are shown in figure 2.6. Since in one diagram, only y_t appears, and in the other diagram, only g_{HVV} appears, the interference of these diagrams makes the relative sign of these couplings contribute to the cross section. Using the coupling modifiers defined in equation 2.13,

the effect on the cross section calculated in leading order and in 4F-scheme by [1] is

$$\sigma_{tHq} = (2.63\kappa_t^2 + 3.58\kappa_V^2 - 5.21\kappa_t\kappa_V)\,\sigma_{tHq}^{\mathrm{SM}}. \tag{2.15}$$

In principle, this can be tested experimentally by calculating the cross section with formula 2.1, using the experimentally determined value for the integrated luminosity and the experimentally determined number of tHq events. In practice, counting the number of tHq events is a big challenge because the events have to be reconstructed and separated from background events. By now, only limits on κ_t can be given: for an SM value of g_{HVV}, i.e. $\kappa_V = 1.0$, data favor positive values of κ_t and exclude values of κ_t outside the ranges of about $[-0.9, -0.5]$ and $[1.0, 2.1]$ at 95 % confidence level [1]. So far, data are consistent with the SM prediction. This thesis is a contribution to the analysis of tHq, i.e. its reconstruction and separation from background processes. Since tHq is a very rare process, its analysis is promised to be very challenging. However, due to the functional form of equation 2.15, any deviation from the SM increases the cross section and thereby the expected number of tHq events.

CHAPTER 3

Experimental setup

This chapter introduces the experimental setup, where data about the process of interest, tHq, can be collected. In order to provoke this process, a particle accelerator (LHC) is used, which is described in section 3.1. For the detection of the final state particles, a detector (ATLAS) is needed, presented in section 3.2. The reconstruction of the various objects of the final state on the basis of the detected signals of ATLAS is explained in section 3.4.

3.1 The Large Hadron Collider (LHC)

3.1.1 Overview

The Large Hadron Collider is an approximatley circular particle accelerator run by the European Organization for Nuclear Research (short: CERN[1]). It is sited near Geneva in a tunnel between 50 m and 175 m below ground [31] (depending on the topology of the location), crossing the Swiss-French border, as can be seen in figure 3.1. This tunnel was originally dug for the Large Electron-Positron Collider (LEP), which was also run by CERN between 1989 and 2000 [32]. The LHC was designed to accelerate both protons and lead-ions and replaced LEP. With a circumference of 26.66 km, it is the worldwide largest hadron collider still today. Run 1 with protons started in 2011 with a collision energy of 7 TeV, which was increased to 8 TeV in 2012 when the Higgs boson was found. After the first long shutdown (LS1), Run 2 with protons started in 2015 with a collision energy of 13 TeV and lasted until 2018. There are also short shutdowns during winter used for maintenance and smaller upgrades. A schematic of the LHC timeline from 2018 is shown in figure 3.2. This thesis contributes to the analysis of data taken during Run 2 with protons.

3.1.2 Functional principle

The LHC accelerates particles using the principle of a synchrotron, which is explained in various textbooks, e.g. in [35]. In contrast to a classic synchrotron, which would eject the particles onto a target, when they have reached their final collision energy, the LHC stores the particles and lets them collide with each other at one of four possible spots on the ring, called interaction points. This property makes the LHC to be a special type of synchrotron, called storage ring [36]. It is designed for colliding

[1] derived from the french name Conseil européen pour la recherche nucléaire

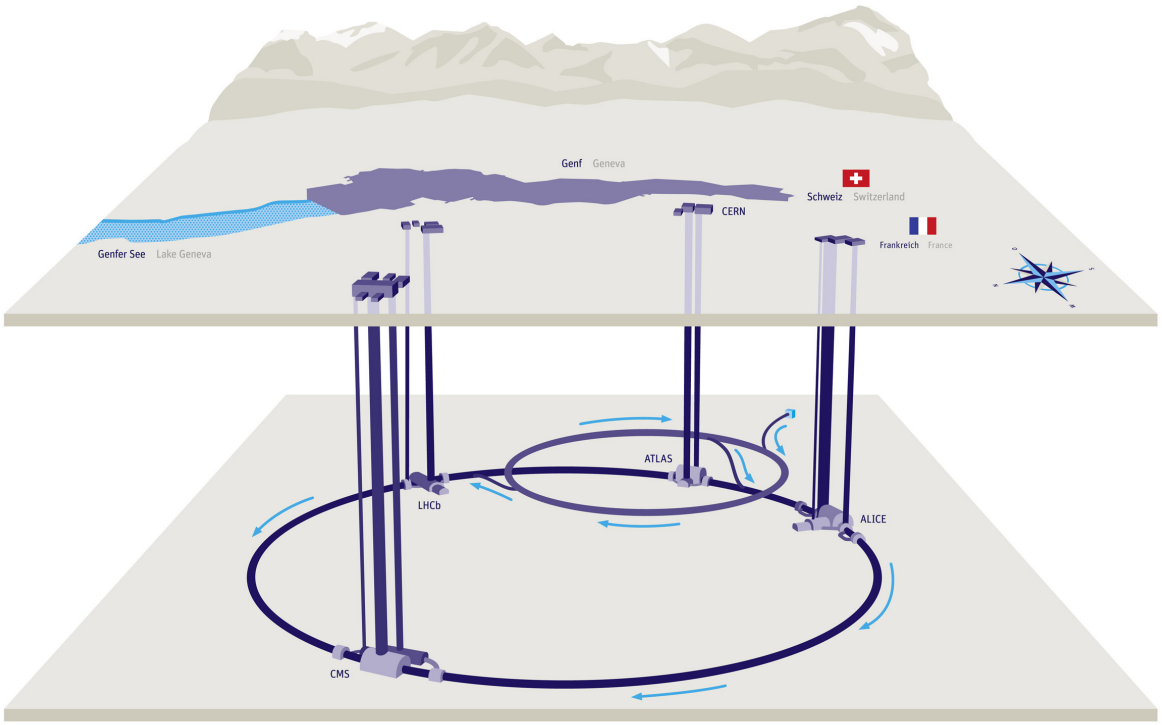


Figure 3.1: Schematic, showing the Large Hadron Collider (LHC) ≈ 100 m belowground near the lake of Geneva in the presence of the Alps. There are four collision points covered by the experiments ATLAS, ALICE, CMS and LHCb with their control centers aboveground. The light blue arrows indicate the direction of travel of the accelerated particles. They are injected into the LHC from the Super Proton Synchrotron (SPS), a circular accelerator surrounding ATLAS. The other pre-accelerators, which inject into the SPS, are indicated by a small light blue box, but not shown in detail. [30]

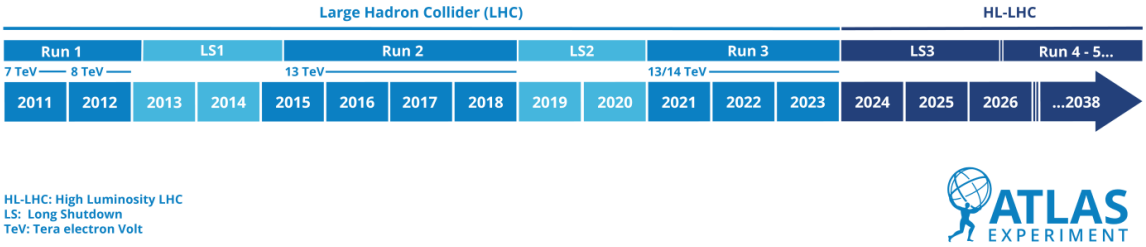


Figure 3.2: The LHC Timeline from 2018. In the meantime, the schedule had to be postponed. Run 3 started in 2022 [33]. "LS1", "LS2" etc. stand for "Long Shutdown 1" and "Long Shutdown 2" respectively. "HL-LHC" stands for "High Luminosity LHC". The values for the energy refer to the collision energy, which is twice the beam energy. Schematic taken from [34].

protons and lead-ions and has two beam pipes, in which the particles propagate in opposite directions (one for accelerating clockwise and one for accelerating counter-clockwise). The particles are injected coming from the Super Proton Synchrotron (SPS), which works as a pre-accelerator of the LHC and has pre-accelerators by itself, which are optimized for different energy ranges. To keep the particles on that trajectory, magnets are used. [37]

3.1.3 Luminosity

The LHC is filled with bunches of $n_p \approx 10^{11}$ protons. It is designed to hold up to $N_b = 2808$ bunches at a bunch spacing of 25 ns in between. The protons are injected with a speed that is already close to the speed of light in the vacuum of $c = 3 \times 10^8$ m/s. With a circumference of the LHC of 26.66 km this corresponds to a revolution frequency of $f = \frac{c}{26.66\,\mathrm{km}} = 11.25\,\mathrm{kHz}$. For collisions, the beams have to be focused and directed onto each other. After focusing the beam size in x-direction is called σ_x^* and the beam size in y-direction is called σ_y^*. With these quantities, the luminosity of the LHC can be calculated via the formula:

$$\mathcal{L} = \frac{N_b n_p^2 f}{4\pi \sigma_x^* \sigma_y^*} F \tag{3.1}$$

The factor $F \leq 1$ takes into account that it is not possible to let the beams collide head-on but with a certain crossing angle, which decreases the effective luminosity. The product $\sigma_x^* \sigma_y^*$ can also be expressed as $\sigma_x^* \sigma_y^* = \beta^* \epsilon / \gamma$ with the betatron function, β^*, evaluated at the interaction point, and the normalized (i.e. energy-independent) emittance, ϵ, divided by the relativistic factor, γ. The design luminosity of the LHC is $\mathcal{L}_{\mathrm{LHC}}^{\mathrm{design}} = 1 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. It has been exceeded by a factor of 1.4 in 2016 and by more than 2 in the years 2017 and 2018 [38]. During Run 2, the LHC deliverd a total integrated luminosity of $\int \mathcal{L} \mathrm{d}t = 156\,\mathrm{cm}^{-2} = 156\,\mathrm{fb}^{-1}$ [39].

3.1.4 Kinematics

The beam energy, E_{beam}, is the kinetic energy of each proton. It was $E_{\mathrm{beam}} = 3.5\,\mathrm{TeV}$ in the beginning of operation of Run 1 of the LHC in 2011, and $E_{\mathrm{beam}} = 4.0\,\mathrm{TeV}$ in 2012 and $E_{\mathrm{beam}} = 6.5\,\mathrm{TeV}$ for Run 2. The total energy of the (proton-proton) collision, $E_{\mathrm{collision}}$, is computed via the Mandelstam variable

$$s = \left(P_1 + P_2\right)^2 = P_1^2 + P_2^2 + 2P_1P_2 = 2m_p + 2P_1P_2 \approx 2P_1P_2 \tag{3.2}$$

The beam energy is so high that even the proton masses of $m_p \approx 1\,\mathrm{GeV}$ can be neglected. Since the three-momenta $\vec{p}_1$ and $\vec{p}_2$ of the two protons have the same magnitude but opposite direction, they add up to zero. Therefore, the collision energy can also be calculated via:

$$E_{\mathrm{collision}} = \sqrt{s} = \sqrt{\left(P_1 + P_2\right)^2} = \sqrt{\left(E_{\mathrm{beam}} + E_{\mathrm{beam}}\right)^2} = 2E_{\mathrm{beam}} \tag{3.3}$$

In general, however, only one parton of each proton participates in the interaction. The other partons are called spectators. The partons participating in the interaction do not carry the full four-momentum of each proton, P_1 and P_2, but only a fraction, x_1P_1 and x_2P_2. The energy available for the interaction is $\sqrt{\hat{s}}$, where $\hat{s}$ is the Mandelstam variable calculated via

$$\hat{s} = \left(x_1P_1 + x_2P_2\right)^2 = x_1^2P_1^2 + x_2^2P_2^2 + 2x_1x_2P_1P_2 \approx 2x_1x_2P_1P_2 \tag{3.4}$$

In the last step, the proton masses have been neglected again, which allows, in comparison with equation 3.2, to see that

$$\hat{s} = \left(x_1 P_1 + x_2 P_2\right)^2 \approx 2 x_1 x_2 P_1 P_2 = x_1 x_2 s \tag{3.5}$$

The energy available for the interaction, therefore, is:

$$E_{\text{interaction}} = \sqrt{\hat{s}} = \sqrt{x_1 x_2 s} = \sqrt{x_1 x_2}\, E_{\text{collision}} = 2\sqrt{x_1 x_2}\, E_{\text{beam}} \tag{3.6}$$

This is the energy available for new particle states which can emerge at the point of interaction (POI). The factors x_1 and x_2 differ from interaction to interaction and, in general, are not equal to each other. Thus, in contrast to the momenta of the protons, which add up to zero, the momenta of the partons do not add up to zero. As a consequence, the total momentum of the final state is not zero and the whole system is boosted along the beam axis.

3.2 The ATLAS experiment (A Toroidal LHC ApparatuS)

3.2.1 Overview

Placed at one of the four interaction points of the LHC, ATLAS (A Toroidal LHC ApparatuS) is characterized as a colliding beam experiment. It consists of various types of detector components, each specialized in a specific task. The whole set of all hardware components is also simply called the ATLAS detector. In order to provide a full event reconstruction, ATLAS covers the most possible solid angle around the point of interaction (POI) and inherits the cylinder-shaped symmetry from the beam pipe. Typically, each detector component can be subdivided into three components. The barrel part is in the central region, close to the POI. For the detection of particles propagating under an acute angle to the beam axis, the detector components have two end-caps (one for each side), which are shaped like an annulus. Shieldings are installed in order to prevent the detectors from being disturbed by interactions of particles with the beam pipe, which become a source of interference for very acute angles. The purposes of ATLAS are particle identification (PID) and the determination of their four-momenta. Since this is important for a wide variety of physics analyses, ATLAS is also characterized as a general-purpose or multi-purpose detector. A schematic of ATLAS, demonstrating its components, is shown in figure 3.5. The components are explained in the subsequent subsections. Technical details of the detector, before any upgrades, are described in [40]. A more descriptive and less technical introduction can be found in [41].

3.2.2 Coordinates

In order to describe the trajectories of the particles detected in ATLAS, a right-handed Cartesian coordinate system (see figure 3.3) is defined with its origin in the point of interaction (POI). The x-axis directs into the center of the approximate circle-shape of the LHC, and the y-axis points upwards in direction of the surface of the earth. The z-axis is parallel to the beam pipe. The projections of the three-momentum vector onto these axes give the vector components in Cartesian coordinates: p^x, p^y and p^z. The cylindrical symmetry of ATLAS allows the additional use of the following set of coordinates, p^{T}, ϕ and θ, which are related to the Cartesian coordinates as follows. The plane transverse to the z-axis is called the transverse plane and the projection of the three-momentum onto this plane is called the

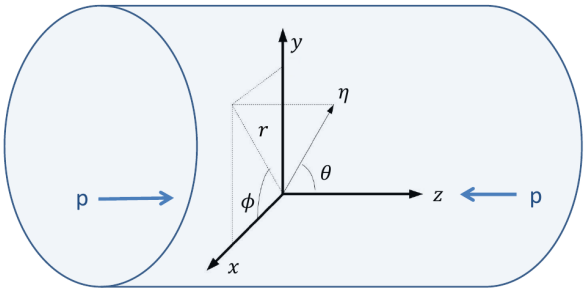


Figure 3.3: A schematic, taken from [42], showing the coordinate system used at ATLAS. In z-direction the beam pipe and the colliding proton beams (p) shown in blue. The y-axis point upwards to the surface of the earth, the x-axis points inwards into the center of the LHC.

transverse momentum, p^{T}.[2] The azimuth angle, ϕ, is defined in the range $[-\pi, \pi)$ as the angle between p^x and p^{T}. The polar angle, θ, is defined in the range $[0, \pi]$ as the angle between p^z and p^{T}. Since the whole system of the final state is in general boosted into z-direction (see section 3.1.4), θ is usually replaced by the Lorentz-invariant pseudorapidity $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$ (see section 2.2.7). Since the set of p^{T}, ϕ and η is a coordinate choice typical for collider physics, the term "collider physics coordinates" is used for that set in this thesis. As an angular distance measure between the objects i and j, the quantity

$$\Delta R_{ij} := \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2} \tag{3.7}$$

is defined.

3.2.3 Basic introduction to the detection of particles

The basic principle behind the detection of particles is the interaction of the particles to be detected with the detector medium. Examples are the ionization of gas or the creation of electron-hole pairs in a semiconductor. Detailed information about different detector media and their applications can be found in various textbooks, e.g. [36] or [44]. For the detection of the various particles, it has to be considered that different particles interact in different ways. Furthermore, only long-lived particles which do not decay before reaching the detector can interact. Those particles are electrons and positrons[3] ($e^{\pm}$), muons ($\mu^{\pm}$), photons (γ) and some of the hadrons, like protons (p), pions ($\pi^{\pm}$, π^0) and neutrons (n), for example. The inner detector (ID) only detects electrically charged particles, i.e. $e^{\pm}$, $\mu^{\pm}$, p and $\pi^{\pm}$ for example. The electromagnetic calorimeter (ECAL) additionally detects photons (γ). The hadronic calorimeter (HCAL) is specialized in the detection of neutral hadrons (e.g. π^0 and n). The muons are the only particles that are even not stopped in the HCAL. Exploiting this for a better reconstruction, muon detectors are sited at the very outside. Neutrinos only interact by the weak interaction and cannot be detected in ATLAS. The different ways how particles interact with the detector are exemplarily shown in figure 3.4.

[2] If "T" is used for the projection onto the transverse plane, it is written in superscript, in analogy to the Cartesian components, which are also projections (onto the coordinate axes). If it is written in subscript, it also refers to the transverse plane but does not mean a projection onto it.

[3] The electron is the only particle, whose antiparticle, the positron, has a different name. As usual in collider physics, the term "electron" refers to both electrons and positrons in the following.

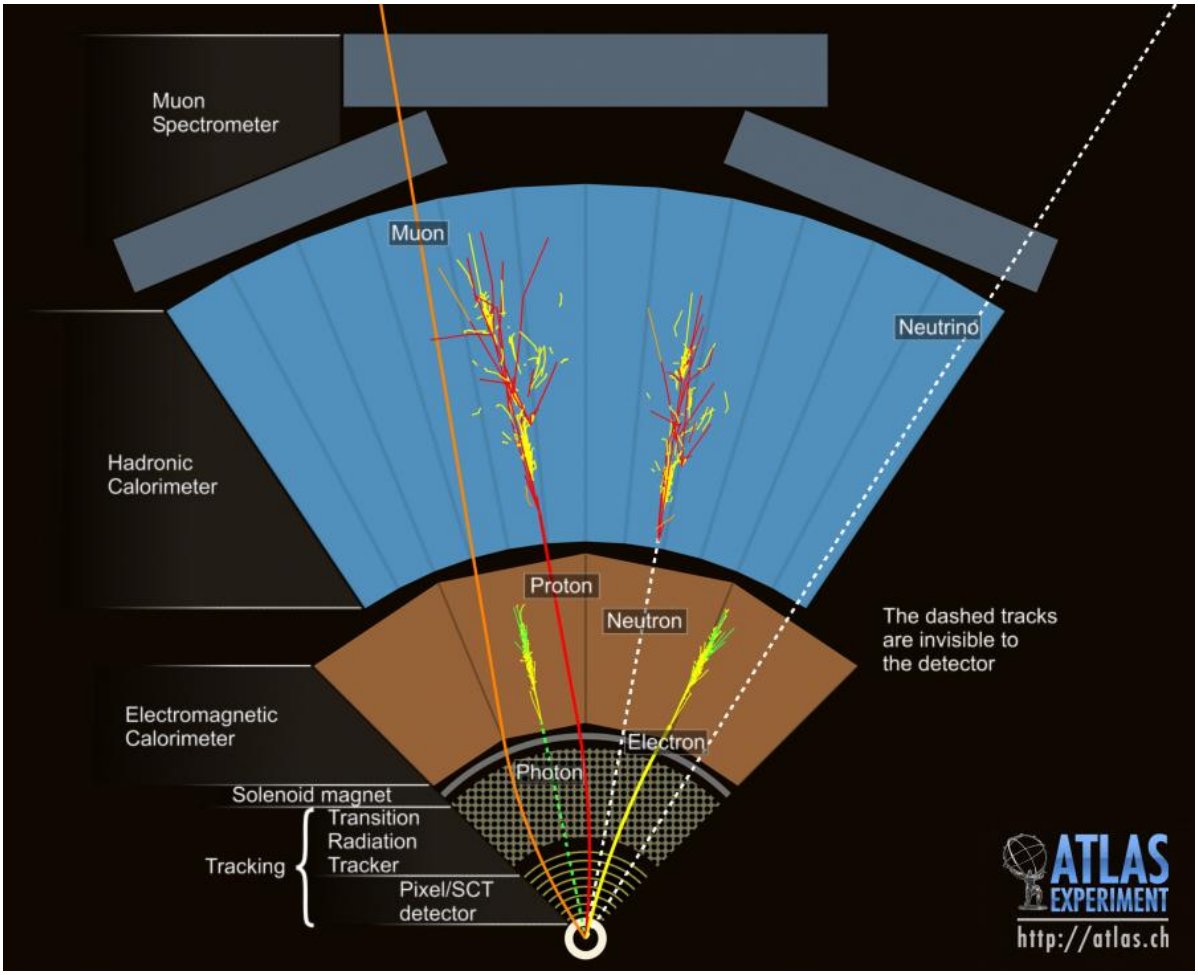


Figure 3.4: Simplified schematic cross section of the ATLAS detector in the barrel region, showing exemplary trajectories of prominent particles. Taken from [43].

3.2.4 Basic introduction to the reconstruction capabilities

The purpose of ATLAS is not only the detection of particles but also the reconstruction of the full four-momentum of each particle, i.e. the three-momentum and the energy. This task is performed differently for electrically charged particles and for those electrically neutral.

Inner detector (ID) For momentum reconstruction, ATLAS is equipped with a magnet system. The magnetic fields bend the trajectory of electrically charged particles. The direction in which the trajectory is bent corresponds to the sign of the electric charge, and the degree corresponds to the transverse momentum. Therefore, the reconstruction of the trajectory, called tracking, is a key ingredient, and the inner detector (ID) is dedicated only to this task. Since every interaction of particles with the detector causes energy loss and thus momentum loss, which would lead to a mismeasurement of momentum, the material density in the ID is as low as possible.

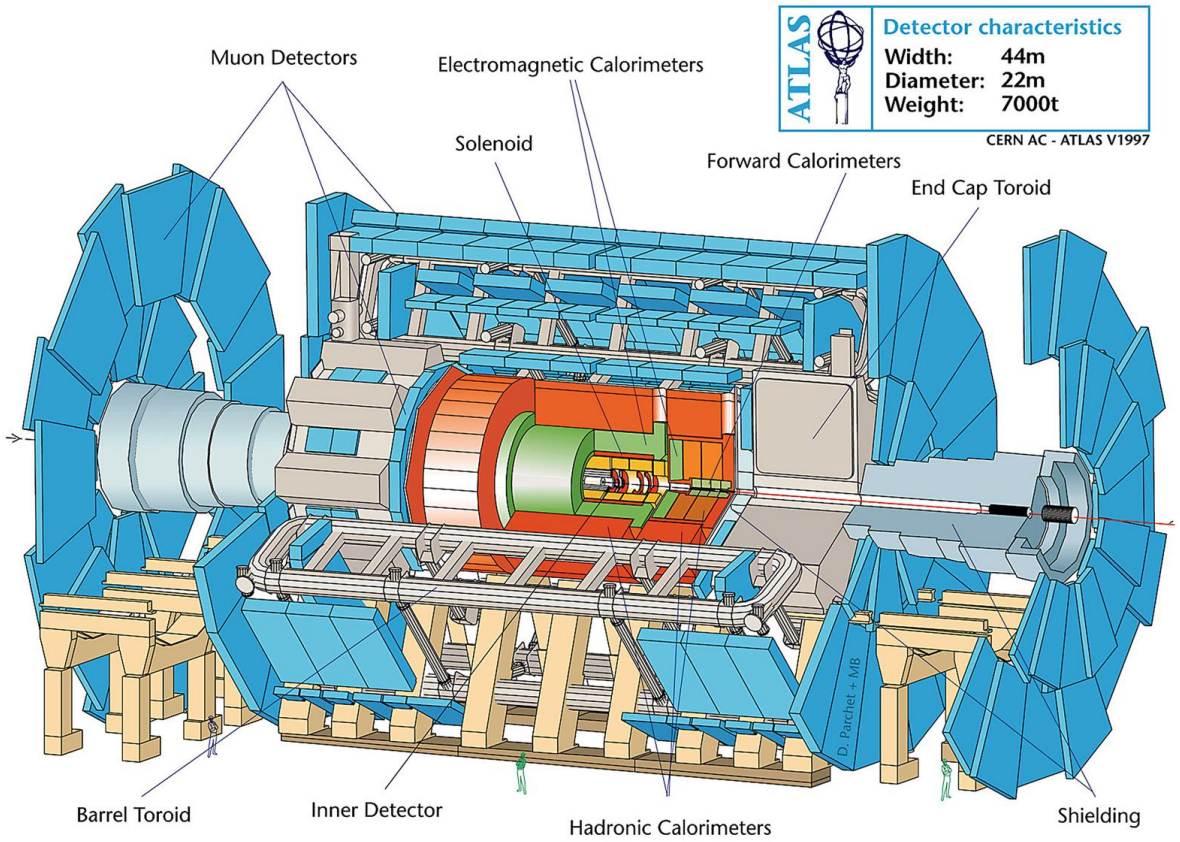


Figure 3.5: Computer-generated layout of the ATLAS detector [45], showing in the order from top to bottom: the muon detectors in blue, the toroidal magnets (both end-cap and barrel) in grey, the shielding in blue-grey, the hadronic calorimeter (HCAL) in orange, the electromagnetic calorimeter (ECAL) and forward calorimeter (FCAL) in green, the transition radiation tracker (TRT) in yellow, the solenoidal magnet (between TRT and ECAL) in red, pixel and semiconductor detector in white (except for the SCT endcaps which are shown in red), the supporting structure in ocher. The components of the inner detector can be viewed in more detail in figure 3.6. In front of the supporting structure figures of human beings are sketched for a comparison of size.

Calorimeters The calorimeters, which surround the ID, are intended to measure the energy by full absorption of the particles and the material density is therefore higher. Photons create electron-positron-pairs and those, in turn, create photons by bremsstrahlung. Thus, both electrons, positrons and photons initiate a cascade of secondary particles, called an electromagnetic shower. The shower stops when the energy of the secondary particles has fallen under a threshold that is necessary for the production of additional secondary particles. Hadrons initiate a hadronic shower through interactions with the nuclei in the calorimeter. The energy is measured by the detection of the secondary particles. In order to spatialize them and relate them to their primary particles, the calorimeters are subdivided into segments. The ECAL is closest to the ID and is dedicated to the detection of electromagnetic showers, which are typically fully absorbed in the ECAL. The HCAL surrounds the ECAL and is dedicated to the detection of hadronic showers, which may also be present in the ECAL to a lower fraction but spread to a larger volume. Electrically neutral particles do not leave a *track* in the ID. For those, only the information from the calorimeters is available and has to be used for reconstructing both energy and momentum using a

dedicated energy calibration.

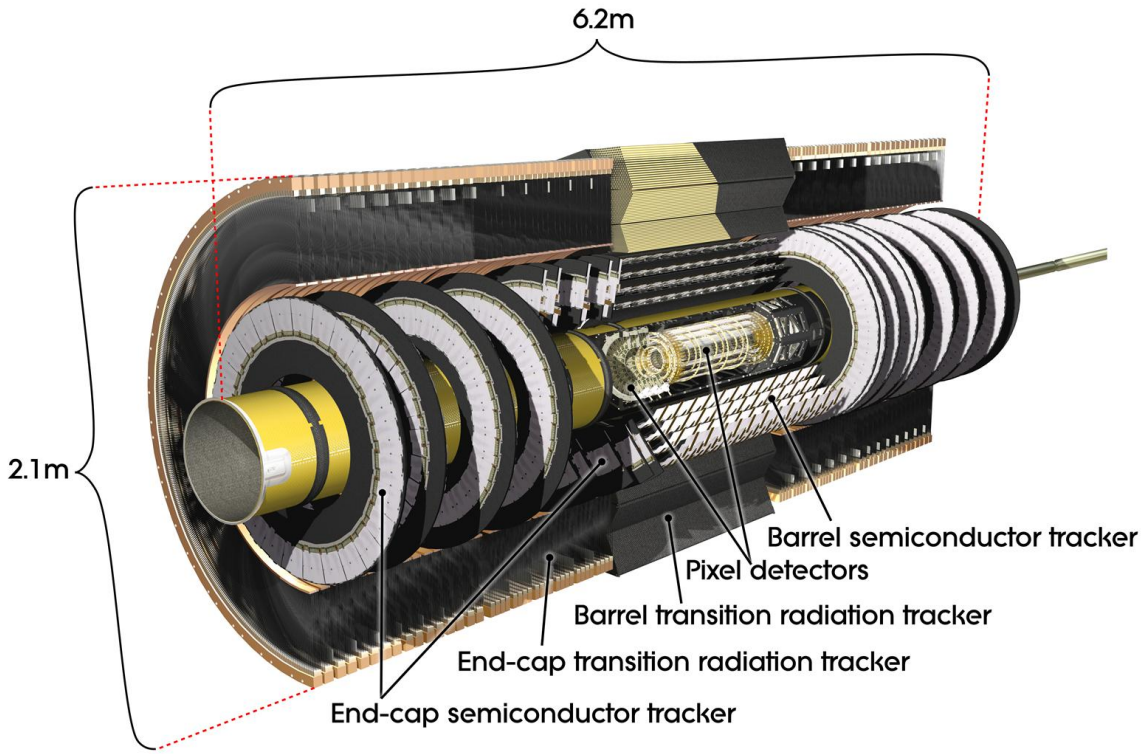


Figure 3.6: Computer-generead layout of the inner detector (ID) of ATLAS [46].

3.2.5 The triggering system

The bunch spacing of 25 ns corresponds to a bunch crossing rate of 40 MHz. ATLAS is not capable of recording every event at that high rate. On the other hand, not all bunch crossings contain interactions interesting for physics analysis. Therefore, a two-level trigger system is installed in order to find those events by triggering for energetic leptons, photons, jets or missing transverse energy [47]. The first level (L1) is purely hardware-based using custom electronics, based on field-programmable gate arrays (FPGA), housed in a separate cavern [40]. Using calorimeter signals of coarser granularity and fast signals from the muon spectrometer, it locates regions of interest (RoI), which are passed to the software-based second-level trigger, called high-level trigger (HLT). Both trigger levels combined reduce the event rate to about 1 kHz.

3.2.6 The inner detector (ID)

The inner detector, shown in figure 3.6, has a diameter of about 2.5 m and allows tracking in the range $|\eta| < 2.5$ and consists of three components.

The pixel detector The component closest to the beam pipe is the silicon pixel detector. It consists of four layers in the barrel part and three discs on each side of the two end-caps. The innermost layer of the barrel part, called the Insertible B-Layer (IBL) [48], has been inserted during LS1 at a radial distance of only 3.3 cm to the beam axis together with a smaller beam pipe replacing the former one [49]. Since it delivered additional 12 million pixels, the whole pixel detector contains over 90 million pixels now. Hits of particles in the pixels are registered with a spatial resolution of about 10 µm ($r\phi$) and of about 100 µm in z-direction.

The semicondutor tracker (SCT) The pixel detector is surrounded by the semiconductor tracker (SCT), which consists of four layers in the barrel region and of each nine disks for the end-caps. The 4088 sensor modules are two-sided, on each side consisting of 768 active silicon microstrips with a pitch of 80 µm and a relative angle of 40 mrad to each other. This geometry allows for an intrinsic resolution of 17 µm perpendicular to the strips ($r\phi$) and of 580 µm parallel to the strips (corresponding to the z-direction in the barrel part and corresponding to the radial direction for the end-caps).

The transition radiation tracker (TRT) The last component of the inner detector is the transition radiation tracker (TRT). It consists of about 300 000 straw tubes of 4 mm diameter in total. 52 544 of those are in the barrel part. They are 144 cm long and aligned parallel to the beam in 73 layers. The residual tubes are 37 cm long and arranged in 20 wheels at the end-caps in 160 layers in total. The space in between the tubes is filled with fibers (barrel) or foils (end caps) of polypropylene, $(C_3H_6)_n$, in CO_2 as environmental gas. Propagating through the multiple transitions between polypropylene and CO_2, the charged particles radiate transition radiation. The straw tubes allow both the detection of the charged particles with a resolution of 130 µm (perpendicular to the tubes) and the detection of the transition radiation, which improves the separation of $e^{\pm}$ from $\pi^{\pm}$. The TRT does not provide spatial resolution in the direction parallel to the tubes. The lower intrinsic resolution is compensated by a large number of hits (at least 36 in the barrel and 22 in the end-caps).

3.2.7 The magnet system

The magnet system consists of four components: one solenoid and three toroids, responsible for the name ATLAS (A Toroidal LHC ApparatuS). The coil of the solenoid has an inner radius of 1.23 m and an outer radius of 1.28 m. With a length of 5.8 m, it surrounds the inner detector and immerses it with a flux density of $B = 2$ T. The central, large toroid has an inner radius of 4.7 m and an outer radius of 10.0 m. Eight separate coils, 25.3 m arranged in an eightfold azimuthal symmetry around the beam axis provide a magnetic field of $B \approx 0.5$ T. Two smaller toroids with both a radius and a length of about 5 m in the end-caps provide a magnetic field of $B \approx 1$ T.

3.2.8 The calorimeters

Geometry The calorimeters are sited between the solenoid and the toroids. More precisely, the very first layer of the barrel part of the electromagnetic calorimeter (ECAL), called presampler and merely 1.1 cm thick, lies inside the solenoid. The barrel part of the hadronic calorimeter (HCAL) surrounds the barrel part of the ECAL and extends from a radial distance of $r \approx 2.25$ m to $r \approx 4.25$ m. The HCAL barrel is subdivided into three parts. One part, called the tile barrel, has approximately the same length as the ECAL barrel and the inner detector. The other two parts, called tile extended barrels, surround the

ECAL end-caps as well as the HCAL end-caps. The end-caps do not touch the beam pipe but surround another calorimeter, called the forward calorimeter (FCAL). The FCAL has both an electromagnetic and a hadronic part and extends the coverage to $|\eta| < 4.9$.

Technology Both the ECAL and the HCAL are sampling calorimeters, i.e. each part (one layer of the ECAL barrel for instance) in itself consists of many alternating layers of two different materials, called active and passive. For the passive layer, a material of high density (lead, iron, copper or tungsten) is chosen in order to absorb the energy of the incident particle and transfer it to secondary particles (showers). The active layer is to detect the secondary particles and its material is either a scintillator (optimized for the detection of secondary photons) or a material, like liquid argon, optimized for the detection of secondary particles electrically charged.

Segmentation In order to allow for a space-resolved readout, each part of the calorimeter is segmented into cells, whose sizes differ for different parts and also depend on η. Most cells of layer 1 and of layer 2 of the ECAL barrel have a size of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The granularity of the HCAL is generally coarser with a typical size of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. All in all, the ECAL provides about 190 000 readout channels and the HCAL about 10 000.

3.2.9 The muon spectrometer (MS)

The muon spectrometer consists of various components. In the barrel part, three layers of monitored drift tubes (MDTs) are either installed on the eight coils of the central toroid or in the space between them. They are positioned between resistive plate chambers (RPCs), of which some are also installed at the very outside of the toroid. For the end-caps, cathode strip chambers (CSCs) are installed between the calorimeters and the small toroids, and MDTs are installed as large wheels at the very outside. Here, the MDTs are positioned between thin gap chambers (TGCs). Among the four different technologies, MDTs deliver the best spatial resolution of 35 µm per chamber. The CSCs deliver additional spatial information, while the RPCs and TPCs are optimized to respond very fast and are used for triggering.

3.3 Low-level object reconstruction

Many different particles propagate through ATALS at the same time. Each detector component consists of multiple layers or segments. The first step in handling this huge amount of data is to reconstruct low-level objects. Since the low-level objects are the basis for the reconstruction of physical objects, described in section 3.4, some of those, namely *tracks* (from the hits in the ID or in the MS) and *topo-clusters* (from the hits in the calorimeters), are introduced in this section.

3.3.1 Reconstruction of *tracks* from hits in the inner detector (ID)

Due to the high luminosity of the LHC, there are lots of trajectories in one bunch-crossing. Therefore, the *track* reconstruction is a complex, multiple-stage procedure. The following description is a rough summary of what is presented in more detail in [50]. At first, hits in the pixel detector and in the SCT are grouped into clusters. For the pixel detector, each cluster is equivalent to one three-dimensional *space point* while for the SCT clusters, clusters from both sides of a strip layer have to be combined for a

three-dimensional measurement. Due to the high event density, clusters may be caused by more than one particle interaction, i.e. have a higher cluster multiplicity. Those clusters are identified based on information gained from simulation. Sets of *space points*, called seeds, are passed to a combinatorial Kalman filter [51] which creates multiple possible *track* candidates per seed. The ambiguity is resolved using track scores which reflect the *track* quality. Clusters assigned to a *track* increase the *track* score weighted by their intrinsic resolution and cluster multiplicity. Holes, i.e. sensitive detector elements without a cluster, decrease the score. In the last steps, those *tracks* are matched to the measurements from the TRT.

3.3.2 Reconstruction of *topo-clusters* from the calorimeter signals

At first, calorimeter cells are merged into prototypes of *topo-clusters*, called *proto-clusters*, using the energy deposition in the cell, E_{cell}, the estimated electronic noise in the cell, σ_{cell}, and a tuple of three parameters, S, P and N with a default setting of $(S, N, P) = (4, 2, 0)$. If one cell fulfills the condition $E_{\text{cell}} > S\sigma_{\text{cell}}$, this cell is taken as a seed for a *proto-cluster*. Neighboured cells are added to the cluster if they fulfill $E_{\text{cell}} > N\sigma_{\text{cell}}$, and the last neighboured cells to be added, are just allowed to fulfill $E_{\text{cell}} > P\sigma_{\text{cell}}$. This procedure allows the energy depositions of different particles to be clustered in the same cluster. Only after the *proto-cluster* is formed, local signal maxima of deposited energy can be identified and the *proto-cluster* is split accordingly into clusters, which only contain one maximum of deposited energy and are called *topo-clusters*. With calorimeter signals alone, it is not possible, in general, to assign a physical meaningful mass to the *topo-clusters*. Therefore, each *topo-cluster* is interpreted as a massless pseudo-particle. The cluster coordinates η_{cluster} and ϕ_{cluster} are calculated as signal-weighted barycenters and correspond to the direction of this pseudo-particle. Incorporating the total energy assigned to the *topo-cluster* allows for the reconstruction of the full four-momentum. [52]

3.3.3 Reconstruction of *tracks* in the muon spectormeter (MS)

In contrast to the inner detector, the components of the MS do not adjoin each other. They have parts of other detector components (e.g. the magnet system) in between (see section 3.2.9) and are therefore called stations. Therefore, the *track* reconstruction starts from reconstructing track segments in each station using a Hough transform (cite). In the second step, those segments are combined and additional information from the muon triggers is included. A first global χ^2 fit of the muon trajectory is used to identify outlier hits, remove them from the track and include hits along the trajectory that have not been used before. After a second fit, ambiguities between tracks are resolved and the final set of tracks is refitted again with a loose constraint on the interaction point and taking also the energy loss in the calorimeters into account. In the end, the track is extrapolated to the beam line [53].

3.4 Physical object reconstruction

The reconstruction of physical objects is the basis for any event reconstruction and thus for many physics analyses. This section focuses on those physical objects which appear in the analysis of tHq in the $2\,e/\mu + 1\tau_{\text{had}}$ decay channel. Since the phenomenon called pile-up affects the object reconstruction, it is introduced shortly in the beginning.

3.4.1 Pile-up

Only a fraction of the events happening in the proton-proton-collisions at the LHC is interesting for physics analysis. The purpose of the high luminosity of the LHC is that also rare processes can be produced with sufficient statistics, but it also gives rise to the problem of pile-up (PU). In-time PU describes the fact of multiple interactions per beam-crossing. Out-of-time PU describes electronic signals in the detector which are caused by interactions of earlier or later bunch crossings. Pile-up makes event reconstruction more complicated but has to be accepted in order to increase the statistics for rare events.

3.4.2 Hadrons and Jets

As explained in section 2.3.4, partons hadronize, i.e if a quark or a gluon is dissociated from a hadron, a new quark-antiquark-pair emerges and new hadrons are formed. Usually, the energy of the dissociated parton is that high that the process of hadronization happens multiple times, i.e. for each dissociated parton multiple hadrons are produced. The kinetic energy of each hadron in the rest frame of the dissociated parton is much lower than the kinetic energy of the dissociated parton in the lab frame. That means that in the lab frame, the trajectories of the hadrons are boosted in direction of the dissociated parton and appear collimated as a bundle. This bundle is called a particle jet (short: jet). Starting point for the reconstruction of jets are the calorimeter cells, including both the electromagnetic and the hadronic calorimeter [52].

Building jets using either *particle flow objects* or *topo-clusters* Charged hadrons create a signal both in the calorimeter and in the tracker. Using the Particle Flow algorithm [54], those signals can be combined creating *particle flow objects* in order to feed them into the jet-finding-algorithm. For neutral hadrons, however, pure *topo-clusters* are taken as input. In general, there are various jet-finding algorithms. In ATLAS, the anti-k_t algorithm is used. Since jets are bundles of particles, particles belonging to one jet have to be close to each other in some sense, i.e. *topo-clusters* that are close to each other are bundled into one jet by using a jet-finding algorithm. Therefore, a distance measure has to be defined for this algorithm. For anti-k_t a distance measure d_{ij} between two *topo-clusters*, labeled i and j, and a distance measure d_{iB} between *topo-cluster* i and the beam B are defined:

$$d_{ij} = \min\left(\left(p_i^T\right)^{-2}, \left(p_j^T\right)^{-2}\right) \cdot \frac{\Delta R_{ij}^2}{R^2} \qquad d_{iB} = \left(p_i^T\right)^{-2} \tag{3.8}$$

ΔR_{ij} is the angular distance measure defined in equation 3.7. R is a radius parameter, which is sometimes also referred to as jet radius, and has to be chosen before the algorithm starts. For most analyses and also for the analysis presented in this thesis, the choice is $R = 0.4$. For all *topo-clusters* of one event d_{ij} and d_{iB} are calculated and stored in a database. Then, the algorithm consists of only one step, which is executed until there is no element left in the database:

- Find the smallest d_{ij} and d_{iB} in the database and
  - if $d_{ij} < d_{iB}$: The *topo-clusters* i and j are considered to belong to the same jet, called ij, which replaces the *topo-clusters* i and j in the database.
  - else: The *topo-cluster* i is considered to be a completed jet and is removed from the database.

In the end, the jets have to be corrected for contributions from pile-up and the energy can be calibrated. The jet energy scale (JES) calibration is obtained by using simulation-based corrections first and applying in-situ corrections thereafter [55]. New techniques on this are reported in [56].

3.4.3 *B*-jets

b-jets are jets that emerge from the hadronization of a *b*-quark. They are distinguished from jets containing a *c*-quark, called *c*-jets, and from jets containing only *u*-, *d*- or *s*-quarks, which are called light-flavor jets and are by far the most frequent type of jets. The discrimination of *b*-jets against these two other types of jets is called *b*-tagging and is a way to identify *b*-quarks which are either directly produced in the proton-proton-collision or are a direct decay product of top quarks produced in the collision. Since a top quark almost always decays into a *b*-quark (see table 2.5), the identification of *b*-jets is also very important for the identification of top quarks.

Properties of *b*-jets Because *b*-quarks have significantly higher mass than the other hadronizing quarks, *b*-jets have typically higher p^{T} [57]. Apart from that, *b*-jets include hadrons containing a *b*-quark (most often B-mesons). That they decay into light leptons sometimes, can also help to distinguish them from light-flavor jets. However, the most important property is the decay length of the B-mesons. It depends on their boost, but is at least $\approx 450\,\mu$m (in the limit of no boost) and ≈ 4.25 mm for a boost due to an energy of 50 GeV. This means that they decay too fast to see any detector, but travel long enough that the decay vertex can be reconstructed and separated from the interaction point (performance could be improved due to additional b-layer). Additionally, the decay of *b*-hadrons leads to large *track* multiplicity.

Tagging of *b*-jets In ATLAS, *b*-tagging works in two stages. In the first stage, various algorithms, called low-level taggers, exploit the aforementioned characteristic properties of *b*-jets from *tracks*, primary vertices and jets in order to provide the high-level tagger with discriminant variables. The high-level tagger is a neural network, called `DL1r`, which gives probabilities for a jet to be a *b*-jet (p_b), a *c*-jet (p_c) or a light-flavor jet (p_{light}) as an output. For practical use, they are combined into one *b*-tagging discriminant defined as [58]

$$D_{\mathrm{DL1r}} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\mathrm{light}}}\right),$$

where f_c is the effective *c*-jet fraction in the background hypothesis, which can be adapted to the analysis in order to optimize the *b*-tagging performance. However, the discrimination achieved is not perfect. Higher rejection rate costs lower reconstruction rate. Therefore, four working points are defined, corresponding to *b*-jet selection effencies of 85 %, 77 %, 70 % and 60 %. These working points correspond to a threshold for the `DL1r` discriminant [59]. For the analysis presented in this thesis, the 70 % working point is used.

3.4.4 Taus

The short decay time of the tau, $\tau = 2.9 \times 10^{-13}$ s (see table 2.3), leads to a decay length of about 87 nm in the limit of no boost and of about 2.4 mm due to a boost of about 50 GeV. Therefore, the tau decays

before interacting with any detector component and hence can only be detected by its decay products. As shown in table 2.4, the tau decays either leptonically, i.e. with an electron or a muon in the final state, or hadronically, i.e. with one or three electrically charged pions or kaons and possibly additional pions ar photons in the final state. Since leptonically decaying taus can not be distinguished from prompt electrons or muons [60], the reconstruction and identification of taus in ATLAS consider only hadronically decaying taus. Therefore, potential tau candidates are jets, reconstructed from *topo-clusters* with the anti-k_t algorithm with radius parameter $R = 0.4$ [61]. One or three *tracks* have to be matched to the *topo-clusters*. Since tracking information are only available for $|\eta| < 2.5$ and the rate of fake-*tracks* is negligible for $p^{\mathrm{T}} > 10\,\mathrm{GeV}$ [50], these two conditions are applied as additional kinematic selection criteria for tau candidates.

Reconstruction and identification Since the jet which is taken as a tau candidate is a decay product of the tau, the vertex of the tau may differ from that originally assigned to the event as a primary vertex. On the other hand, there are many vertices due to (in-time) pile-up, which must not be mismatched to the tau. Therefore, the first step of the tau reconstruction is to find its primary vertex. A dedicated algorithm, called the Tau Vertex Association algorithm, starts with matching *tracks* within $\Delta R = 0.2$ around the jet direction to the *topo-clusters* in order to use them for finding the primary vertex. The jet direction is given by the barycenter of the four-momenta of all *topo-clusters* associated with the jet. After the primary vertex has been found, the four-momenta of all *topo-clusters* are recalculated in the coordinate system of the associated tau vertex. Using this new coordinate system, the region $\Delta R < 0.2$ is called the core region and the region $0.2 < \Delta R < 0.4$ is called the isolation region. *Tracks* within the core region are used for calculating the momentum of the tau and are counted in order to classify the tau decay into 1-prong and 3-prong decays. *Tracks* within the isolation region are only used for the calculation of discriminating variables used for identification [62]. They are passed to a recurrent neural network (RNN) [63], delivering one algorithm for 1-prong and another for 3-prong tau decays. The four working points *tight*, *medium*, *loose* and *very loose* are provided, of which the *loose* working point is used for the analysis in this thesis. The energy calibration of the taus, which is not the same as for jets, is described in [61].

3.4.5 Electrons and positrons

Both electrons and photons produced at the POI are likely to interact with the material in the ATLAS detector they transverse. Electrons are likely to radiate off photons (bremsstrahlung) whereas photons are likely to convert into an electron-positron-pair. While these interactions are exploited for measuring their energy in the ECAL, they complicate the reconstruction of those objects because they may also happen before entering the ECAL, e.g. in the wall of the beam pipe or in the tracker. Therefore, the reconstruction of electrons is highly connected to the reconstruction of photons. In principle, electrons can be distinguished from electrons, since they leave a *track* in the ID and the photons do not. However, photons which converted into an electron-positron-pair before entering the ECAL (called converted photons) leave a *track* as well. By reconstructing the conversion vertex, they can be identified. However, this section focuses on the reconstruction of electrons. Like the photon, the electron is expected to lose its full energy in the calorimeters (mainly in the ECAL) through an electromagnetic shower. Because of this shower, the energy of the primary electron or photon is spread into many *topo-clusters*. For the full reconstruction of the electron or photon, those *topo-clusters* have to be identified and merged into *superclusters*.

Reconstruction The reconstruction of electrons in the central region ($|\eta| < 2.5$) is described in detail in [64]. The following summary does not claim to consider all details to full accuracy:

- *topo-clusters* are matched to *tracks* and the *tracks* are re-fitted accordingly. In the case that an electron radiates off a photon and that creates an electron-positron-pair, one electron can create three *tracks*, or even more if those two processes happen consecutively multiple times.
- The re-fitted *tracks* are matched to *topo-clusters*, which are sorted in descending order of E_{T}. If multiple *tracks* are matched to one *topo-cluster*, the *tracks* are ranked according to quality requirements.
- In descending order of E_{T}, *topo-clusters* fulfilling minimum requirements ($E_{\mathrm{T}} \geq 1\,\mathrm{GeV}$, at least one matched *track* with at least four hits the pixel detector or the SCT) are selected as seeds for the *superclusters*.
- Clusters which are close to the barycenter of the seed cluster (i.e. within the window $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ or within the window $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ if the highest-ranked *track* of the satellite cluster is also the highest-ranked *track* of the seed cluster) are called satellite clusters and associated to the seed cluster if they have not been assigned to a *supercluster* already.
- After the *supercluster* has been built, the *tracks* are matched to *superclusters* in the same way as for the *topo-clusters* and the electron energy is recalibrated because it depends on the matched tracks.

Identification After reconstruction, electrons can still be mimicked by converted photons and non-prompt electrons originating from decays of both light-flavor hadrons and heavy-flavor hadrons. For the identification, the working points *very loose*, *loose*, *medium* and *tight* are defined, where *very loose* is not intended for the analysis in signal regions but only for background studies [65]. The identification uses both cut-based selection criteria, such as the number of hits in the tracker, and likelihood-based discrimination. The likelihood functions L_S and L_B are defined both for signal and background respectively and combine the probabilitiy density functions (PDF) for various discriminative variables, like the ratio of E_{T} in the HCAL to E_{T} in the ECAL, the shower width in the ECAL, the transverse impact parameter relative to the beam axis, information from the TRT and many more. The likelihood functions are combined to the likelihood discriminant d_L, which is defined as:

$$d_L = -\tau^{-1} \ln\left(\frac{L_S + L_B}{L_S} - 1\right)$$

The parameter τ is fixed to 15. The correspondence between the thresholds for d_L, the reconstruction efficiencies, and the background rejection are determined by simulation. The reconstruction efficiency depends on E_{T} and η. For $E_{\mathrm{T}} = 40\,\mathrm{GeV}$, the efficiencies are 93 % (*loose*), 88 % (*medium*) and 80 % for the *tight* working point. For the analysis in this thesis *tight* electrons are used.

Electric charge The sign of the electric charge of a particle can be determined from the direction in which its trajectory is bent by the magnetic field (see section 3.2.4). Although this is principally easy, there are some difficulties. At very high energies or large pseudorapidities, the *tracks* appear almost straight and determining the curvature can fail. Besides an incorrect reconstruction of the curvature, it is not unlikely that the matching of the cluster to the *track* is not correct. As discussed above, this can be

the case if an electron radiates off a photon and that creates an electron-positron pair, which results in three *tracks*. Electrons with incorrectly determined charge can be suppressed using a BDT, the Electron Charge ID Selector Tool (ECIDS) [66].

3.4.6 Muons

For the reconstruction of muons, information from the muon spectrometer (MS), the inner detector (ID) and the calorimeter can be used. Muons are not stopped in the calorimeters but leave a signal of a minimum ionizing particle (MIP). Due to this multiplicity of possible information, there are multiple reconstruction strategies. They are expressed in five different types of muons which correspond to different qualities of reconstruction.

Reconstruction Combined (CB) muons are reconstructed by matching *tracks* in the ID (see section 3.3.1) to *tracks* in the MS (see section 3.3.1) and performing a global track fit. For the reconstruction of inside-out (IO) muons, *tracks* in the ID are extrapolated to the MS and refitted using at least three loosely-aligned MS hits and energy loss in the calorimeters. A *track* in the MS which cannot be matched to a *track* in the ID but extrapolated to the beam axis defines an extrapolated (ME) muon. If there is at least one segment in the MS, but this can be matched to a *track* in the ID under tight angular requirements, the reconstructed muon is called segmented-tagged (ST). Finally, if there is no signal in the MS but a calorimeter signal compatible with a MIP that can be matched to a *track* in the ID, the reconstructed muon is called calorimeter-tagged (CT).

Identification In general, muons can be reconstructed very well and are unlikely to be mimicked by other particles. Therefore, muon identification concentrates on the distinction between prompt muons and non-prompt muons which originate from the decay of either a light-flavored or a heavy-flavored hadron. Since heavy-flavored hadrons decay before interacting with the detector, non-prompt muons from a heavy-flavored hadron decay can be identified by a displaced primary vertex and the profile of the ID *track*, which is less isolated than the ID *track* of a prompt muon. Light-flavored hadrons decay during the flight in the detector and leave a signal before and after their decay. Due to the change of trajectory, those muons are reconstructed with a lower quality. For the distinction between those non-prompt muons and prompt muons, the selection working points *loose*, *medium* and *tight* are introduced, where *tight* muons are a subset of *medium* muons and *medium* muons are a subset of *loose* muons. The definitions of those working points consider the muon type (explained in the paragraph above) in combination with cuts on η and other variables. One of those additional variables, the q/p compability, is defined as

$$q/p \text{ compability} = \frac{|q/p_{\mathrm{ID}} - q/p_{\mathrm{MS}}|}{\sqrt{\sigma^2\left(q/p_{\mathrm{ID}}\right) + \sigma^2\left(q/p_{\mathrm{MS}}\right)}}.$$

It compares the ratio of the electric charge, q, to the momentum, p, either determined by the inner detector (p_{ID}) or determined by the muon spectrometer (p_{MS}), normalized according to their uncertainties $\sigma^2\left(q/p_{\mathrm{ID}}\right)$ and $\sigma^2\left(q/p_{\mathrm{MS}}\right)$. More detailed information can be found in [53]. For the analysis presented in this thesis, the *medium* working point is chosen.

3.4.7 Neutrinos and missing transverse energy

Not all particles can be detected in ATLAS. This does not only concern particles that have only been predicted by theory but have not been detected yet. Neutrinos, which are an integral part of the SM, can not be detected in ATLAS, too, since they interact only by the weak interaction.

However, at least some information about them can be gained by exploiting four-momentum conservation, which states that the sum of the four-momenta of all initial state particles is equal to the sum of the four-momenta of all final state particles. Since the p^z-component cannot be controlled (see section 3.1.4) and the E-component depends on p^z, the conservation of four-momentum can only be applied to p^x and p^y, the components of the transverse plane. Thus, the missing transverse momentum, synonymously also called missing transverse energy[4], $\overrightarrow{\not{E}}_{\mathrm{T}}$, can be reconstructed.

The contribution to $\overrightarrow{\not{E}}_{\mathrm{T}}$ caused by non-detectable particles is also called *real* $\overrightarrow{\not{E}}_{\mathrm{T}}$, while contributions caused by mismeasurements (e.g. limited detector acceptance or pile-up effects) are called *fake* $\overrightarrow{\not{E}}_{\mathrm{T}}$. Only *real* $\overrightarrow{\not{E}}_{\mathrm{T}}$ is interesting for the analysis. Therefore, the reconstruction of $\overrightarrow{\not{E}}_{\mathrm{T}}$ aims to minimize the fraction of *fake* $\overrightarrow{\not{E}}_{\mathrm{T}}$.

Reconstruction of missing transverse energy

Since the protons get collided along the z-axis, both transverse momentum components vanish for both quarks of the initial state. By four-momentum conservation, the transverse momenta of all final state particles add up to zero as well. Therefore, $\overrightarrow{\not{E}}_{\mathrm{T}}$ is calculated as the negative sum of the transverse momenta of all detected particles that have already been reconstructed. Since different choices for working points in the reconstruction of visible objects affect the object definition, it may happen that some *topo-clusters* or *tracks* may be assigned to more than one physical object [67]. To avoid double-counting in the reconstruction of $\overrightarrow{\not{E}}_{\mathrm{T}}$, a procedure called overlapping removal is applied. In this procedure, physical objects enter the calculation in a fixed order. Tracks or clusters that have already been used, are removed from physical objects that have not entered the calculation yet. The physical objects are ordered according to the quality to which they can be reconstructed, i.e.: muons, electrons, photons, hadronic taus and jets. Tracks or clusters that have not been used to that point are called soft terms and can also be added in order to improve the reconstruction efficiency of $\overrightarrow{\not{E}}_{\mathrm{T}}$. For Run 2, unused clusters are excluded, since they are impurified by pile-up to a too high extent, and only unused *tracks* associated with the primary vertex are included. The calculation of $\overrightarrow{\not{E}}_{\mathrm{T}}$ is expressed in equation 3.9.

$$\overrightarrow{\not{E}}_{\mathrm{T}} = \begin{pmatrix} \not{E}^x \\ \not{E}^y \end{pmatrix} = -\left(\sum_{\mu} \begin{pmatrix} p^{\mathrm{x}} \\ p^{\mathrm{y}} \end{pmatrix}_{\mu} + \sum_{e} \begin{pmatrix} p^{\mathrm{x}} \\ p^{\mathrm{y}} \end{pmatrix}_{e} + \sum_{\gamma} \begin{pmatrix} p^{\mathrm{x}} \\ p^{\mathrm{y}} \end{pmatrix}_{\gamma} + \sum_{\tau_{\mathrm{had}}} \begin{pmatrix} p^{\mathrm{x}} \\ p^{\mathrm{y}} \end{pmatrix}_{\tau_{\mathrm{had}}} + \sum_{\mathrm{jets}} \begin{pmatrix} p^{\mathrm{x}} \\ p^{\mathrm{y}} \end{pmatrix}_{\mathrm{jet}} + \sum_{\mathrm{soft}} \begin{pmatrix} p^{\mathrm{x}} \\ p^{\mathrm{y}} \end{pmatrix}_{\mathrm{soft}} \right) \tag{3.9}$$

More information can be found in [68].

[4] If the mass of a particle can be neglected or is set to zero in an intermediate reconstruction step, as for the *topo-clusters* described in section 3.3.2, the values for energy and momentum become the same. This may have provoked an interchangeable use of those two terms.

CHAPTER 4

Analysis strategy and setup

As stated in section 2.5.3, the analysis of single top quark associated Higgs boson production requires the determination of the number of tHq events. The first step for this is the object reconstruction, which is described in section 3.4. The second step is the event reconstruction based on the reconstructed objects. This means reconstructing the intermediate state of the event, based on the reconstructed objects of the final state. This chapter shall explain the basic strategy for event reconstruction.

4.1 Definition of the decay channel

An event of single top quark associated Higgs boson production must contain both a top quark and a Higgs boson. The respective lifetime of both of these particles is so short that they can not be detected in ATLAS and tHq events have to be reconstructed by detection of their decay products. Since both the Higgs boson and the top quark have various decay channels, there are even more possible decay channels for the whole tHq event. Because each decay channel brings its own challenges for reconstruction, it is not possible to address them all at once but to concentrate on one specific decay channel. The work presented in this thesis concentrates on the decay where $H \rightarrow \tau\tau$ and $t \rightarrow l\nu_l b$. Here, l stands for an electron or a muon, which are called light leptons throughout this thesis. Since leptonic tau decays cannot be distinguished from prompt light leptons in the object reconstruction of ATLAS (see section 3.4.4), the decay $t \rightarrow l\nu_l b$ includes also the two-step decay $t \rightarrow b\tau\nu_\tau \rightarrow bl\nu_l\nu_\tau$. Because the tau itself is too short-lived to be detected in ATLAS and has various decay modes, as can be seen in table 2.4, the decay channel has also be specified according to the tau decays. This work specializes in the decay where one tau decays leptonically and the other one hadronically. Also considering the light lepton from the top-quark decay, the channel is called $2\,e/\mu + 1\tau_{\text{had}}$. A Feynman diagram of tHq in the $2\,e/\mu + 1\tau_{\text{had}}$ channel is shown in figure 4.1. The additional quark, q, in the final state is of irrelevant flavor and therefore assumed to be a u-, a d- or an s-quark, since those are the most frequent quark flavors in the given experimental setup of proton-proton-collisions.

4.2 Monte Carlo simulation

Since the event reconstruction is challenging in general, strategies for this are developed and tested on simulated events before applying them to data. Once the strategies are finished, this approach has also

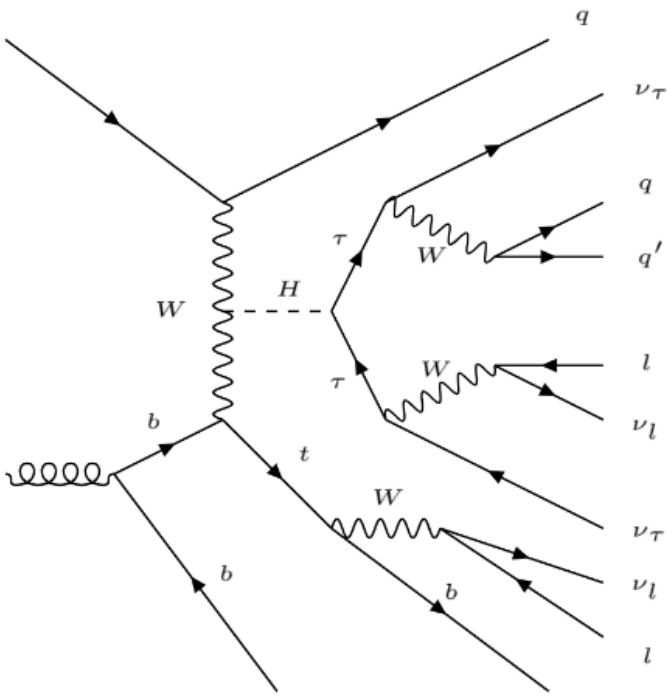


Figure 4.1: Feynman diagram of tHq showing not only its production but also the decay of the Higgs boson and the top quark in the $2\,e/\mu + 1\tau_{\text{had}}$ decay channel. q and q' stand for light quarks of irrelevant flavor. l stands for a light lepton, i.e. an electron or a muon.

the advantage that tests of the SM can simply be conducted by comparing the number of simulated events with the number of data events for the process of interest. Simulated events are generated by Monte Carlo methods [69]. Inputs to those simulations are the interactions of the current status of the Standard Model. An introduction to Monte Carlo methods in particle physics can be found in [70].

4.2.1 Overview

Simulation contains three major steps [71]: generation of the event and immediate decays (see section 4.2.2), simulation of the interactions in the detector (using `Geant4` [72]) and lastly simulation of the digitization, i.e. the transformation of the deposited energy caused by the particle interaction into the readout signal. This can then be compared to the actual readout data of the ATLAS detector, called raw data. Physical observables of simulated events which passed all these steps are referred to as on *reco-level*. Observables calculated before the detector simulation are referred to as on *truth-level*. The experimental conditions in ATLAS, e.g. the amount of pile-up, which depends on the luminosity of the LHC, are not constant for the whole Run 2 of the LHC. Therefore, the simulated samples are generated in different campaigns. Together with their corresponding period of data taking at the LHC and the integrated luminosity, $\int \mathcal{L} \mathrm{d}t$, they are shown in table 4.1. For the whole Run 2, the integrated luminosity of data which can be used for physics analysis amounts to $\int \mathcal{L} \mathrm{d}t = 139\,\mathrm{fb}^{-1}$ [39].

campaign	period of data taking	$\int \mathcal{L} dt$
mc16a	2015 - 2016	$36.21\,\text{fb}^{-1}$
mc16d	2017	$44.31\,\text{fb}^{-1}$
mc16e	2018	$58.45\,\text{fb}^{-1}$
all	2015 - 2018	$138.97\,\text{fb}^{-1}$

Table 4.1: Campaigns of MC simulated samples with the corresponding period of data taking and integrated luminosity, $\int \mathcal{L} dt$, taken from [73].

4.2.2 Event generation

The analysis presented in this thesis is about events, which do not emerge from interactions between protons as a whole but from interactions between partons, i.e. between quarks and gluons. Therefore, parton distribution functions, taken from global fits on measurements, are implemented in the simulation at first. On this level, called parton-level, the following three steps are executed.

- Calculation of the matrix element of the process of interest
- Emergence of the underlying event, i.e. the event which is formed by partons not involved in forming the process of interest
- Partons may radiate off gluons and these gluons may split into new partons which may radiate again and so on. These cascades are called parton showers and have to be simulated for partons of the process of interest, of the underlying events and for partons which are just spectactors.

In the end, hadronization, i.e. the formation of hadrons out of partons, has to be simulated. In case of tHq events, the matrix element is calculated at NLO using `MadGraph5_aMC@NLO` [74], while `PYTHIA 8` [75] is used for the underlying event as well as for parton showering and hadronization.

4.2.3 Processing the samples

After all simulation steps, the simulations are comparable with raw data. The object reconstruction described in section 3.4 is performed depending on the analysis goals. Samples after basic object reconstruction adjusted to single top quark analysis are called `SingleTop Ntuples`. They include fully reconstructed objects. Only b-tagging has to be applied afterwards. Since there are progressive improvements in how well the simulated samples image real data, new versions of `SingleTop Ntuples` are released regularly. For the analysis presented in this thesis, `SingleTop Ntuples` of version 32 have been used. For analysis (B-tagging, event reconstruction, applying selection cuts) a software framework called `tHqLoop` is used. Variables (kinematic variables of reconstructed objects) are saved on an event-by-event basis. The output of `tHqLoop` is saved as a `ROOT`-file, a file type specialized for processing with the analysis framework `ROOT` [76], which is based on C++. Variables in the output of `tHqLoop` can be selected and processed in order to plot histograms.

4.2.4 Properties of the used tHq sample

As the branching ratio $\mathcal{B}(H \to \tau\tau) = 6.26\,\%$ is quite small in comparison to the other decay channels shown in table 2.6, a generic sample of tHq events would contain mostly those events from other

Higgs-boson decays. As events of the decay channels $H \to b\bar{b}$ and $H \to gg$ would not pass the selection cuts, most events of the sample would not be used. This would lead to either bad statistics for the $\tau\tau$-state or to a huge file size. For this reason, a tHq sample specialized on multi-lepton final state analysis is generated. It accepts only events with at least two leptons (both light leptons and taus are accepted) in the final state in order to enrich the sample with events in the signal region.

4.3 Selection cuts

4.3.1 General motivation

Because tHq, the single top quark associated Higgs production, is of special interest in the work presented in this thesis, it is called signal in order to distinguish it from various other processes, which also occur at proton-proton-collisions and are also included in the Run 2 data of the LHC. The other processes are called background. The first step on the way of finding the signal is to set specific critera of the signal process, called selection cuts. Events, which do not fulfill these criteria are discarded in order to increase the ratio of signal to background processes. However, the cuts described in section 4.3.2 do not suffice to eliminate all background events. Since tHq is a very rare process, even after applying the cuts, the signal region is dominated by background events. Table 4.3 shows expected number of events, N_{cut}, that passed the selection cuts in table 4.2 and its weighted[1] statistical error, s_{cut}, based on simulations in comparison to data.

4.3.2 Finding suitable selection cuts for tHq in $2\,e/\mu + 1\tau_{\text{had}}$

For every event to be potentially signal the number of light leptons, hadronic taus and b-jets has to match the number expected for the decay channel described in section 4.1. Including the b-quark and the light quark in the final state (see figure 4.1), the total number of jets must be 2 as a minimum but may be higher, due to additional radiation jets. Jets from spectator quarks are suppressed by the thresholds $|\eta_{\text{jet}}| < 4.5$ and $p_{\text{jet}}^{\text{T}} > 25\,\text{GeV}$. For the b-jet, the light leptons and the hadronic taus, the threshold for η is chosen to be $|\eta| < 2.5$. The threshold $\not{E}_{\text{T}} > 5\,\text{GeV}$ accounts for the missing neutrinos expected from the leptonic decays of the top quark and the tau. The detailed selection cuts are given in table 4.2.

Additional to the selection cuts already mentioned following vetos are defined: muons with $|\eta| < 0.01$ are vetoed because this region is not fully covered by muon chambers for technical reasons. Taus and electrons with $1.37 < |\eta| < 1.52$ are vetoed because the identification in the transition region between barrel part and endcaps of the detector is less reliable.

4.3.3 Signal events of other decay channels

The selection cuts also allow events to pass, which are signal, but of another decay channel. Because the W boson and the Z boson can also decay both into hadronic taus and light leptons, as can be seen in tables 2.1 and 2.2, the decays $H \to WW$ and $H \to ZZ$ can lead to the same final state as $H \to \tau\tau$, for example if both W bosons of $H \to WW$ decay into taus and neutrinos, or if one Z boson of $H \to ZZ$ decays into a pair of taus and the other Z boson into a pair of neutrinos. Both for $H \to WW$ and for $H \to ZZ$, however, a successful mass reconstruction of the $\tau\tau$-state will not result in the Higgs boson

[1] Each simulated event has to be reweighted in order to level out several effects of Monte Carlo mismodelling. This reweighting affects also the statistical errors. Some explanations about that can be found in [77].

- 3 leptons with $|\eta| < 2.5$ and $p_{l_1}^{\mathrm{T}} > 27\,\mathrm{GeV}$, $p_{l_2}^{\mathrm{T}} > 20\,\mathrm{GeV}$ and $p_{l_3}^{\mathrm{T}} > 20\,\mathrm{GeV}$, where l_1, l_2 and l_3 are p^{T}-sorted per definition
- 1 of the leptons has to be tagged to be a hadronically decaying tau
- 2 - 6 jets with $|\eta_{\mathrm{jet}}| < 4.5$ and $p_{\mathrm{jet}}^{\mathrm{T}} > 25\,\mathrm{GeV}$
- 1 of the jets has to be tagged to be a b-jet with $|\eta_{b\text{-jet}}| < 2.5$ and $p_{b\text{-jet}}^{\mathrm{T}} > 25\,\mathrm{GeV}$
- $\not{E}_{\mathrm{T}} > 5\,\mathrm{GeV}$

Table 4.2: Selection cuts for tHq in the $2\,e/\mu + 1\tau_{\mathrm{had}}$ decay channel.

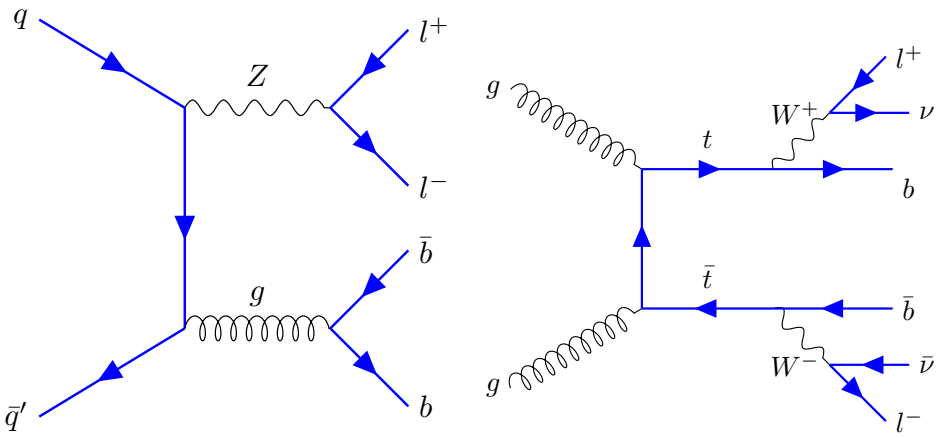
	N_{cut}	s_{cut}
tHq (multi-lepton)	2.04	0.04
tZq	41.1	0.5
$t\bar{t}$	4 111	13
tW	197	5
W+jets	4.8	1.1
Z+jets	2 690	60
Diboson	148.4	2.0
$t\bar{t}W$	52.8	0.6
$t\bar{t}Z$	111.9	0.9
$t\bar{t}H$	52.51	0.22
tWZ / tWH	20.75	0.13
Other	10.7	2.0
Total	7 450	60
Data	7 749	

Table 4.3: Number of events, N_{cut}, that pass the selection cuts listed in table 4.2. s_{cut} means the weighted statistical error. All numbers (except for the line named "data") are `SingleTop Ntuples` of version 32. For tHq, the multi-lepton cut described in section 4.2.4 is used.

mass. Apart from that, the selection cuts also allow that the hadronic tau, which is supposed to originate from the Higgs boson, actually originates from the top quark and that both light leptons originate from the Higgs boson. In this case it is not even necessary for the Higgs boson to decay into a pair of taus but into a pair of any leptons. Although this case is not included in the definition of the decay channel, it can not be rejected by the selection cuts. Also for this case, a successful mass reconstruction assuming the two taus to originate from the same mother particle will not result in the Higgs boson mass.

4.3.4 Sources of background

Obvious background events of tHq are events with a very similar intermediate state, e.g. tZq. The Z boson can also decay into two taus leading to the same final state as tHq. However, as discussed in section 3.4, the object reconstruction does not always work perfect and misidentification is possible.



(a) Z+jets (b) $t\bar{t}$

Figure 4.2: Feynman diagrams of the most dominant background processes, $t\bar{t}$ (b) and Z+jets (a), common non-prompt background.

Therefore, also processes with a very different intermediate, like Z+jets or $t\bar{t}$, can pass the selection cuts. This can happen if light-flavored jets are misidentified as hadronic taus or b-jets or if b-jets are misidentified as light leptons. Leading order Feynman diagrams for the latter case are shown in figure 4.2. Although these background processes can only pass the selection cuts in case of misidentification, they are the main background processes (see table 4.3) because they are very frequent. As already discussed in [77], however, their contribution is much lower, if only leptons with same electric charge (*SS* events) are selected.

4.4 Mass reconstruction

As stated in section 4.3.4, the process tZq can decay through the same decay channel as tHq. Separation of these two processes is especially challenging. One way of facing this challenge is the mass reconstruction of the $\tau\tau$-state, which has the same invariant mass as its mother particle, which is either an Higgs boson or a Z boson. For the mass reconstruction, the four-momenta of the final state decay products of the $\tau\tau$-state have to be added.

4.4.1 Challenges of mass reconstruction

Two difficulties arise:

1. In the $2\,e/\mu + 1\tau_{\text{had}}$-channel there are two light leptons in the final state. Which light lepton belongs to the $\tau\tau$-state and which to the top quark?

2. Both in case of a leptonic tau decay and of a hadronic tau decay, there are neutrinos in the final state which cannot be detected by the ATLAS detector. How to reconstruct their four-momenta?

4.4.2 Current status of lepton assignment

If the light leptons in the final state have opposite electric charge (*OS* events), the leptons can clearly be assigned as follows: Assuming the hadronic tau is originating from the Higgs boson and is identified correctly, also its charge is known. Because the $\tau\tau$-state is electrically neutral, regardless if it originates from the Higgs boson or from the Z boson, the electric charge of the light lepton from the $\tau\tau$-state must be opposite to the electric charge of the hadronic tau. The other light lepton then is assigned to the top quark. First problem of this method is that it fails for the cases in which the hadronic tau originates from the top quark. Second problem is that it can not be applied for events, in which the light leptons in the final state have the same electric charge (*SS* events). For now, these events have to be discarded and statistics are lost. As the second problem concerns a higher fraction of events than the first one, the studies presented in chapter 6 focus on including SS events.

4.4.3 Current status of reconstructing neutrino four-momenta

Visible mass First of all, one method to handle the missing neutrino four-momenta, is to simply ignore them. The mass calculated this way is called visible mass. It is not supposed to provide satisfying results, but it can be seen as a benchmark. Any other, more advanced approch for mass reconstruction has to deliver better results than the visible mass.

Missing transverse energy More advanced approaches exploit the *real* missing transverse energy (see section 3.4.7), $\overrightarrow{\not{E}}_{\mathrm{T}}$ (event). However, this cannot fully be assigned to the $\tau\tau$-state, as there is also one neutrino from the top-quark decay, which also contributes to $\overrightarrow{\not{E}}_{\mathrm{T}}$ (event). Therefore, $\overrightarrow{\not{E}}_{\mathrm{T}}$ ($\tau\tau$) is gained by subtraction:

$$\overrightarrow{\not{E}}_{\mathrm{T}}\,(\tau\tau) = \overrightarrow{\not{E}}_{\mathrm{T}}\,(\text{event}) - \overrightarrow{\not{E}}_{\mathrm{T}}\,(\text{top}) \tag{4.1}$$

Both $\overrightarrow{\not{E}}_{\mathrm{T}}$ ($\tau\tau$) and $\overrightarrow{\not{E}}_{\mathrm{T}}$ (top) are not known and the decomposition is a task in itself. On the basis of `SingleTop Ntuples` of version 29, there are already studies following this approach [77]. It is called "old approach" and will be evaluated and continued in chapter 5.

Transverse mass

After the decomposition, three different methods are already known to reconstruct the Higgs boson mass. The first method, called transverse mass method, uses the following formula:

$$\left(m^{\mathrm{T}}\right)^2 = \left(P_{\mathrm{vis}} + P_{\mathrm{invis}}\right)^2, \tag{4.2}$$

where P_{vis} is the sum of the four-momenta of all visible decay products and P_{invis} is a four-momentum, filled with the components $\left(\sqrt{(\not{E}^x)^2 + (\not{E}^y)^2}, \not{E}^x, \not{E}^y, 0\right)$. Since the Minkowski scalar product $P_{\mathrm{invis}} P_{\mathrm{invis}}$ vanishes, it can be seen that the formula in equation 4.2 is equivalent to the formula in [79]. Major shortcoming of this method is that, as p_ν^z cannot be inferred from $\overrightarrow{\not{E}}_{\mathrm{T}}$, it is simply ignored. Two remaining methods, the collinear approximation and the missing mass calculator, are more advanced and try to overcome this shortcoming. They are explained in the subsequent subsections 4.4.4 and 4.4.5 respectively.

4.4.4 Collinear approximation

Two basic assumptions are needed for the method of collinear approximation:

First assumption $\overrightarrow{\not{E}_{\mathrm{T}}}$ ($\tau\tau$-state) is only due to the invisible decay products of the taus. Their momenta are named $\vec{p}_{\text{invis, 1}}$ and $\vec{p}_{\text{invis, 2}}$ respectively and defined by subtracting the momenta of the visible decay products, $\vec{p}_{\text{visib, 1}}$ and $\vec{p}_{\text{visib, 2}}$, from the momenta of the taus $\vec{p}_{\tau_1}$ and $\vec{p}_{\tau_1}$.

$$\begin{aligned} \vec{p}_{\text{invis, 1}} &= \vec{p}_{\tau_1} - \vec{p}_{\text{visib, 1}} \\ \vec{p}_{\text{invis, 2}} &= \vec{p}_{\tau_2} - \vec{p}_{\text{visib, 1}} \end{aligned} \tag{4.3}$$

In case of a hadronic decay, $\vec{p}_{\text{invis, 1}}$ is the momentum of the tau-neutrino coming from the decay of τ_1. In case of a leptonic decay, $\vec{p}_{\text{invis, 1}}$ is the sum of the momentum of the tau-neutrino and the additional neutrino corresponding to the light lepton. With this definitions, the assumption can be written as:

$$\overrightarrow{\not{E}_{\mathrm{T}}} = \vec{p}_{\text{invis, 1}} + \vec{p}_{\text{invis, 2}} \tag{4.4}$$

In components:

$$\begin{aligned} \not{E}^x &= p^x_{\text{invis, 1}} + p^x_{\text{invis, 2}} \\ \not{E}^y &= p^y_{\text{invis, 1}} + p^y_{\text{invis, 2}} \end{aligned} \tag{4.5}$$

Second assumption The momenta of decay products of each tau are collinear. While this assumption is not true in general, it is approximately true for events with a high initial boost of the $\tau\tau$-state. Two different ways are known how to formulate this idea mathematically and derive formulas for $m_{\tau\tau}$. They both depend on $\vec{p}_{\text{visib, 1}}$, $\vec{p}_{\text{visib, 2}}$ and $\overrightarrow{\not{E}_{\mathrm{T}}}$ and are referred to as formula 1 [79] and formula 2 [80]. They are presented in appendix A. However, for events with low initial boost, this method fails for both formulas and overestimates the mass, leading to a long tail in the mass distribution. The obvious reason is that for a low boost, the assumption is not valid. In the limit case of the taus flying back-to-back to each other, formula 1 diverges [79]. Less obvious but also striking is that for a low boost, $\overrightarrow{\not{E}_{\mathrm{T}}}$ is reconstructed poorly because the angle between the neutrinos gets larger and their momenta partially cancel out in the reconstruction of $\overrightarrow{\not{E}_{\mathrm{T}}}$.

4.4.5 Missing Mass Calculator (MMC)

Also the MMC method assumes that all missing transverse energy is only due to neutrinos. But it tries to overcome the shortcomings of the collinear approximation method by dropping the other assumption of the tau-momenta being collinear. The idea is as follows: In order to fully reconstruct the mass of the $\tau\tau$-state, also all 4 components of both invisible decay products have to be known, leading to 8 unknowns in total. A leptonic tau decay diminishes the number of unknowns by 1, because a single-neutrino state can be assumed to be massless. So, if both taus decay leptonically, only 6 unknowns remain. Additionally, 4 equations constrain the system: equation 4.5 and the mass equations for the two taus, which are produced on-shell. Depending on the number of leptonic tau decays, the solution set has 2 to 4 dimension, and the system of equations can not be solved uniquely. However, not all solutions are equally likely. A global event probability $\mathcal{P}_{\text{event}}$ is defined as a function of a characteristic parameter and

used for minimizing a likelihood-function and thus finding a "most-probable" solution. One example for such a parameter is ΔR_{ij}, where i denotes the visible and j the invisible decay products. More details about this method can be found in [79].

CHAPTER 5

Reconstructing the missing transverse energy from the top quark

5.1 Determining the strategy

5.1.1 Description of the old approach

The old approach for reconstructing $\overrightarrow{\not{E}_{\mathrm{T}}}$ (top), the missing transverse energy from the top quark, is as follows [77]: Assuming that the top quark and the W boson are produced on-shell gives two constraints on the neutrino four-momentum P_ν:

$$m_t^2 = (P_b + P_l + P_\nu)^2 \tag{5.1}$$

$$m_W^2 = (P_l + P_\nu)^2 \tag{5.2}$$

The mass of the neutrino can be neglected.

$$m_\nu = 0 \tag{5.3}$$

So, there is only one kinematic variable left. For convenience, this is chosen to be p_ν^z. As both the neutrino and the light lepton originate from the same mother particle, the W boson, it is reasonable to assume that the kinematics of the neutrino are related to that of the light lepton. As a simple model to start, it is assumed that p_ν^z depends linearly on p_l^z, where the factor α is a parameter to be found out in a second step.

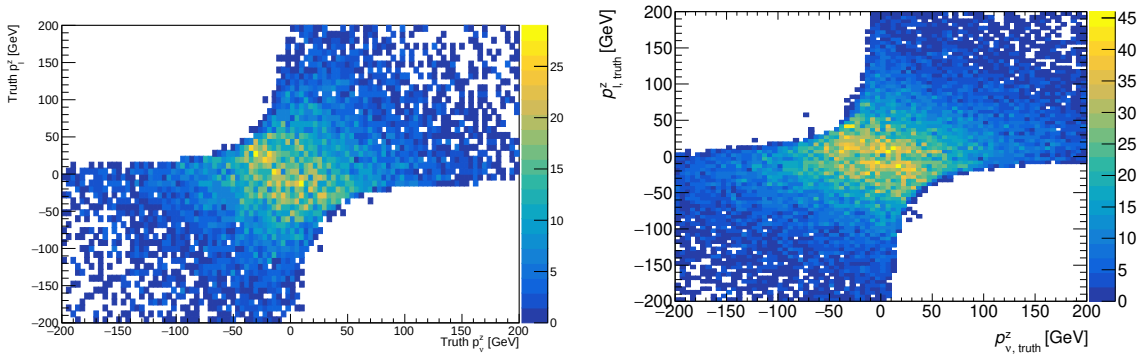
$$p_\nu^z = \alpha \cdot p_l^z \tag{5.4}$$

Treating p_ν^z as a parameter, equations 5.1 and 5.2 represent two-dimensional functions depending on both p_ν^x and p_ν^y to the second power.[1] Those are transformed into one quartic function depending only on p_ν^y. After solving this for p_ν^y its solution can be plugged into one of the second power equations in order to get a solution for p_ν^x. Detailed calculations are documented in [77]. In order to get numerical values as solutions the parameter α is set to $\alpha = 0.5$.

[1] This can be seen by expanding equations 5.1 and 5.2 and using the definition of the Minkowski scalar product in equation 2.3.

5.1.2 Drawbacks of the old approach

The old approach [77] for reconstructing $\overrightarrow{\not{E}}_{\mathrm{T}}$ (top) has its drawbacks: First, the equation for p_ν^y is fourth-power, leading to multiple and even imaginary solutions. More precisely, almost all events belong to one of the following categories: no solution, two real solutions, or two imaginary solutions. The cases of one, three or four solutions are possible but appear to be extremely rare. Events without solutions or with imaginary solutions are discarded. The fraction of events with real solutions depends on the coefficient α in equation 5.4. α is chosen to maximize the ratio of real-solution-cases to imaginary-solution-cases, but still this fraction is less than 50 %. By now there is no recipe for how to include events with imaginary solutions or how to choose one of the two real solutions. Hence, already in [77] truth studies have been performed (on version 29 `ntuples`- and on tZq) in order to evaluate how well p_ν^z is predicted by the constraint of equation 5.4 and in order to find a better constraint. While the truth studies indicated that the constraint of equation 5.4 is not perfect, they did not lead to finding a better constraint.



(a) (b)

Figure 5.1: Correlation of p_l^z with p_ν^z on *truth-level*. Figure (a) is taken from [77] and shows the correlation for tZq-events of version 29 `ntuples`. Only events of campaign mc16a have been used, and events with anti-top quarks have been removed. The remark in [77] that $\alpha = 0.5$ has been chosen before suggests that events without real solutions for p_ν^z on *reco-level* have been removed as well. Figure (b) shows the correlation for tHq events of version 32 `ntuples`, including events with anti-top quarks and of all campaigns.

5.1.3 Comparing old with new truth studies

The truth studies mentioned in the section before are taken on for new analysis. By starting the work on the analysis version 32 `ntuples` are available. Since version 29 `ntuples`, which have been used for the truth studies reported in [77], truth information in `TopPartons` have been renewed. Having implemented them in `tHqLoop`, more statistics are available as well. Performing the truth studies more thoroughly and with the new `ntuple` version, it is expected to evaluate the constraint on p_ν^z more precisely and in the end find a better constraint by searching for correlations on *truth-level*. Additionally, the simulated events are tested for deviations between *reco-level* and *truth-level*, which has also performed on tZq events with version 29 `ntuples`. The direct comparison shows that the deviations are much smaller now. In the end, simulations have to be performed on *reco-level*, so too large deviations would mean that results obtained on *truth-level* may not be transferable to *reco-level*. Details are documented in appendix B.

The first thing to do is to check the correlation between p_l^z and p_ν^z. This was also part of the truth studies in [77]. It is a test for the constraint in equation 5.4. For this, a two-dimensional histogram is filled with events: one axis showing p_l^z, the other axis showing p_ν^z of each event. The number of events per bin is indicated by a color. The scale, which shows which number of events belongs to which color, is shown next to the plotting area. In the context of this thesis, those types of two-dimensional histograms will be called color plots for short. Figure 5.1 shows the color plot of the old truth studies in comparison with the new one. Note that in figure 5.1(a) the color plot is filled with tZq-events, because in version 29 the statistics for tHq had been not enough.

Description Figures 5.1(a) and 5.1(b) have some obvious similarities: The area populated by events is confined by two hyperbolic lines. These are the consequence of the fixed value for the m_W, prohibiting pairs of variates for p_ν^z and p_l^z, which would violate four-momentum conservation. Also common to both color plots is the higher density of events around the origin for relatively low values of p_ν^z and p_l^z respectively. This connected area with event densities coded in yellow and orange will be called a cluster in the context of this thesis.

Evaluation Differences between figure 5.1(a) and figure 5.1(b) can be seen in the clusters. For figure 5.1(a) the shape of the cluster is difficult to determine since it is discontiguous. As described in [77] it is compatible with a weak anti-correlation. In contrast, the cluster in figure 5.1(b) is quite symmetric and indicates no correlation. However, for an absolute symmetric shape events are missing for all p_ν^z and for p_l^z around 0. This can be explained by the selection cut, which sets boundaries for the minimal p^z of the leptons. For the neutrinos, there is no corresponding cut, as they can not be detected anyway. In principle, this shape due to the selection cuts is also expected for figure 5.1(a). Presumably, it is not visible due to less statistics in general.

Conclusion So, if there is a correlation between p_ν^z and p_l^z, it is much too weak to work as a constraint. This can be inferred from figure 5.1(b). Examining figure 5.1(a), it is not that clear but compatible. While on the one hand this is disappointing, on the other hand it means that any constraint founded on a stronger correlation should improve the prediction of $\overrightarrow{\not{E}}_\mathrm{T}$ (top). Therefore, correlations in other kinematic variables of neutrino and lepton are searched for, as reported in the next section.

5.1.4 Searching for new correlations

In general, there is no preferred direction neither for the neutrino nor for the lepton, as there is no preferred flight direction of the objects in tHq at all. However, both particles originate from the W boson with a fixed mass, assuming it to be on-shell. This assumption is affirmed by the hyperbolic lines in figure 5.1. So with fixed m_W and knowing the kinematics of the lepton, the kinematics of the neutrino can at least not be totally arbitrary. In addition, in the rest frame of the W boson, the neutrino and the lepton fly back-to-back to each other due to momentum conservation. In case the rest frame of the W boson would be the lab frame, this constraint would actually easily allow predicting the kinematics of the neutrino. Although the lab frame and the rest frame of the W boson are not the same, and the lab frame is in principle different for every event, hopefully, some relations due to this constraint survive the transformation between the frames.

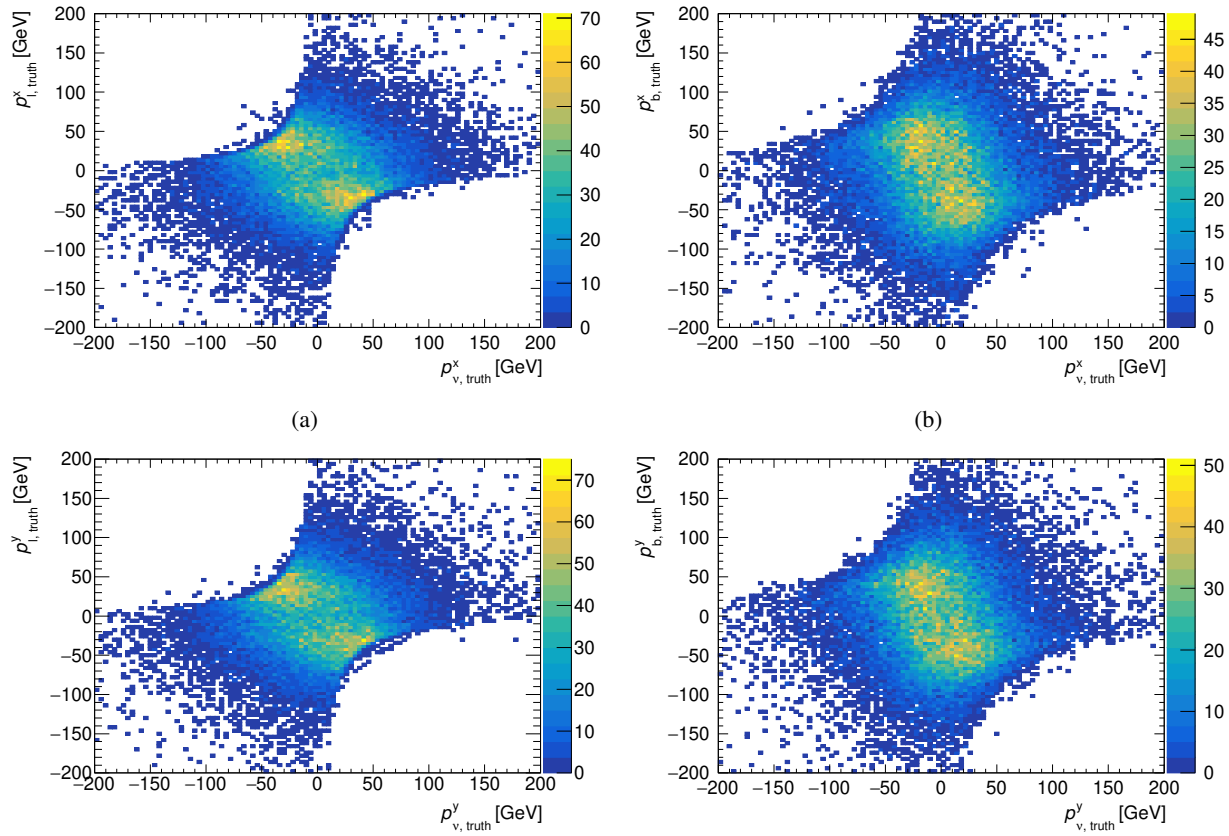


Figure 5.2: Correlation of p_l^x with p_ν^x (a), of p_b^x with p_ν^x (b), of p_l^y with p_ν^y (c) and of p_b^y with p_ν^y (d) on *truth-level* for tHq events of version 32 `ntuples`.

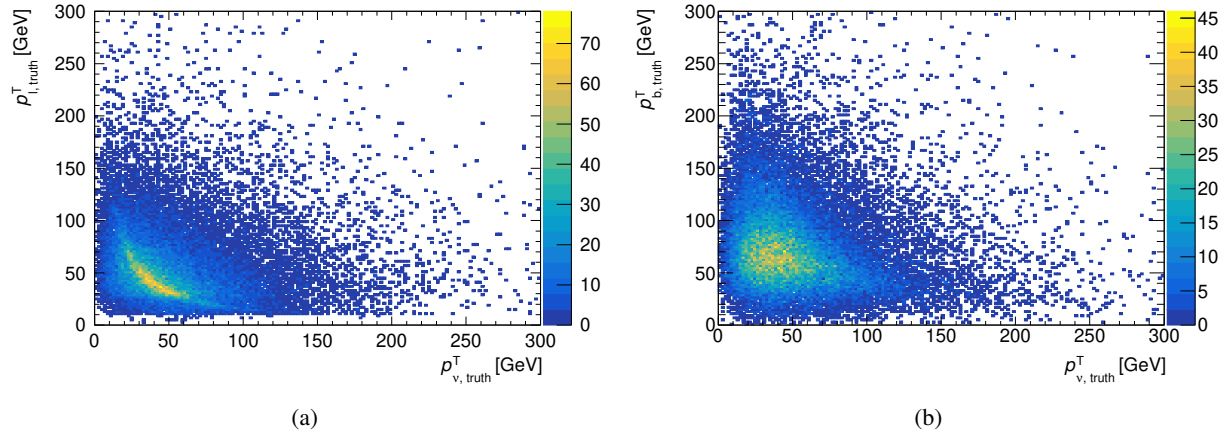
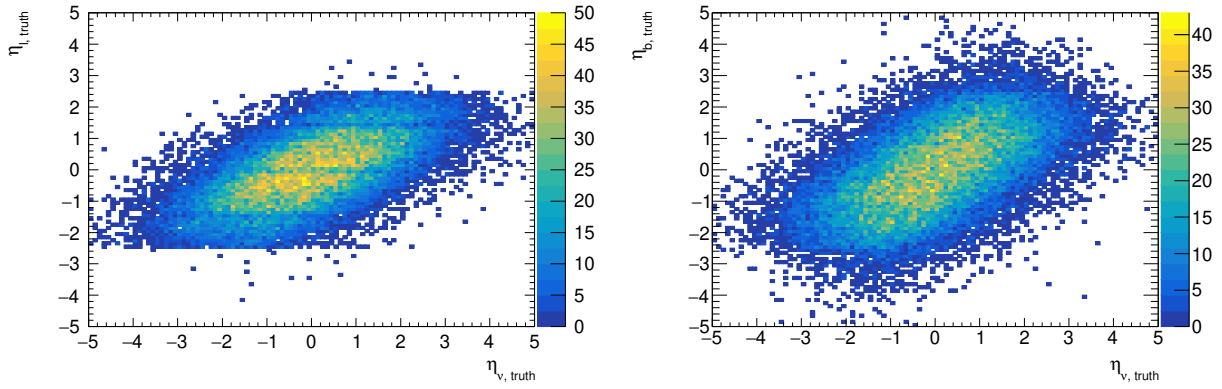
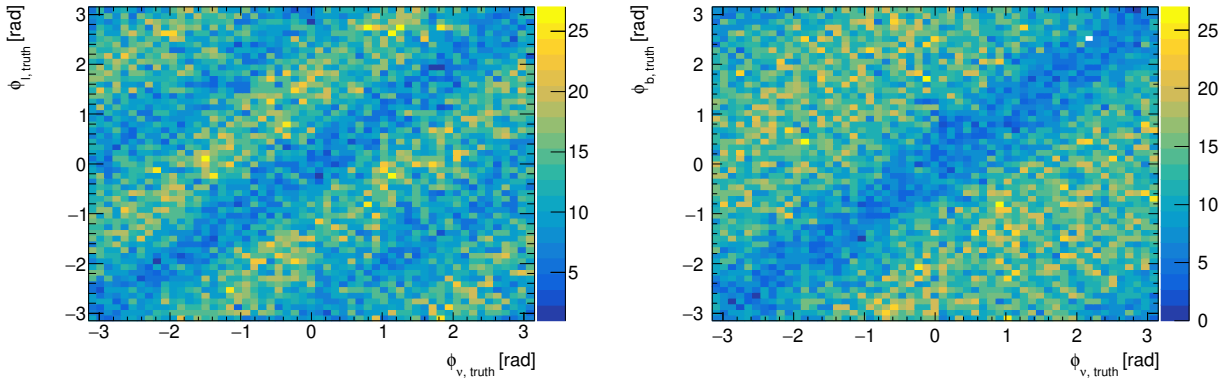


Figure 5.3: Correlation of p_l^T with p_ν^T (a)) and of p_b^T with p_ν^T (b) on *truth-level* for tHq events of version 32 `ntuples`.



(a) (b)

Figure 5.4: Correlation of η_l with η_ν (a) and of η_b with η_ν (b) on *truth-level* for tHq events of version 32 `ntuples`.



(a) (b)

Figure 5.5: Correlation of ϕ_l with ϕ_ν (a) and of ϕ_b with ϕ_ν (b) on *truth-level* for tHq events of version 32 `ntuples`.

The search for correlations covers all three-momentum components and begins with p^x and p^y, as p^z has already been handled. Additionally, the studies also include the kinematic variables p^{T}, ϕ and η, which are more suitable for the symmetries of the collider physics setup. Because also the top quark is assumed to be on-shell, also the kinematics of the b-quark should be correlated to the kinematics of the neutrino, and the search for correlations is extended to include also the b-quark. However, the correlation is expected to be smaller because there are more degrees of freedom with three decay particles.

Color plots for p^x and p^y are shown in figure 5.2 and for p^{T}, ϕ and η in figures 5.3, 5.5 and 5.4 respectively. Comparing all of those, the most distinct correlation pattern can be seen for p^{T} and ϕ. As expected, the kinematics of the neutrino are more strongly correlated to that of the lepton than to that of the b-quark.

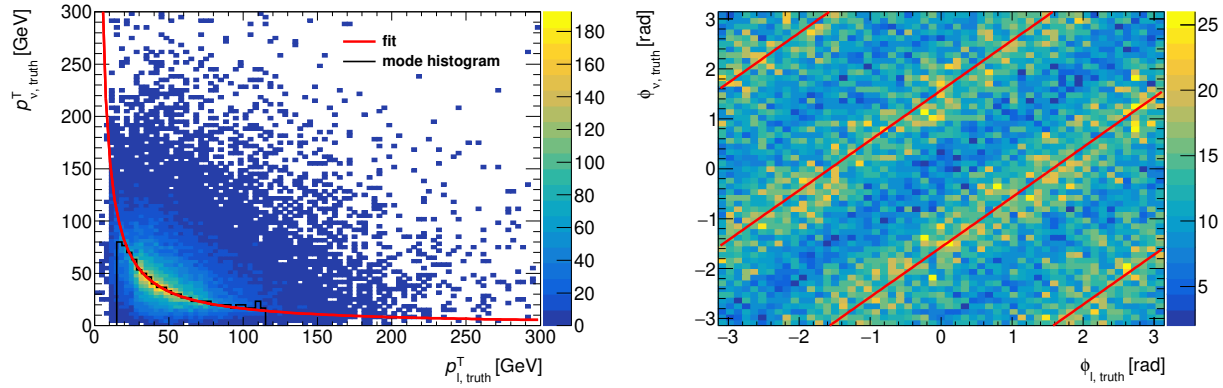
5.1.5 Deciding for a new strategy

The most distinct correlations have been found in "collider physics coordinates". In contrast, the old approach using the mass constraints uses Cartesian coordinates, because the addition of four-vectors

results in adding the Cartesian components, which is not true for any other set of coordinates. It is possible to formulate the fourth-power equation for p_ν^z and plugin $p^x = p^{\mathrm{T}} \cos\phi$ or $p^y = p^{\mathrm{T}} \sin\phi$ to solve the second-power polynomial for the remaining coordinate. However, as both p^{T} and ϕ are needed to calculate one of p^x and p^y, it is appealing to use both p^x and p^y. Unfortunately, this would over-constrain the system, probably tightening the issue of events without a solution. This leads to the idea and decision to not use the mass constraints at all but to simply use $p^x = p^{\mathrm{T}} \cos\phi$ and $p^y = p^{\mathrm{T}} \sin\phi$ for $\not{E}^x$ and $\not{E}^y$ respectively. This strategy has major advantages. First of all, it is much simpler and therefore less error-prone and - maybe even more important - it gets rid of the issue of multiple and imaginary solutions. The latter point has two positive consequences: Firstly, no events have to be discarded, and secondly for the rest of the events there is no problem with the ambiguity of solutions. With the new strategy, an unambiguous solution shall be obtained for every event. The mass constraints can be used for optimization utilizing a kinematic fit for instance.

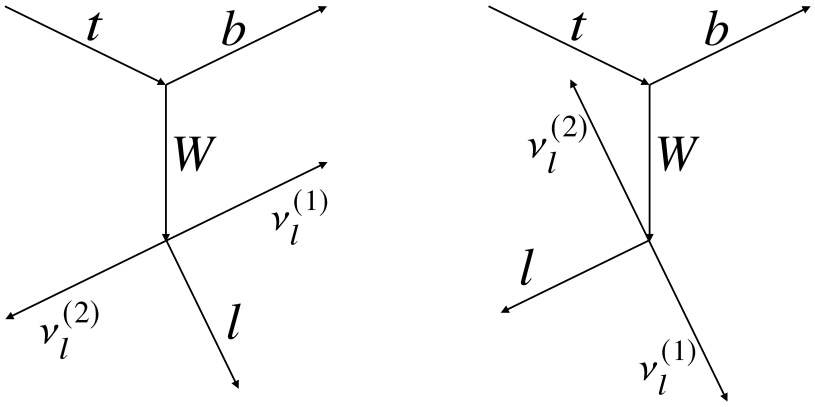
5.2 Transforming correlations into predictions

Having the strategy fixed to using the correlations between p_l^{T} and p_ν^{T}, respectively between ϕ_l and ϕ_ν, it remains the task to transform the correlations into an analytic relation, which, of course, is not exactly true for all events but approximately true for as many events as possible. Values for p_ν^{T} and ϕ_ν that are predicted using this analytic relation will be called $p_{\nu,\,\mathrm{pred}}^{\mathrm{T}}$ and $\phi_{\nu,\,\mathrm{pred}}$ respectively in order to clearly distinguish them from $p_{\nu,\,\mathrm{truth}}^{\mathrm{T}}$ and $\phi_{\nu,\,\mathrm{truth}}$, which have been considered so far.



(a) (b)

Figure 5.6: Correlation of p_ν^{T} with p_l^{T} (a) and of ϕ_ν with ϕ_l (b), both on *truth-level*. In figure (a) for every position in a selective range of the horizontal axis, the mode of all vertical bins corresponding to this very horizontal position is taken as an entry of a 1D-histogram, called mode (drawn in black). The red line shows a fit of the model $p_\nu^{\mathrm{T}} = \frac{f}{p_l^{\mathrm{T}}}$ to the mode histogram, which gives $f = (1\,615.98 \pm 5.15)\ \mathrm{GeV}^2$. For figure (b) the red lines are drawn by eye to highlight areas enriched with events, which are interpreted to be favored. The four red lines fulfill $\phi_\nu = \phi_l \pm \frac{\pi}{2}$ and $\phi_\nu = \phi_l \pm \frac{3\pi}{2}$ respectively, which can be reduced to $\phi_\nu = \phi_l \pm \frac{\pi}{2}$, as $\phi + \frac{3\pi}{2} = \phi - \frac{\pi}{2}$ and $\phi - \frac{3\pi}{2} = \phi + \frac{\pi}{2}$.



(a) Best case:
$\nu_l^{(2)}$ clearly favored, $\nu_l^{(1)}$ clearly disfavored

(b) Worst case:
both $\nu_l^{(2)}$ and $\nu_l^{(1)}$ are equally (dis-)favored.

Figure 5.7: Flight directions of the top quark and its decay products both for $\phi_l - \phi_b = \pm\frac{\pi}{2}$ (a) and $\phi_\nu = \phi_b \pm \pi$ (b) in a schematic. The case of $\phi_l - \phi_b = \pm\frac{\pi}{2}$, shown in (a) leads to one solution that is parallel to the b-quark (suppressed) and one anti-parallel (favored). Because of this clear distinction, this case is called "best case". In the case $\phi_\nu = \phi_b \pm \pi$, shown in (b), both neutrino-solutions traverse perpendicularly to the b-quark and are equally (dis-)favored. Therefore this case is called "worst case".

5.2.1 Transverse momentum p_ν^{T}

The analytic relation between p_ν^{T} and p_l^{T} is extracted from the correlation seen in figure 5.3(a) by applying an analytic fit. As p_ν^{T} is unknown and should be determined as a function of p_l^{T}, p_ν^{T} is plotted on the y-axis and p_l^{T} on the x-axis. For fitting, however, the 2D histogram has to be transformed into a 1D histogram. In order to do so for every position of the horizontal axis the mode of all vertical bins corresponding to this very horizontal position is taken as an entry of a 1D-histogram, called mode histogram. A coarser binning than in figure 5.3(a) is chosen in order to prevent the position of the mode bin being dependent on statistical fluctuations. The range for the mode histogram is limited to the range $[15\,\mathrm{GeV}, 115\,\mathrm{GeV}]$, where the mode is unique. A function of the form $p_\nu^{\mathrm{T}} = \frac{f}{p_l^{\mathrm{T}}}$ (where f is a fit parameter) is fitted to the mode histogram leading to $f = (1\,615.98 \pm 5.15)\,\mathrm{GeV}^2$. Both the mode histogram and the fit function are plotted onto the color plot allowing a consistency check. This can be seen in figure 5.6(a).

5.2.2 Azimuth angle ϕ_ν

The correlation of ϕ_ν with ϕ_l, already shown in figure 5.5(a), is shown again in figure 5.6(b). Axes are swapped so that ϕ_ν is shown on the y-axis and ϕ_l on the x-axis, which is consistent with figure 5.6(a). The correlation pattern consists of parallel bands of alternating color, which indicates the event density. Events lying in the yellow/orange band with higher event density are interpreted to be favored and events in the blue/turquoise band with lower event density are interpreted to be suppressed. In order to expose the favored events, the yellow/orange bands with higher event density are highlighted by red lines. Without a fit, it can be seen by eye that the two long red lines fulfill the conditions $\phi_\nu = \phi_l \pm \frac{\pi}{2}$. The two short red lines fulfill $\phi_\nu = \phi_l \pm \frac{3\pi}{2}$, which is equivalent to $\phi_\nu = \phi_l \pm \frac{\pi}{2}$, since because $\phi = \phi + k\,2\pi\ \forall k \in \mathbb{Z}$, it

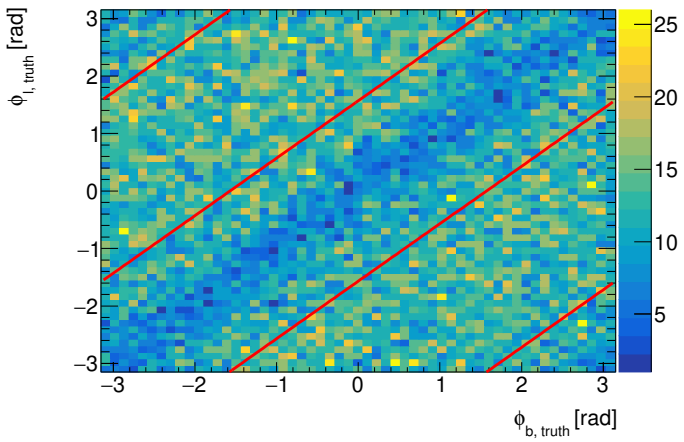


Figure 5.8: Correlation of ϕ_l with ϕ_b on *truth-level*. As in figure 5.5(a) the four red lines fulfill $\phi_l = \phi_b \pm \frac{\pi}{2}$ and $\phi_l = \phi_b \pm \frac{3\pi}{2}$ respectively. They allow to see that the event density for $\phi_l = \phi_b \pm \frac{\pi}{2}$ is similar to the event density for $\phi_l = \phi_b \pm \pi$ (the area between the first and the second and between the third and the fourth red line, counted from the top). Only for $\phi_l = \phi_b$ the event density drops.

is $\phi + \frac{3\pi}{2} = \phi - \frac{\pi}{2}$ and $\phi - \frac{3\pi}{2} = \phi + \frac{\pi}{2}$. Nevertheless, the ambiguity of the two solutions in $\phi_\nu = \phi_l \pm \frac{\pi}{2}$ remains. In fact, this ambiguity is crucial, because the difference between the two solutions is π, i.e. picking the wrong solution assumes a flight direction for the neutrino which is opposite to the actual direction: the largest deviation one could think of.

Problem of ambiguity

Because this is so crucial and can not be resolved with the ϕ_ν-ϕ_l-correlation alone, the ϕ_ν-ϕ_b-correlation is investigated, which is shown in figure 5.5(b). It implies that events with $\phi_\nu = \phi_b$ are suppressed. This relation alone is not very strong, but it might help with solving the ambiguity. However, it does not solve the ambiguity in every case. It only helps if one solution of $\phi_\nu = \phi_l \pm \frac{\pi}{2}$ leads to $\phi_\nu = \phi_b$ and the other one to the opposite. But this depends on the relative angle between b-quark and light lepton. Figure 5.7 shows in a schematic the flight directions of the top quark and its decay products both for $\phi_l - \phi_b = \pm\frac{\pi}{2}$ (figure 5.7(a)) and $\phi_\nu = \phi_b \pm \pi$ (figure 5.7(b)). The case of $\phi_l - \phi_b = \pm\frac{\pi}{2}$, shown in figure 5.7(a), using the constraint $\phi_\nu = \phi_l \pm \frac{\pi}{2}$ leads to one solution that is parallel to the b-quark (suppressed) and one anti-parallel (favored). It is called "best case". In case of $\phi_\nu = \phi_b \pm \pi$, shown in figure 5.7(b), both neutrino-solutions fulfill $\phi_\nu = \phi_l \pm \frac{\pi}{2}$. The correlation shown in figure 5.5(b) does not allow for preferring one solution over the other in that case because they are equally likely. Therefore, this latter case is called "worst case". In order to find out the respective fractions of those two cases, the ϕ_l-ϕ_b-correlation, shown in figure 5.8, is investigated: Like the correlation between ϕ_ν and ϕ_b has shown that the case $\phi_\nu - \phi_b = 0$ is suppressed, the correlation between ϕ_l and ϕ_b also shows that the case $\phi_l - \phi_b = 0$ is suppressed. The relative angles of $\phi_l - \phi_b = \pi$ and $\phi_l - \phi_b = \pm\frac{\pi}{2}$, marked by red lines in figure 5.8, are equally favored. The relative angles $\phi_l - \phi_b = \pm\frac{\pi}{2}$ form the "best case", while $\phi_l - \phi_b = 0$ and $\phi_l - \phi_b = \pi$ form the "worst case". This means that due to the suppression of events with $\phi_l - \phi_b = 0$, the "worst case" should be to some extent less frequent than the "best case". This is promising for this approach to be tested.

Choosing the correct solution

The definition $\Delta\phi_{bl} := \phi_b - \phi_l$ is equivalent to the angle, measured beginning with the direction of the light lepton in the shortest path (ϕ defined in the range $[-\pi, \pi)$) onto the direction of the b-quark. If this path goes anti-clockwise, the angle is positive (as it is in figure 5.7(a)). In this case, the angle onto the direction of the favored neutrino solution is clockwise, i.e. $\phi_\nu = \phi_l - \frac{\pi}{2}$ is chosen to be the correct solution. One could also argue this way: As those neutrino solutions are favored which go in the opposite direction of the b-quark, the angle $\Delta\phi_{bl} := \phi_b - \phi_l$ and $\Delta\phi_{\nu l} := \phi_\nu - \phi_l$ must have the opposite sign. Without loss of generality:

$$\Delta\phi_{bl} > 0 \implies \Delta\phi_{\nu l} < 0 \implies \Delta\phi_{\nu l} = -\frac{\pi}{2} \implies \phi_\nu = \phi_l - \frac{\pi}{2}$$
$$\Delta\phi_{bl} < 0 \implies \Delta\phi_{\nu l} > 0 \implies \Delta\phi_{\nu l} = +\frac{\pi}{2} \implies \phi_\nu = \phi_l + \frac{\pi}{2}$$

The case $\Delta\phi_{bl} = 0$ is still ambiguous but should be very rare and is treated as $\Delta\phi_{bl} > 0$.

5.2.3 Giving an analytic formula

Now it is possible to sum up the preceding results and give analytic formulas:

$$p^{\mathrm{T}}_{\nu,\,\mathrm{pred}} = \frac{f}{p^{\mathrm{T}}_l} \text{ with } f = (1\,615.98 \pm 5.15)\,\mathrm{GeV}^2 \tag{5.5a}$$

$$\phi_{\nu,\,\mathrm{pred}} = \begin{cases} \phi_l - \frac{\pi}{2}\,, \text{ if } & \phi_b - \phi_l \geq 0 \\ \phi_l + \frac{\pi}{2}\,, \text{ if } & \phi_b - \phi_l < 0 \end{cases} \tag{5.5b}$$

Transformed into Cartesian coordinates, this leads to:

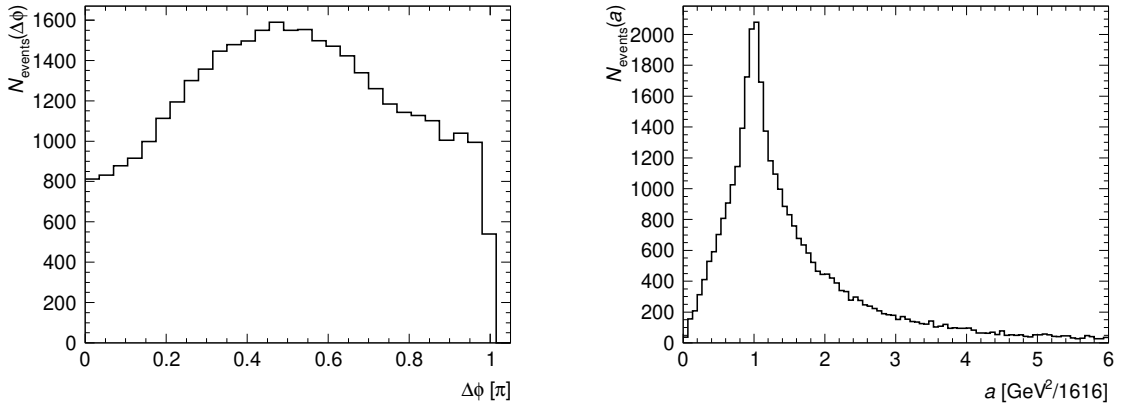
$$p^{x}_{\nu,\,\mathrm{pred}} = p^{\mathrm{T}}_{\nu,\,\mathrm{pred}} \cos\phi_{\nu,\,\mathrm{pred}} \tag{5.6a}$$
$$p^{y}_{\nu,\,\mathrm{pred}} = p^{\mathrm{T}}_{\nu,\,\mathrm{pred}} \sin\phi_{\nu,\,\mathrm{pred}} \tag{5.6b}$$

5.3 Explaining correlations with analytic functions for the event density

Section 5.1 reported the search for correlations between two variables. In a two-dimensional histogram, called color plot, which was filled with all events as a function of these two variables, areas enriched with events have been found. This phenomenon will be expressed by introducing a new quantity ρ, called event density, which is the number of events per histogram area. In a color plot it can be understood as a function of the two variables naming the two coordinate axes. Areas enriched with events mean that the event density is not constant over the whole histogram area but increased in these areas. It was possible to describe those areas with higher event density by formulating constraints between the two variables, namely $\phi_\nu - \phi_l = \pm\frac{\pi}{2}$ and $p^{\mathrm{T}}_\nu p^{\mathrm{T}}_l = f \approx 1\,616\,\mathrm{GeV}^2$. In section 5.2 these constraints have been used to predict p^x_ν and p^y_ν as a function of p^{T}_l, ϕ_l and ϕ_b. This section is about giving reasons why the event density is increased for the regions described by the constraints named above.

5.3.1 Basic considerations

Since the mass of the top quark is clearly higher than the mass of the W boson, the W boson can be assumed to be on-shell. The masses of the lepton and the neutrino from the W-boson decay, on the other hand, are negligibly small. By its decay, the rest mass of the W boson is therefore transformed into kinetic energy of lepton and neutrino for the most part. Likewise, the kinetic energy of the W boson is transformed into the kinetic energy of lepton and neutrino. The larger the kinetic energy of the W boson is compared to its rest mass, the stronger lepton and electron tend to fly in the same direction as the W boson, because of the boost. In the limit of infinite boost, lepton and neutrino fly collinear, i.e. with an infinitely small relative angle. In the opposite limit of no boost, they are expected to fly back-to-back to each other, i.e with a relative angle of π. The correlations of the previous section showed that the relative angle is $\frac{\pi}{2}$, i.e. it is neither the first limit case nor the second but a case in between. Why the relative angle is exactly $\frac{\pi}{2}$, however, is not clear and will be treated in the following subsections.



(a) (b)

Figure 5.9: Number of events as a function of $\Delta\phi := |\phi_l - \phi_\nu|$ (a) and as a function of $a := p_l^{\mathrm{T}} p_\nu^{\mathrm{T}}$ (b). The distribution of figure (b) is expected to peak at the fit value $f \approx 1\,616\,\mathrm{GeV}^2$ obtained from the fit in figure 5.6(a). The abscissa is scaled by this value for clarity.

5.3.2 Introducing one-dimensional event density

The first step is to reduce complexity by reducing the dimension of ρ. This is done by applying the constraints as follows. With the definition $\Delta\phi := |\phi_l - \phi_\nu|$, the constraint between the two variables ϕ_l and ϕ_ν can be expressed as a specific value for one variable, namely $\Delta\phi = \frac{\pi}{2}$.[2] Analogously, with the definition $a := p_l^{\mathrm{T}} p_\nu^{\mathrm{T}}$, the constraint between the two variables p_l^{T} and p_ν^{T} can be expressed as a specific value for the variable a, namely $a = f \approx 1\,616\,\mathrm{GeV}^2$. The event density that was displayed in figure 5.6(b) as a function of ϕ_l and ϕ_ν is now considered as a function of $\Delta\phi$ and called $\rho_{\Delta\phi}$. It

[2] Taking the absolute value in the definition of $\Delta\phi$ ignores the rotational direction between ϕ_l and ϕ_ν but this is fine since due to the rotational symmetry, there is no preferred rotational direction. This is consistent with the previous observation that the ambiguity between $\phi_l - \phi_\nu = +\frac{\pi}{2}$ and $\phi_l - \phi_\nu = -\frac{\pi}{2}$ could only be resolved by taking ϕ_b into account. In the following discussion, however, only the kinematic variables of the W-boson decay will be considered and thus, the symmetry remains unbroken.

is obtained by plotting the number of events as a function of $\Delta\phi$ in a one-dimensional histogram and dividing it by the bin width $\Delta(\Delta\phi)$. Analogously, the event density that was displayed in figure 5.6(a) as a function of p_l^{T} and p_ν^{T} is now considered as a function of a and called ρ_a. It is obtained by plotting the number of events as a function of a in a one-dimensional histogram and dividing it by the bin width Δa. In formulas:

$$\rho_a := \frac{N_{\text{events}}(a)}{\Delta a} \tag{5.7a}$$

$$\rho_{\Delta\phi} := \frac{N_{\text{events}}(\Delta\phi)}{\Delta(\Delta\phi)} \tag{5.7b}$$

For illustration, the one-dimensional histograms $N_{\text{events}}(\Delta\phi)$ and $N_{\text{events}}(a)$ are shown in figure 5.9.

Relating one-dimensional event density to two-dimensional event density

Figure 5.9(a) shows the number of events as a function of $\Delta\phi$.

The histogram shows a broad maximum at $\Delta\phi = \frac{\pi}{2}$ with about 1600 events. The distribution has an approximate symmetry around this maximum but this symmetry is not extact. On the left handside the distribution falls of to about 800 events for $\Delta\phi = 0$ while on the right hand side it falls of to about 1000 events for $\Delta\phi = \pi$. That the last bin shows less than 600 events is a binning effect. However, this broad maximum the one-dimensional distribution is conform with the pattern found in the two-dimensional distribution in figure 5.6(b). Moreover, the slight asymmetry could not be seen that clearly in figure 5.6(b).

Figure 5.9(b) shows the number of events as a function of a. The distribution has a steep peak at $a = f$. It is asymmetric and has a long tail for values $a > 1$. The events in the region around $a = f$ correspond to those events on or near the fit curve in figure 5.6(a). The events with $a < 0$ lie left-hand-sided and below the fit curve and the events with $a > 0$ lie right-hand-sided and above the fit curve. For events that are right-hand-sided and above the fit curve, higher distances to the fit curve are possible than for events that are left-hand-sided and below. Both the peak position and the asymmetric shape with the long tail for $a > 0$ conform with the two-dimensional histogram in figure 5.6(a).

The one-dimensional histograms in figure 5.9 are peaked for those values of $\Delta\phi$ and a which correspond to the constraints extracted from figure 5.6. The reasons for the peaks in the one-dimensional histograms will also explain the constraints found in the two-dimensional histograms. Since the bin widths $\Delta(\Delta\phi)$ and Δa are constant, peaks in the distributions of $N_{\text{events}}(\Delta\phi)$ and $N_{\text{events}}(a)$ are equivalent to the peaks of $\rho_{\Delta\phi}$ and ρ_a. Therefore, the following subsections are about giving reasons for peaks in the distribution of $\rho_{\Delta\phi}$ and ρ_a.

5.3.3 Link to theoretical principles

One possible way to explain the peaks in the distributions shown in figure 5.9, is to derive an analytical function for $\rho_{\Delta\phi}$ and ρ_a from theoretical principles. One important theoretical principle is that of equation 2.1, which relates the number of events of one interaction with the integrated luminosity and the cross section. In a more precise nomenclature, this reads as:

$$N_{\text{events}}^{\text{total}} = \sigma \times \int \mathcal{L}\,\mathrm{d}t \tag{5.8}$$

The new nomenclature has been chosen because the symbol N_{events}, as used as a label in the histograms, does not mean the total number of events, as used in equation 2.1. The total number of events is obtained as the integral of the respective histograms:

$$N_{\text{events}}^{\text{total}} = \int_0^{\infty} N_{\text{events}}(a)\,\mathrm{d}a = \int_0^{\pi} N_{\text{events}}(\Delta\phi)\,\mathrm{d}\Delta\phi \tag{5.9}$$

For a fixed value of a, or $\Delta\phi$ respectively, $N_{\text{events}}(a)$ and $N_{\text{events}}(\Delta\phi)$ represent a part of $N_{\text{events}}^{\text{total}}$, which depends on the bin width Δa and $\Delta\,(\Delta\phi)$. This can be illustrated by writing $\Delta N_{\text{events}}^{\text{total}}$ for N_{events}. Thereby, ρ_a and $\rho_{\Delta\phi}$, can also be written as:

$$\rho_a = \frac{N_{\text{events}}(a)}{\Delta a} = \frac{\Delta N_{\text{events}}^{\text{total}}(a)}{\Delta a} \tag{5.10a}$$

$$\rho_{\Delta\phi} = \frac{N_{\text{events}}(\Delta\phi)}{\Delta\,(\Delta\phi)} = \frac{\Delta N_{\text{events}}^{\text{total}}(\Delta\phi)}{\Delta\phi}\,, \tag{5.10b}$$

where the Δ in front of $\Delta N_{\text{events}}^{\text{total}}$ directly corresponds to the Δ in front of Δa, i.e these two quantities are defined such that if one Δ is scaled up or down, the other Δ gets automatically also scaled, since a specific amount of events $\Delta N_{\text{events}}^{\text{total}}(a)$ directly corresponds to a specific interval Δa. That precondition being given, both Δ's can be scaled down so that the event densities ρ_a and $\rho_{\Delta\phi}$ are equal to the derivates of $N_{\text{events}}^{\text{total}}$ with respect to a and $\Delta\phi$:

$$\rho_a = \frac{\Delta N_{\text{events}}^{\text{total}}(a)}{\Delta a} = \frac{\mathrm{d}N_{\text{events}}^{\text{total}}(a)}{\mathrm{d}a} \tag{5.11a}$$

$$\rho_{\Delta\phi} = \frac{\Delta N_{\text{events}}^{\text{total}}(\Delta\phi)}{\Delta\,(\Delta\phi)} = \frac{\mathrm{d}N_{\text{events}}^{\text{total}}(\Delta\phi)}{\mathrm{d}\Delta\phi} \tag{5.11b}$$

5.3.4 Deriving analytical functions for ρ_a, $\rho_{\Delta\phi}$ and $\rho_{\Delta\eta}$

At that point, the dependence of $N_{\text{events}}^{\text{total}}$ on a and $\Delta\phi$ is not apparent and has to be revealed in the following. $\int\mathcal{L}\mathrm{d}t$ is a property of the LHC and constant with respect to kinematic event variables. σ is the cross section of tHq considering the decay channel and the selection cuts. As stated in equation 2.9, this depends on the matrix element (which can be obtained by evaluating the Feynman diagram in figure 4.1) integrated over the phase space. The matrix element is a product of the initial state, propagator and final state, i.e. it depends only linearly on those terms. This implies that the matrix element for the whole process can be written as a product of matrix elements for each decay step, and this means that the matrix element of the W boson decay from the top quark can be factored out of the rest. Using mathematical formulas this means

$$\sigma = \int |\mathcal{M}_{tHq}|^2\,\mathrm{d}\rho = \int |\mathcal{M}_{W\,\text{decay}}\mathcal{M}_{\text{rest}}|^2\,\mathrm{d}\rho = \int |\mathcal{M}_{W\,\text{decay}}|^2\;|\mathcal{M}_{\text{rest}}|^2\,\mathrm{d}\rho \tag{5.12}$$

The matrix element for the W boson decay depends on the spin configuration. Since for the analysis in this thesis no spin configuration is favored, the matrix element has to be spin averaged. The result is: [6]

$$|\mathcal{M}_{W\,\text{decay}}|^2 = \frac{1}{3} g_W^2 m_W^2 \ , \tag{5.13}$$

where g_W is the weak coupling constant. As this expression does not depend on the phase space variables, it can be pulled out of the integral.

$$\sigma = \frac{1}{3} g_W^2 m_W^2 \times \int |\mathcal{M}_{\text{rest}}|^2 \, \mathrm{d}\rho \tag{5.14}$$

This expression for σ is now plugged into formula 5.8 leading to:

$$\begin{aligned} N_{\text{events}}^{\text{total}} &= \frac{1}{3} g_W^2 m_W^2 \times \int |\mathcal{M}_{\text{rest}}|^2 \, \mathrm{d}\rho \times \int \mathcal{L} \mathrm{d}t \\ &= k \times m_W^2 \ \text{ using the constant } \quad k := \frac{1}{3} g_W^2 \times \int |\mathcal{M}_{\text{rest}}|^2 \, \mathrm{d}\rho \times \int \mathcal{L} \mathrm{d}t \end{aligned} \tag{5.15}$$

In principle, all factors in equation 5.15 are constants, but there is an implicit dependence of m_W^2 on a and $\Delta\phi$ due to four-momentum conservation. Showing this will need some additional steps:

$$\begin{aligned} m_W^2 &= \left(P_l + P_\nu\right)^2 = P_l^2 + P_\nu^2 + 2 P_l P_\nu = m_l^2 + m_\nu^2 + 2 P_l P_\nu \\ m_W^2 &\approx 2 P_l P_\nu = 2 (E_l E_\nu - \vec{p}_l \cdot \vec{p}_\nu) \end{aligned} \tag{5.16}$$

Since $m_W^2 \gg m_l^2 + m_\nu^2$, the masses of the light lepton and the neutrino can be neglected leading to equation 5.16. This allows the transformation into the coordinates p^{T}, ϕ and η, which have already been used for searching correlations. Using the relations

$$\begin{aligned} p^x &= p^{\mathrm{T}} \cos(\phi) \qquad p^y = p^{\mathrm{T}} \sin(\phi) \qquad p^z = p^{\mathrm{T}} \sinh(\eta) \\ E &= \sqrt{\left(p^{\mathrm{T}}\right)^2 + \left(p^z\right)^2} = p^{\mathrm{T}} \sqrt{1 + \sinh^2(\eta)} = p^{\mathrm{T}} \cosh(\eta) \end{aligned} \tag{5.17}$$

and plugging them into equation 5.16 above leads to:

$$\begin{aligned} m_W^2 &= 2\, p_l^T \ p_\nu^T \left(\cosh\eta_l \cosh\eta_\nu - \cos\phi_l \cos\phi_\nu - \sin\phi_l \sin\phi_\nu - \sinh\eta_l \sinh\eta_\nu\right) \\ &= 2\, p_l^T \ p_\nu^T \left(\cosh(\eta_l - \eta_\nu) - \cos(\phi_l - \phi_\nu)\right) \\ &= 2a \left(\cosh\Delta\eta - \cos\Delta\phi\right) \ , \end{aligned} \tag{5.18}$$

where the definitions $a = p_l^{\mathrm{T}} p_\nu^{\mathrm{T}}$, $\Delta\phi = |\phi_l - \phi_\nu|$ and $\Delta\eta := \eta_l - \eta_\nu$ have been applied in the last step. Plugging equation 5.18 into equation 5.15 leads to:

$$N_{\text{events}}^{\text{total}} = k \times 2a \left(\cosh\Delta\eta - \cos\Delta\phi\right) \tag{5.19}$$

Now, the dependence of $N_{\text{events}}^{\text{total}}$ on a and $\Delta\phi$ is apparent and the event densities ρ_a and $\rho_{\Delta\phi}$ can be

calculated by derivation. For prospective discussions $\rho_{\Delta\eta}$ is also calculated:

$$\rho_a = \frac{\mathrm{d}N^{\text{total}}_{\text{events}}}{\mathrm{d}a} = k \times 2\,(\cosh\Delta\eta - \cos\Delta\phi) = \rho_a(\Delta\eta, \Delta\phi) \tag{5.20a}$$

$$\rho_{\Delta\phi} = \frac{\mathrm{d}N^{\text{total}}_{\text{events}}}{\mathrm{d}\Delta\phi} = k \times 2a \sin\Delta\phi = \rho_{\Delta\phi}(a, \Delta\phi) \tag{5.20b}$$

$$\rho_{\Delta\eta} = \frac{\mathrm{d}N^{\text{total}}_{\text{events}}}{\mathrm{d}\Delta\eta} = k \times 2a \sinh\Delta\eta = \rho_{\Delta\eta}(a, \Delta\eta) \tag{5.20c}$$

For further analysis of the functional form of the event densities, it is important to note that the variables a, $\Delta\phi$ and $\Delta\eta$ are not independent, but related to each other by equation 5.18, as m_W^2 is fixed. So it is possible to write a as a function of the other variables

$$a = \frac{m_W^2}{2\,(\cosh\Delta\eta - \cos\Delta\phi)} \tag{5.21}$$

This allows writing equations 5.20 in an alternative functional form:

$$\rho_a(a) = k \times \frac{m_W^2}{a} \tag{5.22a}$$

$$\rho_{\Delta\phi}(\Delta\eta, \Delta\phi) = k \times m_W^2 \frac{\sin\Delta\phi}{\cosh\Delta\eta - \cos\Delta\phi} \tag{5.22b}$$

$$\rho_{\Delta\eta}(\Delta\eta, \Delta\phi) = k \times m_W^2 \frac{\sinh\Delta\eta}{\cosh\Delta\eta - \cos\Delta\phi} \tag{5.22c}$$

5.3.5 Discussing functions for the event densities

Discussing $\rho_{\Delta\phi}$

In the functional form of equation 5.20b, $\rho_{\Delta\phi}$ linearly depends on $\sin\Delta\phi$, which vanishes for $\Delta\phi = 0$ and $\Delta\phi = \pi$ and is maximal for $\Delta\phi = \frac{\pi}{2}$. The pattern in figure 5.9(a) shows an offset of about 800 events for $\Delta\phi = 0$ and of about 1000 events for $\Delta\phi = \pi$. In constrast to the functional form of equation 5.20b, it does not vanish at these positions but it has also local minima there. Like the functional form, it is also maximal for $\Delta\phi = \frac{\pi}{2}$. So the positions of both maxima and minima in the distribution of $\Delta\phi$ are consistent with the positions of maxima and minima of the function for $\rho_{\Delta\phi}$. The offset in figure 5.9(a) means that some events are distributed uniformly and do not follow the function for $\rho_{\Delta\phi}$. But this is no issue since according to the constraint, $\Delta\phi$ is not the only quantity with impact on the event density $\rho_{\Delta\phi}$. That this offset differs for $\Delta\phi = 0$ and $\Delta\phi = \pi$ cannot be explained with the functional form of equation 5.20b, but it can be explained with the functional form of equation 5.22b. For this purpose, the prominent cases $\Delta\phi = 0$, $\Delta\phi = \frac{\pi}{2}$ and $\Delta\phi = \pi$ are considered.

The cases $\Delta\phi = \frac{\pi}{2}$ and $\Delta\phi = \pi$ are easy to handle:

$$\rho_{\Delta\phi} = k \times m_W^2 \times \begin{cases} \frac{1}{\cosh\Delta\eta - 0} = 1 & \text{for } \Delta\phi = \frac{\pi}{2} \\ \frac{0}{\cosh\Delta\eta + 1} = 0 & \text{for } \Delta\phi = \pi \end{cases} \tag{5.23}$$

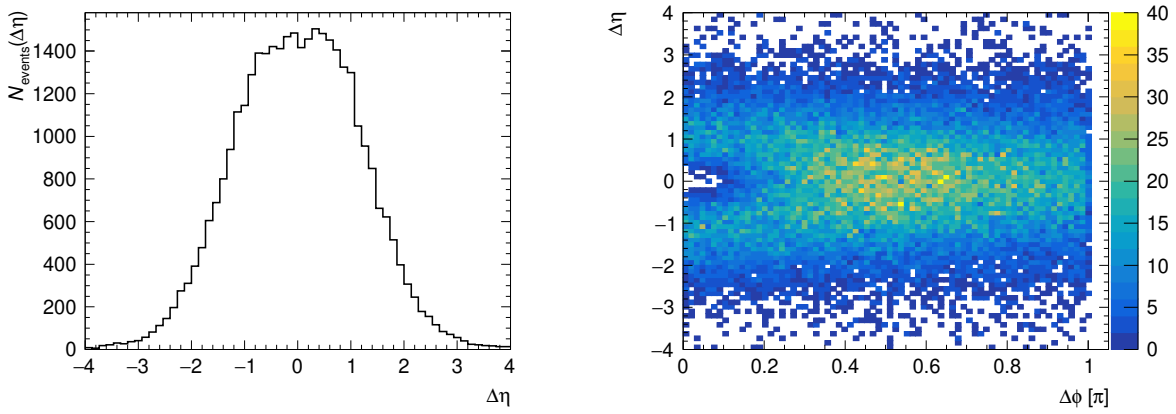
These two cases conform with the analysis of equation 5.20b, i.e. $\rho_{\Delta\phi}$ vanishes for $\Delta\phi = \pi$ and is maximal for $\Delta\phi = \frac{\pi}{2}$. However, the case $\Delta\phi = 0 \Rightarrow \cos\Delta\phi = 1$ needs to be considered more carefully, as equation 5.22b is not well-defined for the case $\Delta\eta = 0 \Rightarrow \cosh\Delta\eta = 1$, because both the nominator and the denominator vanish. So this case is considered by the approximation $\Delta\phi \ll 1$ and $\Delta\eta \ll 1$:

$$\frac{\sin\Delta\phi}{\cosh\Delta\eta - \cos\Delta\phi} \xrightarrow[\Delta\phi\ll 1]{} \frac{\Delta\phi}{\cosh\Delta\eta - 1 + \frac{\Delta\phi^2}{2}} \xrightarrow[\Delta\eta\ll 1]{} \frac{\Delta\phi}{\frac{\Delta\eta^2}{2} + \frac{\Delta\phi^2}{2}} \tag{5.24}$$

The result depends on which variable goes faster to zero.

$$\frac{\Delta\phi}{\frac{\Delta\eta^2}{2} + \frac{\Delta\phi^2}{2}} \longrightarrow \begin{cases} \frac{2\Delta\phi}{\Delta\eta^2} & \xrightarrow[\Delta\phi\to 0]{} 0 \quad \text{for } \Delta\eta \gg \Delta\phi \\ \frac{2}{\Delta\phi} & \xrightarrow[\Delta\phi\to 0]{} \infty \quad \text{for } \Delta\eta \ll \Delta\phi \end{cases} \tag{5.25}$$

The case for $\Delta\phi \to 0$ with $\Delta\eta \gg \Delta\phi$ results in 0, exactly as the case $\Delta\phi = \pi$. But for $\Delta\eta \ll \Delta\phi$ the case $\Delta\phi \to 0$ lets $\rho_{\Delta\phi}$ diverge, in contrast to $\Delta\phi = \pi$. Apparently, the divergence of $\rho_{\Delta\phi}$ for the case $\Delta\phi \to 0$ with $\Delta\eta \ll \Delta\phi$ does not lead to a higher number of events but to a lower. This seems odd at first but can be explained by the fact that the functions for ρ_a, $\rho_{\Delta\phi}$ and $\rho_{\Delta\eta}$ are constrained. An arbitrarily high event density is not possible because $N^{\text{total}}_{\text{events}}$ is fixed. That means that values at which the event density diverges are forbidden by the constraint, and the number of events for these values is suppressed. This explains why the offset of about 800 events at $\Delta\phi = 0$ is below the offset of about 1000 events at $\Delta\phi = \pi$.



(a) (b)

Figure 5.10: Number of events as a function of $\Delta\eta := \eta_l - \eta_\nu$ (a) and as a function of both $\Delta\phi := |\phi_l - \phi_\nu|$ and $\Delta\eta$ (b).

Discussing $\rho_{\Delta\eta}$

The distribution of $\Delta\eta$ is shown in figure 5.10(a). It is symmetric with a broad peak at $\Delta\eta = 0$. The function $\rho_{\Delta\eta}(a, \Delta\eta)$ as in equation 5.20c is linear in $\sinh\Delta\eta$. It vanishes for $\Delta\eta = 0$, diverges for $\Delta\eta \to \infty$ and has no local extremum. Hence, it cannot explain the peak in the distribution of $\Delta\eta$. The function $\rho_{\Delta\eta}(\Delta\eta, \Delta\phi)$ as in equation 5.22c is more complicated. So it is considered for characteristic

cases separately. Firstly, for $\Delta\eta \to \infty$, the term $\cos\Delta\phi$ in the denominator can be neglected and the function becomes linear to $\tanh\Delta\eta$, which does not diverge for $\Delta\eta \to \infty$ but also vanishes for $\Delta\eta = 0$. The same happens also for smaller $\Delta\eta$ in the case that $\cos\Delta\phi = 0 \Leftrightarrow \Delta\phi = \frac{\pi}{2}$. Secondly, for $\Delta\eta \ll 1$, equation 5.22c can be approximated:

$$\rho_{\Delta\eta}(\Delta\eta, \Delta\phi) = k \times m_W^2 \frac{\sinh\Delta\eta}{\cosh\Delta\eta - \cos\Delta\phi} \xrightarrow[\Delta\eta \ll 1]{} k \times m_W^2 \frac{\Delta\eta}{1 + \frac{\Delta\eta^2}{2} - \cos\Delta\phi} \tag{5.26}$$

In case $\Delta\phi \gg \Delta\eta$, this equation goes to 0 for $\Delta\eta \to 0$. For the case $\Delta\phi \ll \Delta\eta$, the approximation $\Delta\phi \ll 1$ is also applied:

$$\frac{\Delta\eta}{1 + \frac{\Delta\eta^2}{2} - \cos\Delta\phi} \xrightarrow[\Delta\phi \ll 1]{} \frac{\Delta\eta}{\frac{\Delta\phi^2}{2} + \frac{\Delta\eta^2}{2}} \tag{5.27}$$

Analogously to the case study in equation 5.25, the result depends on which variable goes faster to zero.

$$\frac{\Delta\eta}{\frac{\Delta\eta^2}{2} + \frac{\Delta\phi^2}{2}} \longrightarrow \begin{cases} \frac{2\Delta\eta}{\Delta\phi^2} & \xrightarrow[\Delta\eta \to 0]{} 0 \quad \text{for } \Delta\phi \gg \Delta\eta \\ \frac{2}{\Delta\eta} & \xrightarrow[\Delta\eta \to 0]{} \infty \quad \text{for } \Delta\phi \ll \Delta\eta \end{cases} \tag{5.28}$$

In contrast to the case $\Delta\phi \gg \Delta\eta$, where $\rho_{\Delta\eta}$ vanishes, $\rho_{\Delta\eta}$ diverges in the case $\Delta\phi \ll \Delta\eta$ for $\Delta\eta \to 0$. But as seen in the section before, this divergence should not argue for a higher event density. In order to illustrate that once more, figure 5.10(b) shows the number of events as a function of $\Delta\phi$ and $\Delta\eta$ in a 2D histogram. As already seen in figures 5.9(a) and 5.10(a), the event density is higher for $\Delta\phi = \frac{\pi}{2}$ and is symmetric around 0. In general, the event density is higher in the region $-1 < \Delta\eta < 1$, in agreement with figure 5.10(a). But there is a region around the point $(\Delta\phi, \Delta\eta) = (0, 0)$ where the event density vanishes. It was not possible to detect this phenomenon in the one-dimensional histograms in figure 5.9(a) and 5.10(a), but it confirms the explanation that due to the constraint, the number of events is suppressed in regions where the functions for the event density diverge. However, also with the functional form for $\rho_{\Delta\eta}$ of equation 5.22c the broad peak at $\Delta\eta = 0$ cannot be explained because it does not have any local extremum. But for the same reason the peak is also not in contrast to the functional form for $\rho_{\Delta\eta}$. However, the shape of the distribution of $\rho_{\Delta\eta}$ can be explained in another way. The symmetry is a direct consequence of the mirror symmetry at $z = 0$. That smaller values for $\Delta\eta$ are favored is due to the fact that the mass of the W boson is much higher than the mass of its decay products. Therefore, the transverse boost of the W boson will be transferred to its decay products to a high fraction.

Discussing ρ_a

In the functional form of equation 5.20a, ρ_a is inversely proportional to a, i.e. it diverges for $a \to 0$. Since it cannot explain the finite value $a = f$, the functional form of equation 5.22a is considered, although that does not depend on a explicitly. However, it has a local maximum for $\cos\Delta\phi = -1 \Rightarrow \Delta\phi = \pi$. For all values of $\Delta\eta$ it is maximal but never diverges. So, the value $\Delta\phi = \pi$ should maximize ρ_a. In order to find a value for a that corresponds to that local maximum, $\cos\Delta\phi = -1$ is plugged into equation 5.21

together with $\Delta\eta = 0$, since the distribution for $\Delta\eta$ is peaked there. This leads to:

$$a = \frac{m_W^2}{2\,(\cosh\Delta\eta - \cos\Delta\phi)} = \frac{m_W^2}{2(1-(-1))} = \frac{m_W^2}{4} = \frac{(80.38\,\mathrm{GeV})^2}{4} = 1\,615.24\,\mathrm{GeV}^2 \tag{5.29}$$

That this value is so close to the fit value $f = (1\,615.98 \pm 5.15)\,\mathrm{GeV}^2$ is astonishing but reassures the consistency of the arguments explained before.

Conclusion

The distribution of $\rho_{\Delta\phi}$ in figure 5.9(a) could be directly explained by analytic functions for $\rho_{\Delta\phi}$, as given in equation 5.20b and 5.22b because this function has local extrema in the dependence of $\Delta\phi$ for which the event density is increased. The function for ρ_a only has local extrema in the dependence on $\Delta\phi$ in the form of equation 5.22a. The value of a for which ρ_a is maximized cannot be given directly but rather indirectly by using the value of $\Delta\phi$ for which ρ_a is maximized and using $\Delta\eta = 0$, which empirically is the average and most probable value for $\Delta\eta$. This value is not the position of a local extremum of a function for $\rho_{\Delta\eta}$ and the distribution of $\rho_{\Delta\eta}$ cannot be explained by analytical functions, but only through qualitative considerations. But this is not an issue, since $\rho_{\Delta\eta}$ has not been used for prediction.

5.3.6 Summary

After some basic considerations in the beginning, this section was dominated by calculations. As a summary, the number of events is proportional to the W-boson mass and the event density to its derivate. The derivative with respect to $\Delta\phi$ has a local maximum at $\frac{\pi}{2}$ and for this value also the event density is maximized. The derivates with respect to a or $\Delta\eta$ do not have local maxima in a or in $\Delta\eta$, but the derivative with respect to a has one in $\Delta\phi$. Using the local maximum for $\Delta\eta = 0$ which was found empirically and assuming the maximum at $\Delta\phi = \pi$, the fit value for a could be reproduced. The correlations between the kinematic variables of the lepton and the neutrino should appear at every leptonic top-quark decay and therefore, may be exploited for other analyses as well.

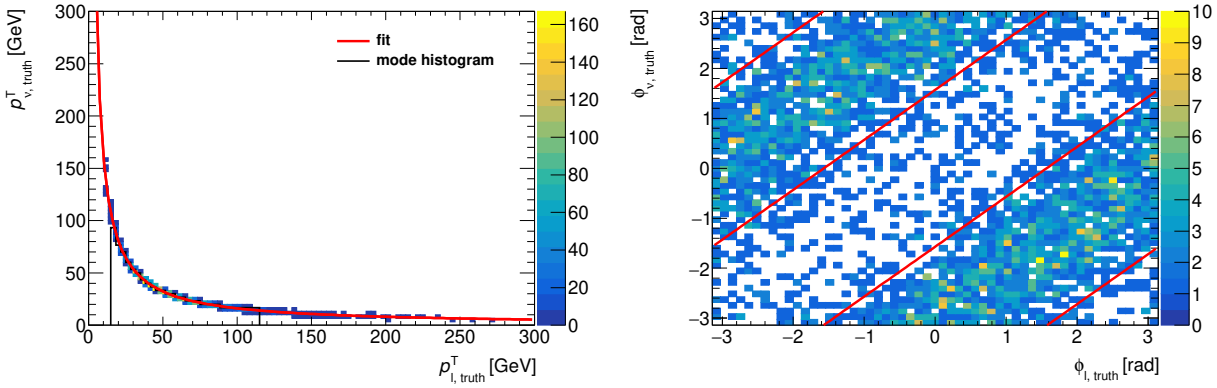
5.4 Predicting p_ν^x and p_ν^y

Section 5.2 reported about transforming the correlations that have been found between p_l^T and p_ν^T and between ϕ_ν and ϕ_l into predictions, formulated in equations 5.5. The discussion of ρ_a at the end of section 5.3, however, indicated that equations 5.5a and 5.5b cannot be fulfilled by one event at the same time, since equation 5.5a implies $\phi_\nu - \phi_l = \pi$ and equation 5.5b implies $\phi_\nu - \phi_l = \pm\frac{\pi}{2}$. In this section, this contrast is investigated on simulated events and, after that, taken as a starting point for further optimizing the prediction procedure.

5.4.1 Selecting events with good prediction of p_ν^T and ϕ_ν

The aforementioned contrast between $\phi_\nu - \phi_l = \pi$ and $\phi_\nu - \phi_l = \pm\frac{\pi}{2}$ is considered from two perspectives. Both perspectives investigate the correlations between $\phi_{\nu,\,\mathrm{truth}}$ and $\phi_{l,\,\mathrm{truth}}$ and between $p^\mathrm{T}_{\nu,\,\mathrm{truth}}$ and $p^\mathrm{T}_{l,\,\mathrm{truth}}$, as already shown in figure 5.6. The first perspective selects only events fulfilling $|p^\mathrm{T}_{\nu,\,\mathrm{pred}} - p^\mathrm{T}_{\nu,\,\mathrm{truth}}| <$

2 GeV and the second perspective selects only events fulfilling $|\phi_{\nu,\,\text{pred}} - \phi_{\nu,\,\text{truth}}| < 0.1$ rad. The two perspectives are treated one after the other.



(a) (b)

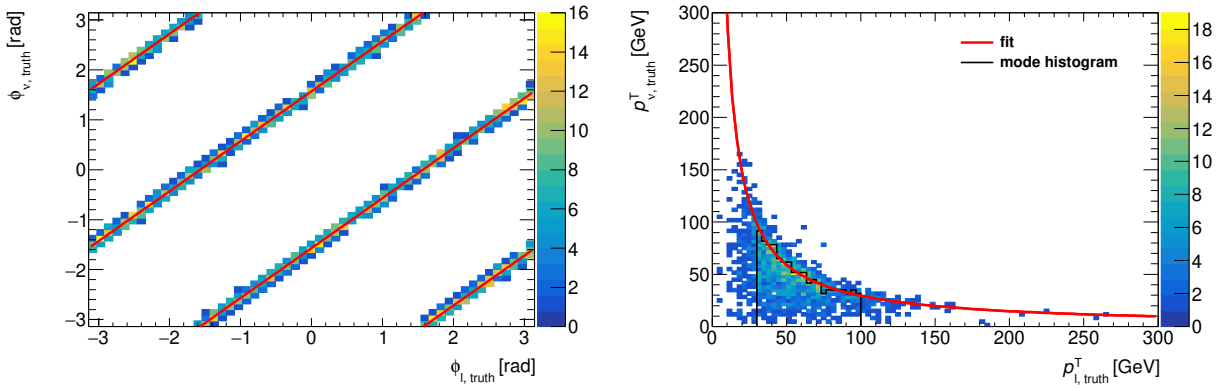
Figure 5.11: Correlation of p^{T}_{ν} with p^{T}_{l} (a) and of ϕ_{ν} with ϕ_{l} (b) (both on *truth-level*), as in figure 5.6, but only for events fulfilling $|p^{\text{T}}_{\nu,\,\text{pred}} - p^{\text{T}}_{\nu,\,\text{truth}}| < 2\,\text{GeV}$. In (a) both the fit curve and the mode histogram are exactly the same as in figure 5.6(a), but after applying the fit and calcualting $p^{\text{T}}_{\nu,\,\text{pred}}$, events which do not fulfill $|p^{\text{T}}_{\nu,\,\text{pred}} - p^{\text{T}}_{\nu,\,\text{truth}}| < 2\,\text{GeV}$ have been excluded. Figure (b) shows the same events as figure (a). The red lines still fulfill $\phi_{\nu} = \phi_{l} \pm \frac{\pi}{2}$, as in figure 5.6(b). But due to the event selection, the red line does not highlight the areas enriched with events but borders them.

Selecting events with good p^{T}_{ν}-prediction

For the first perspective only events fulfilling $|p^{\text{T}}_{\nu,\,\text{pred}} - p^{\text{T}}_{\nu,\,\text{truth}}| < 2\,\text{GeV}$ are selected. For illustration, those events, together with the fit function fitted to all events, are plotted in figure 5.11(a). The diagram works as a cross-check. As expected, only events close to the fit curve can be seen. The same events are selected for the histogram in figure 5.11(b) showing the correlation between ϕ_{ν} and ϕ_{l}. As in figure 5.6, the red lines still display the conditions $\phi_{\nu} = \phi_{l} \pm \frac{\pi}{2}$. In contrast to figure 5.6, the event density is not increased for $\phi_{\nu} = \phi_{l} \pm \frac{\pi}{2}$ but for $\phi_{\nu} = \phi_{l} \pm \pi$. However, the condition $\phi_{\nu} = \phi_{l} \pm \pi$ is not fulfilled to the same precision as $|p^{\text{T}}_{\nu,\,\text{pred}} - p^{\text{T}}_{\nu,\,\text{truth}}| < 2\,\text{GeV}$, which means that the relation between equation 5.5a and $\phi_{\nu} = \phi_{l} \pm \pi$ is not purely deterministic but includes also a random component. This random component is also consistent with the expectations, since also in figure 5.6 and in the discussion in section 5.3 random components have been seen. However, the random component does not harm the general observation that for events fulfilling $|p^{\text{T}}_{\nu,\,\text{pred}} - p^{\text{T}}_{\nu,\,\text{truth}}| < 2\,\text{GeV}$ the best description for the correlation between $\phi_{\nu,\,\text{truth}}$ and $\phi_{l,\,\text{truth}}$ is $\phi_{\nu} = \phi_{l} \pm \pi$, while for all events it is $\phi_{\nu} = \phi_{l} \pm \frac{\pi}{2}$. This contrast has already been implied in the discussion of ρ_{a} at the end of section 5.3 and is clearly demonstrated by figure 5.11(b).

Selecting events with good ϕ_{ν}-prediction

For the second perspective, only events fulfilling $|\phi_{\nu,\,\text{pred}} - \phi_{\nu,\,\text{truth}}| < 0.1$ rad are selected. For a cross-check, they are plotted in figure 5.12(a) together with the four red lines displaying the condition $\phi_{\nu} = \phi_{l} \pm \frac{\pi}{2}$. As expected, all events in the diagram are close to those red lines. The same events are



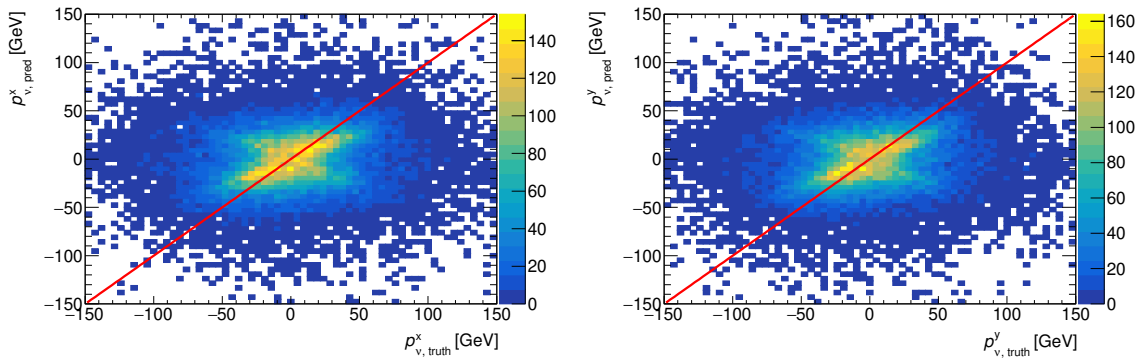
(a) (b)

Figure 5.12: Correlation of p_ν^{T} with p_l^{T} (b) and of ϕ_ν with ϕ_l (a) (both on *truth-level*), as in figure 5.6, but only for events fulfilling $|\phi_{\nu,\,\mathrm{pred}} - \phi_{\nu,\,\mathrm{truth}}| < 0.1\,\mathrm{rad}$. For each event respectively, $\phi_{\nu,\,\mathrm{pred}}$ is determined using ϕ_l and ϕ_b on *truth-level* as in equation 5.5b. In order to test the condition, $\phi_{\nu,\,\mathrm{truth}}$ is taken from the same event. Figure (a) confirms that the selected events are close to the red lines fulfilling the condition $\phi_\nu = \phi_l \pm \frac{\pi}{2}$, which has been used for prediction. Figure (b) shows the same events as figure (a). The mode histogram and fit have been calculated on the basis of this very set of events. The fit results in $f = (2\,955.14 \pm 90.47)\,\mathrm{GeV}^2$.

selected for the histogram in figure 5.12(b) showing the correlation between p_ν^{T} and p_l^{T}. In contrast to figure 5.6(a), where the events are spread over almost the whole histogram area, almost all events in figure 5.12(b) are confined to a smaller area limited by $p^{\mathrm{T}}_{\nu,\,\mathrm{truth}} = 150\,\mathrm{GeV}$ and $p^{\mathrm{T}}_{l,\,\mathrm{truth}} = 150\,\mathrm{GeV}$. However, also the event distribution in figure 5.12(b) suggests using the form $p_\nu^{\mathrm{T}} = \frac{f}{p_l^{\mathrm{T}}}$ for a fit function. The fit range is changed from $[15\,\mathrm{GeV}, 115\,\mathrm{GeV}]$ to $[30\,\mathrm{GeV}, 100\,\mathrm{GeV}]$ in order to meet the diminished number of events. In comparison to the result of $f_{\mathrm{all}} = (1\,615.98 \pm 5.15)\,\mathrm{GeV}^2$ from the fit using all events, the fit using only events fulfilling $|\phi_{\nu,\,\mathrm{pred}} - \phi_{\nu,\,\mathrm{truth}}| < 0.1\,\mathrm{rad}$ results in $f_{\mathrm{selected}} = (2\,955.14 \pm 90.47)\,\mathrm{GeV}^2$. This is roughly twice the value of f_{all}. In order to illustrate this difference, figure C.1 in appendix C shows the fit function with the result f_{selected} into a diagram showing the correlation between p_ν^{T} and p_l^{T} using all events. The statistical error of f_{selected} is much larger than that of f_{all} due to the smaller number of events. All in all, the fact that the fit parameters differ roughly by a factor of 2, confirms that equations 5.5a and 5.5b are in conflict with each other, as implied by the discussion of ρ_a at the end of section 5.3.

5.4.2 Introducing different approaches for predicting p_ν^x and p_ν^y

The investigation in the previous section showed that the number of events fulfilling both equation 5.5a and 5.5b is very small. This may have the consequence that the prediction of p_ν^x and p_ν^y, which combines both predictions of equation 5.5 in equation 5.6, may be satisfying also only for a small number of events. This would be disadvantageous for the mass reconstruction of the $\tau\tau$-state. A possible way to compensate for that is to combine those predictions of p_ν^{T} and ϕ_ν which have been found to be compatible with each other in the previous section. In total, there are three approaches to be compared regarding the prediction of p_ν^y and p_ν^y. The first approach uses $\phi_\nu = \phi_l \pm \frac{\pi}{2}$ and $f_{\mathrm{all}} = (1\,615.98 \pm 5.15)\,\mathrm{GeV}^2$ obtained from the fit with all events. It is called *conventional*. The approach using $\phi_\nu = \phi_l \pm \pi$, which fits best for those



(a) (b)

Figure 5.13: Correlation of $p^x_{\nu,\text{ pred}}$ with $p^x_{\nu,\text{ truth}}$ (a) and of $p^y_{\nu,\text{ pred}}$ with $p^y_{\nu,\text{ truth}}$ (b) for the *conventional* approach ($\phi_\nu = \phi_l \pm \frac{\pi}{2}$, p^{T} fit parameter $f = (1\,615.98 \pm 5.15)\,\text{GeV}^2$). The red line indicates the identity.

events with good p^{T}-prediction, is called *back-to-back*. At last, the approach using $\phi_\nu = \phi_l \pm \frac{\pi}{2}$ and $f_{\text{selected}} = (2\,955.14 \pm 90.47)\,\text{GeV}^2$ obtained only from events which fit best to the ϕ_ν-prediction using $\phi_\nu = \phi_l \pm \frac{\pi}{2}$, is called *new p^T-fit*.

Evaluation of the different approaches for predicting p^x_ν and p^y_ν

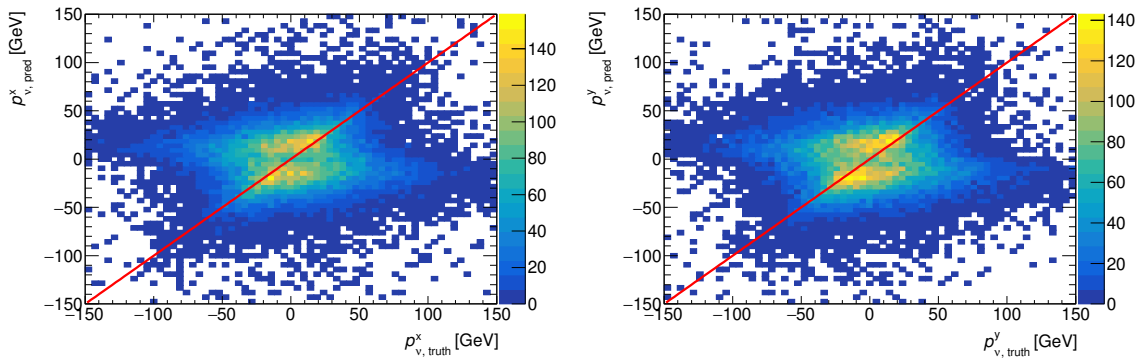
In order to evaluate the prediction, $\phi_{\nu,\text{ pred}}$ and $p^{\text{T}}_{\nu,\text{ pred}}$ are plotted against $\phi_{\nu,\text{ truth}}$ and $p^{\text{T}}_{\nu,\text{ truth}}$ respectively in a color plot for all approaches. They are shown in figure 5.13 (*conventional*), 5.14 (*back-to-back*) and 5.15 (*new p^T-fit*). In the limit of a perfect prediction, all events should lie on a diagonal line indicating the identity. For evaluation, this line is drawn in red on top of the color plot in the same plotting area. For each approach p^x_ν is shown on the left hand side and p^y_ν on the right hand side.

General properties of the prediction of p^x_ν and p^y_ν Some general properties of the prediction can be seen in all plots. There is a cluster of events (yellow and orange) around the origin (±50 GeV) with low p^x/p^y and a cloud of events (light blue and dark blue) (±150 GeV). The larger the distance of the events from the red line, the worse the prediction. In all plots, there is a not negligible amount of events that are quite far away from the identity and therefore badly predicted. As the cluster of events is around the origin, there is no large systematic bias in the prediction, meaning that the prediction for the cluster events is not too bad in any case.

Evaluation of the prediction of p^x_ν and p^y_ν in detail For all three approaches, the differences between these approaches for p^x_ν or p^y_ν are larger than the difference between the prediction of p^x_ν and p^y_ν for each approach. For this reason, in the following subsections the plots are discussed in more detail for each approach respectively but without separating in p^x_ν and p^y_ν.

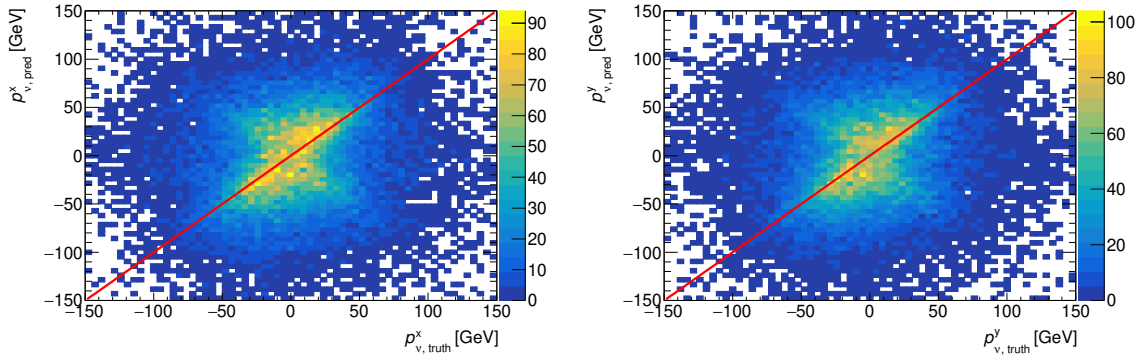
5.4.3 Conventional approach

For the *conventional* approach, the shape of the cluster fits quite well to the identity, though there are some events which deviate from the identity line by a wrong sign in $p^x_{\nu,\text{ pred}}$, giving the cluster some kind



(a) (b)

Figure 5.14: Correlation of $p^x_{\nu,\text{ pred}}$ with $p^x_{\nu,\text{ truth}}$ (a) and of $p^y_{\nu,\text{ pred}}$ with $p^y_{\nu,\text{ truth}}$ (b) for the *back-to-back* approach ($\phi_\nu = \phi_l \pm \pi$, p^{T} fit parameter $f = (1\,615.98 \pm 5.15)\,\text{GeV}^2$). The red line indicates the identity.



(a) (b)

Figure 5.15: Correlation of $p^x_{\nu,\text{ pred}}$ with $p^x_{\nu,\text{ truth}}$ (a) and of $p^y_{\nu,\text{ pred}}$ with $p^y_{\nu,\text{ truth}}$ (b) for the *new* p^T*-fit* approach ($\phi_\nu = \phi_l \pm \frac{\pi}{2}$, p^{T} fit parameter $f = (2\,955.14 \pm 90.47)\,\text{GeV}^2$). The red line indicates the identity.

of an "x-shape". These deviating events, lying on the "anti-identity", are due to the ambiguity in the ϕ-prediction. The ambiguity could be resolved by taking the neutrino-solution antiparallel to the b-quark, but as can be seen in figure 5.5(b), the number of events with the neutrino going parallel to the b-quark is suppressed but not vanishing. This explains why the "anti-identity"-line in figure 5.13 is weaker but not vanishing. The distribution of the events in the cloud seems to be almost random, being similar to a circle around the origin. However, the shape is not a circle but an ellipse, which is more flat, meaning that the range in $p^x_{\nu,\text{ truth}}$ is wider than that for $p^x_{\nu,\text{ pred}}$. The shape can be explained by examining the distributions of $p^{\text{T}}_{\nu,\text{ truth}}$ plotted on the y-axis and $p^{\text{T}}_{l,\text{ truth}}$ plotted on the x-axis in figure 5.6(a). Because of the inverse proportion of the fit function, large values in p^{T}_l are mapped to small values in p^{T}_ν. Large values of p^{T}_ν would be mapped by small values for p^{T}_l. But events with small p^{T}_l do not pass the selection cuts. This is why the distribution of $p^{\text{T}}_{\nu,\text{ pred}}$ is missing those events and so is the distribution of $p^x_{\nu,\text{ pred}}$. The p^y_ν-plot looks very similar to the p^x_ν-plot in all of these aspects.

5.4.4 Back-to-back approach

For the *back-to-back* approach, the shape of both the cluster and the cloud of events is quite distorted in comparison to the *conventional* approach. The x-shape cluster is torn apart into two parallel lines. The new shape can be described as a z-shape instead of an x-shape but without the diagonal line of the z, which would lie on the identity. The plot illustrates what happens if the *back-to-back* approach is applied to those events, for which it is not true. In the *back-to-back* approach, the angle of $p^x_{\nu,\,\mathrm{pred}}$ is always the opposite of p^x_l. The following consideration concerns events with $p^{\mathrm{T}}_l \approx 30\,\mathrm{GeV}$ and $\phi_l \approx 0$, which leads to $p^x_l \approx p^{\mathrm{T}}_l$. For those events, the fit function predicts $p^{\mathrm{T}}_l \approx p^{\mathrm{T}}_\nu$, as can be seen in figure 5.6(a). If for those events ϕ_ν is predicted by $\phi_\nu = \phi_l + \pi$, these events are mapped to $p^x_\nu = -p^x_l$ always, which explains why for events with $p^x_l > 0$ there are quite many events on the left hand side of the identity giving the upper horizontal line and for events with $p^x_l < 0$ there are quite many events on the right hand side from the identity, giving the lower horizontal line. With a prediction for ϕ_ν which is correct for more events, the chance is higher for them to be mapped along the identity line. This trend can also be seen in the cloud of events, which is not an ellipse anymore but rather a rhomboid. Also for the *back-to-back* approach the p^y_ν-plot looks very similar to the p^x_ν-plot in all of these aspects.

5.4.5 New p^{T}-fit approach

For this approach, the cluster has again the x-shape as for the *conventional* approach. Here, the x-shape can be seen more clearly for p^x_ν than for p^y_ν and in general, the cluster for p^y_ν seems to contain more events than the cluster for p^y_ν. Also, the angle between the two lines of the x seems to be nearly perpendicular here, whereas the x is flatter for the *conventional* approach. In addition, the cluster appears to be larger here than for the *conventional* approach. But the cloud also looks larger here (less white space in the plotting area), meaning that for more events the distance to the identity is higher. As for the *conventional* approach, the "anti-identity" is weaker here and most events of the cluster build a shape with a positive slope. However, the slope is a bit larger than 1, while for the *conventional* approach this was a bit smaller than 1. This can be explained with the other p^{T}-prediction. Since the fit-parameter is higher here, the prediction $p^{\mathrm{T}}_{\nu,\,\mathrm{pred}}$ is also higher than for the *conventional* approach. Apparently, for most events it is too high. That $p^{\mathrm{T}}_{\nu,\,\mathrm{pred}}$ is too high for the *new p^T-fit* approach also explains why the cloud is larger here. Higher $p^{\mathrm{T}}_{\nu,\,\mathrm{pred}}$ leads to higher $p^x_{\nu,\,\mathrm{pred}}$ and $p^y_{\nu,\,\mathrm{pred}}$. It counteracts the flattening of the $p^x_{\nu,\,\mathrm{pred}}$-distribution due to the p^{T}-selection of the leptons, already discussed in the paragraph about the *conventional* approach.

5.5 Testing mass reconstruction with predicted p^x_ν and p^y_ν

The *back-to-back* approach for predicting p^x_ν and p^y_ν clearly gives less good results than the other two. These two, i.e. the *conventional* and the *new p^T-fit* approach, are both worth testing for the mass reconstruction of the $\tau\tau$-state. For the mass reconstruction of the $\tau\tau$-state the predicted values for p^x_ν and p^y_ν are taken as $\overrightarrow{\not{E}}_{\mathrm{T}}$ (top) and, as indicated in equation 4.1, are subtracted from $\overrightarrow{\not{E}}_{\mathrm{T}}$ (event) in order to get $\overrightarrow{\not{E}}_{\mathrm{T}}\,(\tau\tau)$, the missing transverse energy assigned to the $\tau\tau$-state. $\overrightarrow{\not{E}}_{\mathrm{T}}\,(\tau\tau)$ is plugged into a mass reconstruction method mentioned in section 4.4.3. The visible mass is not affected by the prediction of p^x_ν and p^y_ν and therefore not investigated here. Since the implementation of the MMC is not available

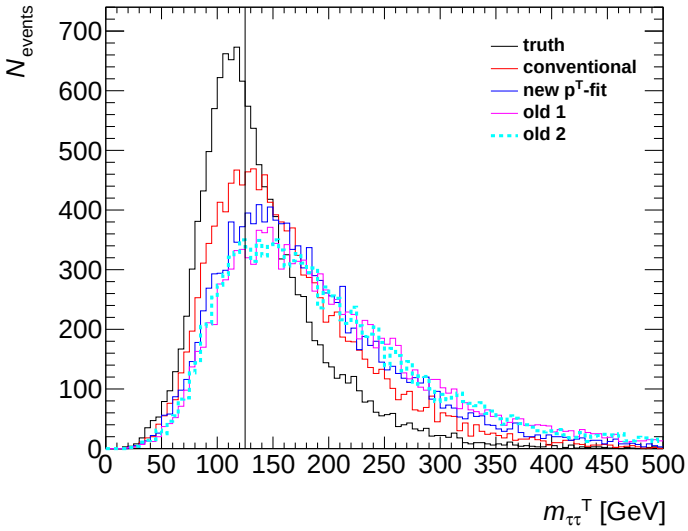
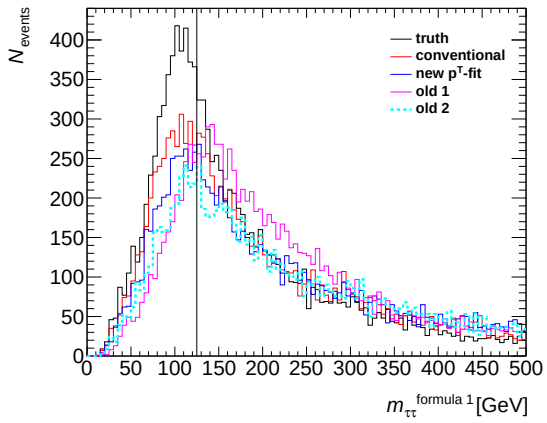


Figure 5.16: Number of tHq events (`SingleTop Ntuples` version 32) as a function of the reconstructed $\tau\tau$-state mass using the transverse mass method and different inputs for $\overrightarrow{\not{E}}_{\mathrm{T}}$ $(\tau\tau)$. The inputs for $\overrightarrow{\not{E}}_{\mathrm{T}}$ $(\tau\tau)$ are calculated via $\overrightarrow{\not{E}}_{\mathrm{T}}$ $(\tau\tau) = \overrightarrow{\not{E}}_{\mathrm{T}}$ (event) $- \overrightarrow{\not{E}}_{\mathrm{T}}$ (top), where $\overrightarrow{\not{E}}_{\mathrm{T}}$ (top) is filled with p_{ν}^{x} and p_{ν}^{y}. The legend in each histogram indicates what values for p_{ν}^{x} and p_{ν}^{y} are taken. "truth" means p_{ν}^{x} and p_{ν}^{y} are taken on *truth-level*. All other approaches use information on *reco-level* exclusively in order to predict p_{ν}^{x} and p_{ν}^{y}. "conventional" uses the *conventional* approach and "new p^{T}-fit" the *new p^{T}-fit* approach (cf. section 5.4). "old 1" and "old 2" stand for the two solutions of the *old* approach (cf. section 5.1).

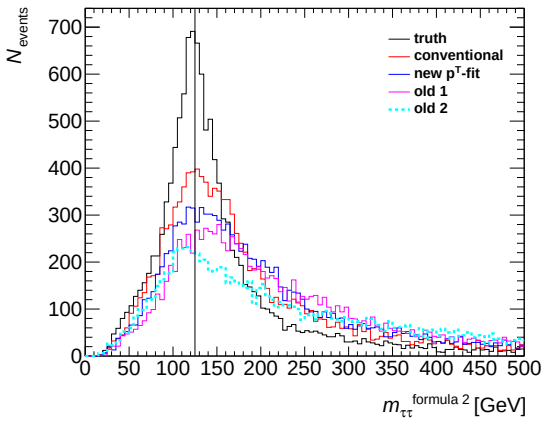
now, only the transverse mass method and the two formulas for the collinear approximation are used for mass reconstruction. For each mass reconstruction method, the different approaches for predicting p_{ν}^{x} and p_{ν}^{y} are compared to each other. This is done by showing the respective mass distributions, which are different because of the different input for $\overrightarrow{\not{E}}_{\mathrm{T}}$ $(\tau\tau)$, in one plot. In the next paragraph, this will be demonstrated using the example of the transverse mass method, shown in figure 5.16. For each mass reconstruction method, a new plot is generated. In order to compare the different mass reconstruction methods to each other, the respective plots are all arranged next to each other in one figure, 5.17.

Comparing the different predicting approaches

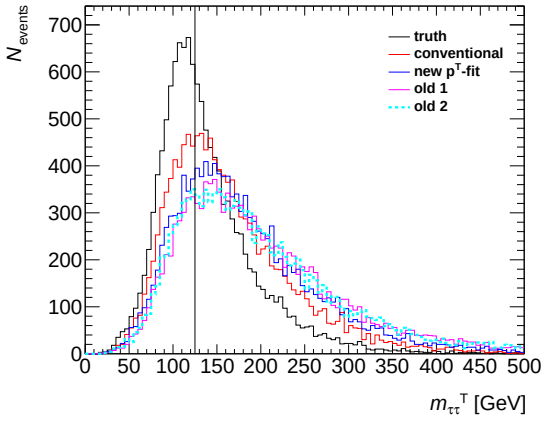
All mass distributions shown in figure 5.16 are generated using the transverse mass method but use different input for $\overrightarrow{\not{E}}_{\mathrm{T}}$ (Higgs). The old approach, as explained before in section 5.1, delivers either no real solution or two equally valued solutions. Events with two solutions are both plotted in figure 5.16, called *old 1* and *old 2*. Events without real solutions are discarded. In the first step, also the other distributions contain only events with real solutions for the *old* approach. This way, the different approaches for predicting p_{ν}^{x} and p_{ν}^{y} can be compared easily. Those different approaches are the *conventional* and *new p^{T}-fit* approach discussed above in section 5.4 as well as the old approach solving the quartic equation. Finally, figure 5.16 also contains the solutions of the transverse mass method, if $\overrightarrow{\not{E}}_{\mathrm{T}}$ $(\tau\tau)$, calculated with $p_{\nu,\,\mathrm{truth}}^{x}$ and $p_{\nu,\,\mathrm{truth}}^{y}$ (instead of $p_{\nu,\,\mathrm{pred}}^{x}$ and $p_{\nu,\,\mathrm{pred}}^{y}$) is taken as input (called *truth*). This allows getting an upper boundary for the best mass reconstruction using the transverse mass method in the ideal case of



(a)



(b)



(c)

Figure 5.17: Number of tHq events (`SingleTop Ntuples` version 32) as a function of the reconstructed $\tau\tau$-state mass using different mass reconstruction methods and different inputs for $\overrightarrow{\not{E}}_T$ $(\tau\tau)$. Those methods are the collinear approximation with formula 1 (equation A.5, figure (a)) and with formula 2 (equation A.10, figure (b)) as well as the transverse mass method (equation 4.2, figure (c)). The inputs for $\overrightarrow{\not{E}}_T$ $(\tau\tau)$ are calculated via $\overrightarrow{\not{E}}_T$ $(\tau\tau) = \overrightarrow{\not{E}}_T$ (event) $- \overrightarrow{\not{E}}_T$ (top), where $\overrightarrow{\not{E}}_T$ (top) is filled with p_ν^x and p_ν^y. The legend in each histogram indicates what values for p_ν^x and p_ν^y are taken. "truth" means p_ν^x and p_ν^y are taken on *truth-level*. All other approaches use information on *reco-level* exclusively in order to predict p_ν^x and p_ν^y. "conventional" uses the *conventional* approach and "new p^T-fit" the *new p^T-fit* approach (cf. section 5.4). "old 1" and "old 2" stand for the two solutions of the *old* approach (cf. section 5.1).

perfectly predicted p_ν^x and p_ν^y. Thus, it can be seen how much the mass reconstruction is theoretically able to be improved by improving the prediction of p_ν^x and p_ν^y.

5.5.1 The transverse mass method

The distributions of the reconstructed mass using the transverse mass method can be examined in figure 5.16 as well as in figure 5.17(c). The distribution using the *truth* input has by far the largest peak

and the smallest tail, which is as expected. In spite of using the *truth* input, the peak position is below the Higgs mass of 125 GeV, but this is also expected, since for the transverse mass method the missing energy along the beam axis cannot be used. The second-highest peak is that of the mass distribution using the *conventional* approach. The two solutions of the *old* approach are very close to each other and give the weakest results. The distribution of the *new p^T-fit* approach is between the distributions of the *conventional* approach and of the *old* approach. It lies approximately in the middle of them, but a bit closer to the distributions of the *old* approach.

5.5.2 Formula 1 of the collinear approximation

Figure 5.17(a) shows the mass reconstruction of the $\tau\tau$-state using formula 1 of the collinear approximation. As for the transverse mass method, the distribution using the *truth* input has by far the largest peak and its position is also below the Higgs mass, even further below. In contrast to the transverse mass method, its width, including its tail, is quite close to the width of the other distributions. In further contrast, all other distributions are quite close to each other, but with a second glance they can be separated from each other. The second-highest peak is that of the mass distribution using the *conventional* approach, outperforming the *new p^T-fit* approach. The two solutions of the old approach give the weakest results; solution 1 has a quite high peak, but at a too high mass, and solution 2 is hardly weaker than the *new p^T-fit* approach.

5.5.3 Formula 2 of the collinear approximation

The reconstrued mass distributions of the $\tau\tau$-state using formula 2 of the collinear approximation are shown in figure 5.17(b). Also here, the mass distribution using the *truth*-input has by far the largest peak. As for the transverse mass method, also its width is clearly smaller and the tail falls off faster. But in contrast to the other methods, the mass distribution using the *truth*-input is at the correct position only for formula 2. As for the other methods, the *conventional* approach shows the highest peak of all predicting approaches, the *new p^T-fit* approach shows the second-highest and *old 1* and *old 2* the lowest. The differences in the distributions between the different predicting approaches are the largest for this method.

5.5.4 Conclusion and outlook

Predicting approaches

Both new approaches, *conventional* and *new p^T-fit*, are better than both solutions of the old approach. The search for new correlations was worthwhile and the exploitation of them is not only easier, since there is no quartic equation to solve and no cases without real solutions, but it also leads to a clear improvement in mass reconstruction, if the same events are used as for the *old* approach. The *conventional* approach is clearly the best. Attempts to optimize this with the *new p^T-fit* or *back-to-back* approach have been without success. Comparing the distributions in more detail reveals that the *new p^T-fit* approach lies somewhere in the middle between the *conventional* approach and the *old* approach. It shows that optimizations in the prediction make a difference in the end. The *truth* distribution shows that there is space for even more optimization, but on the other hand, at least for the used mass reconstruction method, this optimization is also limited. Ways to further improve the prediction, presumably have to be fundamentally different from the *new p^T-fit* and *back-to-back* approach. Maybe there are correlations between kinematic variables of the *b*-jet and the light lepton from the top quark which can be exploited in order to improve the prediction.

For example, it is thinkable to divide the sample into clusters depending on kinematic variables of the b-jet and conduct the p^{T}-fit for each cluster separately. On the other hand, from what has been exposed about the reasons for the correlations, there is no indication that this clustering would help.

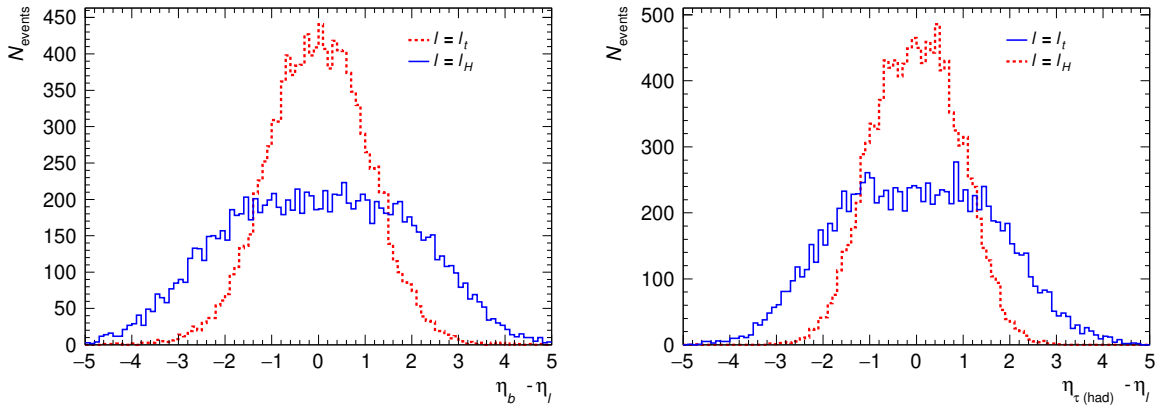
Mass reconstruction methods

Formula 1 of the collinear approximation is by far the weakest, as can be seen by the scale of the y-axis. Formula 2 is the best, the only distribution with the correct peak position and the distribution with the smallest peak width. The transverse mass method is in between, with the same scale on the y-axis as formula 2, but with a peak width that is a bit larger and a peak position that is a bit too low. The different reconstruction methods also differ in the extent to which they are sensitive to different predicting approaches. From the three reconstruction methods that have been investigated here, it can be concluded that the better the peak is for the *truth* input, the better it is also for the other inputs and the more sensitive it is to the different approaches. Short: For a poor reconstruction method the predicting approach does not matter, but for a good one it does. On the basis of this insight, it is expected that the missing mass calculator further improves the mass reconstruction and also shows differences between the predicting approaches more clearly. This may motivate further attempts to improve the prediction of p_ν^x and p_ν^y.

CHAPTER 6

Studies on assignment

The selection cuts of the $2\,e/\mu + 1\tau_{\text{had}}$-channel select events with two light leptons in the final state, called l_1 and l_2. In the reconstruction of tHq there are two possible origins for each light lepton, the top quark and the $\tau\tau$-state of the Higgs boson decay. For the mass reconstruction of the $\tau\tau$-state, the correct assignment of l_1 and l_2 to their respective origin is crucial. After assignment, the light lepton originating from the top quark is labeled as l_t and the light lepton originating from the $\tau\tau$-state of the Higgs boson decay is labeled as l_H. The strategy for assignment explained in section 4.4.2 works only for *OS* events. This chapter reports on developing strategies for assignment which also work for *SS* events.



(a) (b)

Figure 6.1: Number of events as a function of differences in the kinematic variable η. Figure 6.1(a) shows $\eta_b - \eta_l$, where η_b corresponds to the b-jet and figure 6.1(b) shows $\eta_{\tau_{\text{had}}} - \eta_l$, where $\eta_{\tau_{\text{had}}}$ corresponds to the hadronic tau. In both diagrams, η_l corresponds to the light lepton originating either from the top quark (dashed line in diagram 6.1(a) and solid line in figure 6.1(b)) or to the Higgs boson (solid line in diagram 6.1(a) and dashed line in figure 6.1(b)). Both histograms in each diagram only contain events, where both light leptons have the same sign of electric charge (*SS* events).

6.1 Exploration of kinematic variables on *truth-level*

A strategy working also for *SS* events should be independent of the charge of the light leptons but may rely on kinematic variables. The main idea behind this approach is that l_H shares some kinematic properties with the hadronic tau because it has the same mother particle, and that l_t shares some with the b-jet for the same reason. The following studies are executed exclusively on *truth-level* for *SS* events.

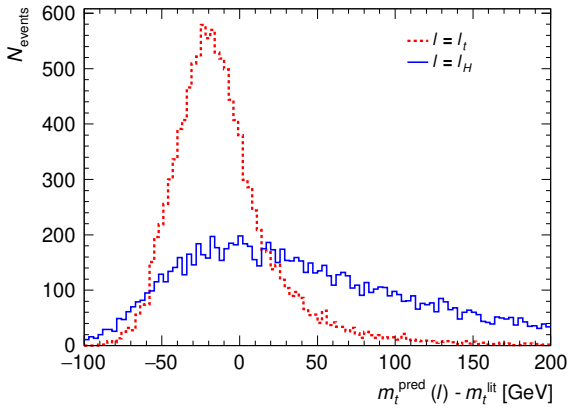


Figure 6.2: Number of events as a function of $m_t^{\text{pred}}(l) - m_t^{\text{lit}}$, where $m_t^{\text{lit}} = 173\,\text{GeV}$ is the literature value of the mass of the top quark (taken from table 2.3). $m_t^{\text{pred}}(l)$ is calculated using formula 6.1. It is a predictor for the top quark mass and depends on the choice of the light lepton. The dashed line is used for the light lepton from the top quark (correct choice) and the solid line for the light lepton from the Higgs boson (wrong choice). Both histograms only contain events, where both light leptons have the same sign of electric charge (*SS* events).

6.1.1 Explore differences in η as example

One example of these kinematic variables is η. Using the names η_b for the b-jet, $\eta_{\tau_{\text{had}}}$ for the hadronic tau, η_{l_t} for the light lepton from the top quark and η_{l_H} for the light lepton from the Higgs boson, it is expected that η_b should be closer to η_{l_t} than to η_{l_H}, and analogously it is expected that $\eta_{\tau_{\text{had}}}$ should be closer to η_{l_H} than to η_{l_t}. In order to check this expectation, for every event, differences in η are calculated and filled into histograms shown in figure 6.1. Figure 6.1(a) shows the distribution of $\eta_b - \eta_{l_H}$ in comparison to $\eta_b - \eta_{l_t}$ and figure 6.1(b) shows $\eta_{\tau_{\text{had}}} - \eta_{l_H}$ in comparison to $\eta_{\tau_{\text{had}}} - \eta_{l_t}$. In both cases the dotted histogram line is used for the lepton for which a narrower distribution is expected, i.e. in figure 6.1(a) the dotted line shows $\eta_b - \eta_{l_t}$ and the solid line shows $\eta_b - \eta_{l_H}$, while in figure 6.1(b) the dotted line shows $\eta_{\tau_{\text{had}}} - \eta_{l_H}$ and the solid line shows $\eta_{\tau_{\text{had}}} - \eta_{l_t}$. Both in figure 6.1(a) and in figure 6.1(b) the distribution of the dotted line is clearly narrower, which confirms the expectation. Additionally, the narrower distributions show that only very few events fulfill $|\eta_b - \eta_{l_t}| > 2.5$ and $|\eta_{\tau_{\text{had}}} - \eta_{l_H}| > 2.5$, while the fraction of events fulfilling $|\eta_b - \eta_{l_H}| > 2.5$ and $|\eta_{\tau_{\text{had}}} - \eta_{l_t}| > 2.5$ is much larger, which can be inferred from the wider distributions with the solid histogram line. This is interesting since the selection cuts require $|\eta| < 2.5$ for all leptons. This means that particles 1 and 2 with $|\eta_1 - \eta_2| > 2.5$ fly in opposite directions relative to the beam axis. Due to the boost along the beam axis, it is expected that $|\eta_1 - \eta_2| \lesssim 2.5$ if particle 1 and particle 2 have the same mother particle. However, in order to use the differences in η for assignment, they have to be compared event by event. This is reported in

section 6.1.3.

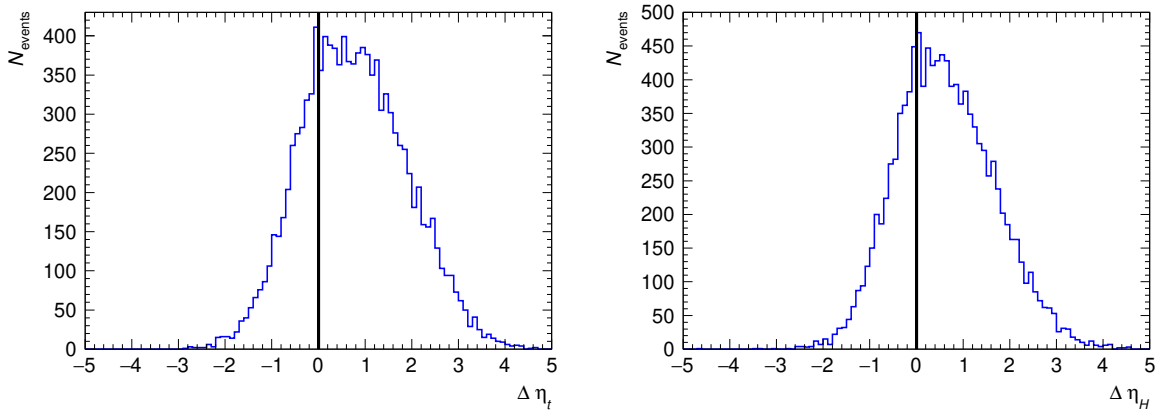
6.1.2 Explore reconstructed masses

Another idea tries to exploit the predictions for the neutrino from the top quark, reported in chapter 5 and the mass constraint for the top quark, which has not been used for predicting p_ν^x and p_ν^y. The idea is to reconstruct m_t by adding up the according four-momenta and to do this calculation for both light leptons. p_ν^{T} and ϕ_ν are predicted as depicted in equations 5.5a and 5.5b. Figure 5.4(a) shows that as a first-order approximation, it can be assumed that $\eta_\nu = \eta_l$. So the formula for m_t^{pred} depends on the lepton l explicitly by P_l and implicitly by P_ν^{pred}:

$$m_t^{\mathrm{pred}}(l) := \sqrt{\left(P_b + P_l + P_\nu^{\mathrm{pred}}(l)\right)^2} \tag{6.1}$$

It is expected that the result of m_t^{pred} using l_t should be closer to the literature value of m_t than the result using l_H. In order to check this expectation, $m_t^{\mathrm{pred}}(l)$ is calculated for every event for both leptons. The deviation from the literature value of $m_t^{\mathrm{lit}} = 173\,\mathrm{GeV}$ is shown in figure 6.2. In contrast to figure 6.1 the distributions are not symmetric around 0 and have a tail towards positive values. The distribution of $m_t^{\mathrm{pred}}(l_t)$ has a clear peak at $\approx -20\,\mathrm{GeV}$. A possible reason, why the peak is significantly below 0, could be that the assumption $\eta_\nu = \eta_l$ leads to minimal angular momentum between light lepton and neutrino.

The peak position of $m_t^{\mathrm{pred}}(l_H)$ is difficult to estimate because the distribution is very broad and shows statistical fluctuations. Thus, the deviation of the peak position is either not statistically significant or small enough to be neglected here. While the distributions in figure 6.2 do not differ significantly according to the peak position, their difference in shape is significant. The peak of $m_t^{\mathrm{pred}}(l_t)$ is much narrower and its tail much flatter. In other words, calculating $m_t^{\mathrm{pred}}(l)$ with incorrect assignment leads to many more events with much too high masses. While the deviation of the peak position is not as ideal for assignment as expected, the different shape of the distribution is all the more promising.



(a) (b)

Figure 6.3: Number of events as a function of $\Delta\eta_H$ (a) and of $\Delta\eta_t$ (b). These variables are defined in equations 6.3 and 6.4 and are constructed such, that they are expected to give positive values to as many events as possible. Both histograms only contain events, where both light leptons have the same sign of electric charge (*SS* events).

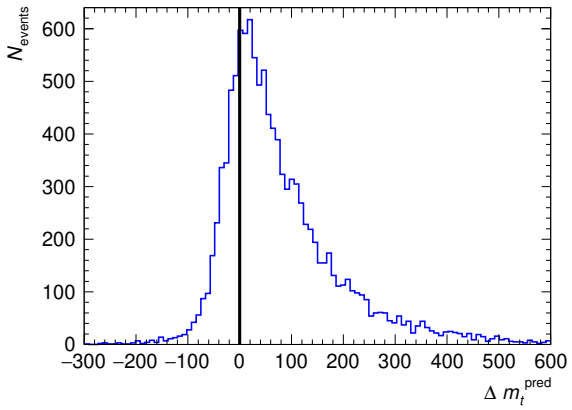


Figure 6.4: Number of events as a function of Δm_t^{pred}, as it is defined in equation 6.2. This variable is constructed such, that it is expected to give positive values to as many events as possible. The histogram only contains events, where both light leptons have the same sign of electric charge (*SS* events).

6.1.3 Construct new variables

The significant deviation in the shape of the distributions in figure 6.2 leads to the idea to calculate the masses with both leptons and take that lepton to be correct which leads to a smaller mass. Therefore the new variable Δm_t^{pred} is defined as:

$$\Delta m_t^{\text{pred}} := m_t^{\text{pred}}(l_1) - m_t^{\text{pred}}(l_2) \tag{6.2}$$

In order to compare the differences in η event by event, the following variables are defined:

$$\Delta\eta_H := |\eta_{\tau_{\text{had}}} - \eta_{l_t}| - |\eta_{\tau_{\text{had}}} - \eta_{l_H}| \tag{6.3}$$

$$\Delta\eta_t := |\eta_b - \eta_{l_H}| - |\eta_b - \eta_{l_t}| \tag{6.4}$$

These variables are constructed such, that, based on the results above, they are expected to give positive values for most events. Figures 6.4 and 6.3 allow for checking this expectation. A vertical line at 0 is drawn in order to discriminate events with positive values from the other. For all three kinematic variables, the distribution is asymmetric in favor of events with a positive value, i.e. the expectations are confirmed. This property can be used for assignment in the way that the light leptons are assigned such that the result for the kinematic variables above is positive. So for each variable, an explicit rule for assignment can be formulated:

For $\Delta\eta_H$: If $|\eta_{\tau_{\text{had}}} - \eta_{l_1}| - |\eta_{\tau_{\text{had}}} - \eta_{l_2}| \geq 0$, then: $l_1 = l_t$ and $l_2 = l_H$
else: $l_1 = l_H$ and $l_2 = l_t$ (6.5)

For $\Delta\eta_t$: If $|\eta_b - \eta_{l_1}| - |\eta_b - \eta_{l_2}| \geq 0$, then: $l_1 = l_H$ and $l_2 = l_t$
else: $l_1 = l_t$ and $l_2 = l_H$ (6.6)

$$\text{For } \Delta m_t^{\text{pred}}: \quad \text{If } m_t^{\text{pred}}(l_1) - m_t^{\text{pred}}(l_2) \geq 0, \text{ then: } l_1 = l_H \text{ and } l_2 = l_t \quad \text{else: } l_1 = l_t \text{ and } l_2 = l_H \tag{6.7}$$

As the variables are positive for most events and not for all, the assignment is not expected to be always correct, but mostly. However, if the rate of correctly assigned light leptons is significantly larger than 50 % (which is expected for random assignment), the strategy is worth being applied. This has to be checked on *reco-level*.

6.2 Truth matching

So far studies have been conducted on full *truth-level*, but in the end, they have to be applied on *reco-level*. For this transition, it is imported to understand that in the data file, which contains the kinematic variables of all simulated events, there is no direct correspondence with the variables on *truth-level* and those on *reco-level*.

6.2.1 Variable structure on *truth-level*

Variables on *truth-level* refer to the particles of the process; the top quark, the Higgs boson and their respective decay products. The variables of the decay products are naturally assigned to their mother particles. This allows a consistent naming scheme also for particles that cannot be detected in ATLAS. This naming scheme is convenient for *truth-level*-studies and independent of selection cuts. Examples for variable names on *truth-level* are `truth_l_from_t_pt` for the transverse momentum of the lepton from the top quark-decay and `truth_H_tauM_FromH_pt` for the transverse momentum of the negative tau from the Higgs boson-decay.

6.2.2 Variable structure on *reco-level*

Variables on *reco-level* refer to the particles detected in the detector and the naming scheme follows the selection cuts. Consequently, the variables of the final state particles are related to the particle types, as they are determined through the procedures described in chapter 3. In order to distinguish particles of the same type, they are sorted according to their experimental value for p^{T}. Examples for variable names on *reco-level* are `pt_lep1`, `pt_lep2` and `pt_lep3` for the transverse momenta of all three leptons, including the hadronic tau.

6.2.3 Strategy for truth matching

As the naming schemes for on *truth-level* fundamentally differs from that on *reco-level*, there is no genuine matching between the variable names (except for particles for which the selection cuts select exactly one of its type, as for the b-jet). In fact, a requirement for truth matching is the assignment itself. So, doing the truth matching before the assignment is a challenge that needs an extra strategy.

Using kinematic properties

One possible strategy is using kinematic properties also for truth matching. For example, it is possible to sort the variables also on *truth-level* with respect to p^{T} and match them accordingly. A drawback of this

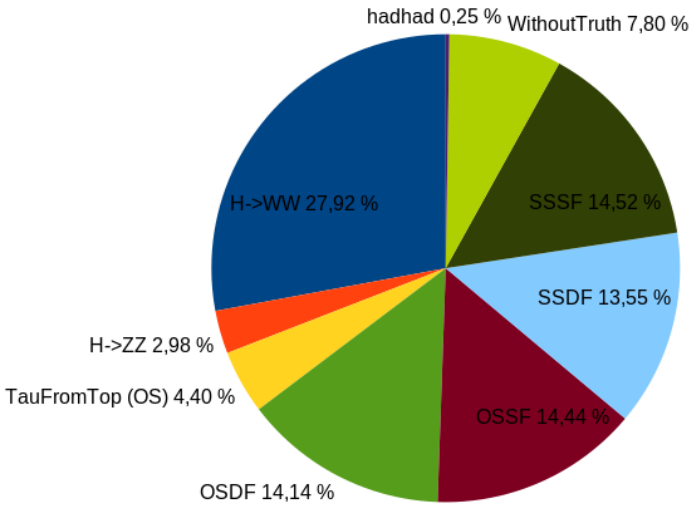


Figure 6.5: Composition of the tHq sample with multilepton filter (see section 4.2.4) using `SingleTop Ntuples` of version 32 and after applying the selection cuts described in section 4.3.2. The events of $H \rightarrow \tau\tau$ with hadronic tau originating from the Higgs boson are categorized by charge and flavor constellations (*OSDF*, *OSSF*, *SSDF* and *SSSF*) of the two light leptons. "TauFromTop (OS)" means that the hadronic tau originates from the top quark. Further explanation in section 6.2.4.

strategy is, that the values for p^{T} differ between *truth-level* and *reco-level*, as on *reco-level* distortions due to the detector are taken into account. In some cases that would lead to a different sorting on *truth-level* and *reco-level* and to incorrect matching, where it is difficult to estimate the rate of incorrect matchings. Taking other kinematic variables than p^{T}, e.g. η or ϕ, into account should minimize this rate, but in the end, the whole method seems to be unnecessarily complicated. A simpler strategy is favored.

Using information on charge and flavor

For *OS* events the truth matching is very simple, as the assignment strategy is as well. This opens the idea to use charge and flavor information of the light leptons for truth matching. For this strategy both the *OS* events and the *SS* events are categorized, whether the light leptons have different flavor (*DF*) or same flavor (*SF*). In short, those four categories are called *OSDF*, *OSSF*, *SSDF* and *SSSF*. For the *SSSF* events, the matching does not work. For all other categories, the light leptons can unambiguously be matched, because the two light leptons differ in at least one property (charge, flavor or both). The assignment strategy is tested for the first three categories separately. If there is any effect due to charge or flavor, it will be seen by comparing the *OSDF* with the *OSSF* events and the *OSDF* with the *SSDF* events. But if there is no effect, which is expected, the results of the first three categories can be transferred to the *SSSF* events.

6.2.4 Sample splitting

In order to test the assignment strategy for the *OSDF*, *OSSF* and *SSDF* events separately, the event sample is split. For splitting, information on *truth-level* is used, which allows relating every event of the tHq sample according to its decay channel. It is important to know that this relating happens after

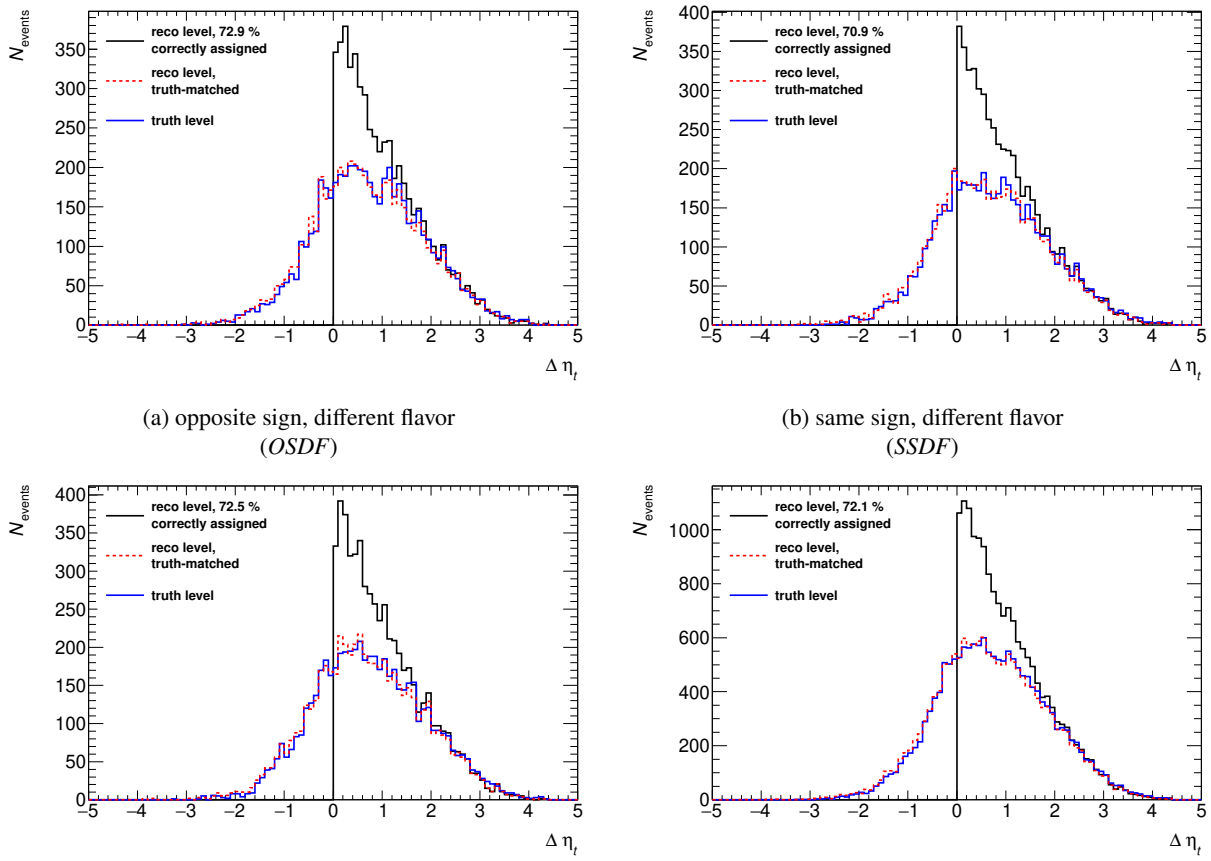
applying the selection cuts, which are applied on *reco-level* (see chapter 4). For this reason, besides the decay channel of interest, $H \rightarrow \tau\tau$, also $H \rightarrow WW$ and $H \rightarrow ZZ$ are found in the sample. Because information on *truth-level* slightly differs from the information on *reco-level*, a small number of events with two hadronic taus are in the sample (called "hadhad") as well. Only events of $H \rightarrow \tau\tau$ are selected for further categorization into *SSSF*, *SSDF*, *OSSF* and *OSDF*, because they are the most important for mass reconstruction (see section 4.3.3). Additionally, for *OS* events, both light leptons may originate from the $H \rightarrow \tau\tau$ (see section 4.4.2). Those events, labeled with "TauFromTop (OS)", will be separated from the other *OS* events and not be included in *OSDF* or *OSSF*. Events "WithoutTruth" that, due to technical reasons, lack information on *truth-level*, have to be separated as well.

Sample composition

The sample composition with all the aforementioned categories is shown in figure 6.5. Unfortunately, with about 8 % the fraction of events without truth information is unexpectedly high, but within the scope of this analysis, it has simply to be accepted. That the fraction of "hadhad" events is so small (< 0.3 %) argues for a good tau reconstruction on *reco-level*. Adding up ≈ 28 % for events of $H \rightarrow WW$ and ≈ 3 % for events of $H \rightarrow ZZ$, ≈ 61 % remain for $H \rightarrow \tau\tau$. It is due to the selection cuts (described in section 4.3.2) and the multilepton filter (see section 4.2.4) that the fraction of $H \rightarrow \tau\tau$ events is much higher than the branching ratios of those decays suggest (22.0 % for $H \rightarrow WW$, 6.26 % for $H \rightarrow \tau\tau$ and 2.73 % for $H \rightarrow ZZ$, see table 2.6). However, the fraction of $H \rightarrow ZZ$ events to $H \rightarrow WW$ events in the sample (≈ 28 % and ≈ 3 %) does not deviate that much from the fraction of their branching ratios (≈ 22.0 % and ≈ 2.73 %). This is no issue, as there is no reason, why the suppression of $H \rightarrow WW$ and $H \rightarrow ZZ$ should differ too much. In principle, this could be checked by considering all subsequent decays of $H \rightarrow WW$, $H \rightarrow ZZ$ and $H \rightarrow \tau\tau$, adding up the branching ratios of those decay channels which pass the selection cuts and comparing them to those decay channels which do not pass. On the other hand, the fractions of *OSDF* (14.14 %), *OSSF* (14.44 %), *SSDF* (13.55 %) and *SSSF* events (14.52 %) are very similar. As the branching ratios of the leptonic tau decays are very similar as well (see table 2.4), this is a convincing argument that the sample splitting went correct. At least for the following analysis, it is not necessary to check the sample composition by calculations.

6.2.5 Evaluation of truth matching

Of course, also for the strategy using the information on charge and flavor, the matching is not correct to 100 % because reconstruction of flavor and charge is neither (as has been discussed in chapter 3), but this rate can be determined easily. The three categories used for the truth matching (*OSDF*, *OSSF* and *SSDF*) contain 16531 events. For all those events, it is checked, if all pdgIds of the three leptons on *truth-level* are the same as on *reco-level*. It turns out that for 457 out of 16531 events not all pdgIds of the three leptons on *truth-level* are the same as on *reco-level*, which corresponds to a rate of 2.76 %. For those events, the matching explained above fails. Additionally, there may be events, for which two or more pdgIds are simultaneously wrong but in a way that the errors cancel each other out and can not be detected (e.g. an electron is misidentified as a positron and a positron is misidentified as an electron in the same event). But this effect should concern less than $2.76\ \% \times 2.76\ \%$ of the events, i.e. it should be very rare and negligible.



(a) opposite sign, different flavor (*OSDF*)

(b) same sign, different flavor (*SSDF*)

(c) opposite sign, same flavor (*OSSF*)

(d) histograms of figures (a), (b) and (c) added up (*OSDF* + *SSDF* + *OSSF*)

Figure 6.6: Number of events as a function of $\Delta\eta_t$ as defined in equation 6.4 for different sign and flavor combinations of the light leptons in comparison. Blue solid line: $\Delta\eta_t$ is calculated on *truth-level*. Red dotted line: $\Delta\eta_t$ is calculated using variables on *reco-level* that have been matched to the variables on *truth-level*. Black solid line: $\Delta\eta_t$ is calculated on *reco-level* after it has been used for assigning the light leptons to the Higgs boson and the top quark. Comparing the pdgIds of the assigned variables with the matched variables allows calculating the rate of correctly assigned leptons in percent.

6.3 Testing kinematic variables on *reco-level*

Using the truth-matching strategy explained before, it is possible to determine the rate of correctly assigned light leptons on *reco-level*. Beforehand, the truth-matching strategy has to be tested.

6.3.1 Testing the truth-matching strategy

Description Figure 6.6 shows the distribution of $\Delta\eta_t$, separately for *OSDF* (figure 6.6(a)), *SSDF* (figure 6.6(b)), *OSSF* (figure 6.6(c)) and all of those three categories together (figure 6.6(d)). In each of those four diagrams, the histogram with the blue solid line shows the distribution on full *truth-level* (the same as in figure 6.3(a) but with another set of events). For the histogram depicted by the red dashed line, variables on *reco-level* have been used for calculation, after they have been matched to the variables on *truth-level* using the truth-matching strategy described in section 6.2.3. The black line shows the distribution using variables on full *reco-level*, i.e. without truth-matching but using $\Delta\eta_t$ for assignment instead. The rate of correctly assigned light leptons (short: assignment rate) is shown in the legend of each diagram.

Evaluation Some general properties can be seen in any of those four diagrams. The black-lined histogram has only entries for positive values, but this is expected, as the variable plotted is the same which is used for assignment. It has been plotted for cross-checking. Both the blue and the red histogram have more entries for positive values than for negative values, as has been seen before for the studies on *truth-level* only. The red dashed line has only small deviations from the blue line, which is as expected since the values for η slightly differ between *reco-level* and *truth-level*. If the deviations were too large, this would imply bad truth-matching. So the histograms are consistent with expectations and for each diagram the truth-matching works fine. Comparing the diagrams to each other, only small differences between those can be seen. The diagram for *OSDF* events in figure 6.6(a) has a bump at $\Delta\eta_t \approx 1$. The assignment rate varies between 70.9 % and 72.9 %.

Conclusion May these differences be systematic or statistical fluctuations, the crucial factor is, that the differences between *truth-level* and truth-matched *reco-level* are small and for every category, they are small to the same extent, which means that the truth matching in itself does not depend on the category. The differences are a bit smaller in figure 6.6(d), which is reasonable, as this histogram has better statistics. Therefore, for $\Delta\eta_t$ the assignment strategy can also be applied to *SSSF* events.

6.3.2 Comparing assignment rate of kinematic variables

The rate of correctly assigned light leptons (short: assignment rate), which is only known for $\Delta\eta_t$ by now, shall be compared to the other variables which have already been investigated on *truth-level*. The most promising of those variables is Δm_t^{pred}. As there is no simple strategy to predict the neutrino momenta from the Higgs boson decay, unfortunately, it is not possible to construct a variable analogous to Δm_t^{pred} using the decay products of the Higgs boson. But it is possible to use visible decay products.

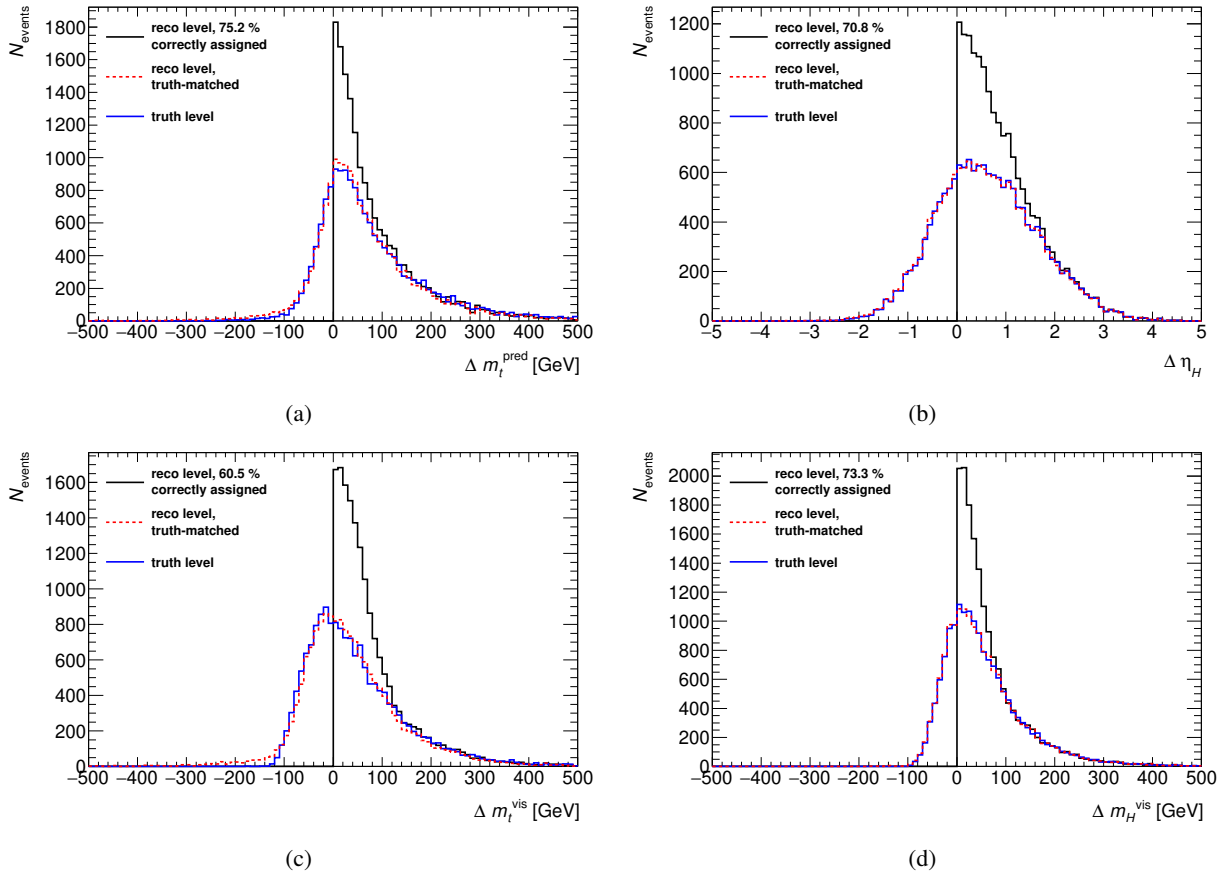


Figure 6.7: Number of events as a function of Δm_t^{pred} (a), $\Delta\eta_H$ (b), Δm_t^{vis} (c) and Δm_H^{vis} (d) as defined in equations 6.2, 6.3, 6.8 and 6.9 for comparing their assignment rate. For all histograms, only $H \rightarrow \tau\tau$ events (ignoring $H \rightarrow WW$ and $H \rightarrow ZZ$) of the categories *OSDF*, *OSSF* and *SSDF* are used. Events, where the hadronic tau originates from the top quark (called "TauFromTop (OS)") as well as events, where the two light leptons have the same flavor and same sign in electric charge (called *SSSF*) are not included.

For comparison, this is also done for the top quark. Hence, the following variables are defined:

$$\Delta m_t^{\text{vis}} := \sqrt{\left(P_b + P_{l_H}\right)^2} - \sqrt{\left(P_b + P_{l_t}\right)^2} \tag{6.8}$$

$$\Delta m_H^{\text{vis}} := \sqrt{\left(P_{\tau_{\text{had}}} + P_{l_t}\right)^2} - \sqrt{\left(P_{\tau_{\text{had}}} + P_{l_H}\right)^2} \tag{6.9}$$

In order to compare those variables, figure 6.7 shows histograms for Δm_t^{pred} 6.7(a), $\Delta\eta_H$ 6.7(b), Δm_t^{vis} 6.7(c) and Δm_H^{vis} 6.7(d) as defined in equations 6.2, 6.3, 6.8 and 6.9. With a rate of 75.2 %, Δm_t^{pred} has the highest assignment rate. It is followed by Δm_H^{vis} with 73.3 %, $\Delta\eta_t$ (figure 6.6(d)) with 72.1 % and $\Delta\eta_H$ with 70.8 %. Δm_t^{vis} has by far the lowest assignment rate with 60.5 %. Also for those variables, the truth-matching strategy is tested the same way as explained in section 6.3.1 and with the same conclusion as for $\Delta\eta_t$, i.e. that the assignment strategy can also be applied for *SSSF* events. The corresponding histograms can be found in figure **??** ($\Delta\eta_H$), **??** (Δm_t^{pred}), **??** (Δm_t^{vis}) and **??** (Δm_H^{vis}) in appendix **??**.

6.4 Optimization of the assignment rate

6.4.1 Combining variables

So far the assignment rate over all events has been the only parameter considered for each variable. Under the assumption that all events have to be included in the assingment strategy and that only one variable is used, this is an obvious strategy in order to find the best variable. But in principle, it is also possible to exclude events from assignment that are difficult to assign or to combine variables. From this point of view, other properties of the variables are interesting to be investigated. Already in section 6.1.3, a strong asymmetry in the distribution of Δm_t^{pred} with a long tail to positive values has been seen on *truth-level* (see figure 6.4). Also, the distribution of Δm_H^{vis} in figure 6.7(d) shows this asymmetry, but with the beneficial property that there are no entries for negative values beyond a limit at around $-100\,\text{GeV}$. For events with $|\Delta m_H^{\text{vis}}| < 100\,\text{GeV}$ the value for Δm_H^{vis} is always positive and the assignment should be always correct. Therefore, it is feasible to exclude events with $|\Delta m_H^{\text{vis}}| < 100\,\text{GeV}$ from assignment in order to optimize the assignment rate. Instead of excluding those events, it is possible to use a different variable for assigning those events, e.g. Δm_t^{pred}, since this variable has the highest assignment rate in general. In fact, combining these two variables in the previously explained way leads to a total assignment rate of 79.1 %, which is a significant improvement compared to 75.2 %, which was achieved using only Δm_t^{pred}.

Optimization However, the value for the assignment rate achieved so far is not the highest possible, as can be seen by considering the numbers. 3705 out of 16531 events passed the cut $|\Delta m_H^{\text{vis}}| > 100\,\text{GeV}$ and 3666 events of those have been assigned correctly using Δm_H^{vis}, which is 98.9 %. The remaining 12826 events have been assigned using m_t^{pred}, but with 9404 events only 73.3 % have been assigned correctly. In comparison, the assignment rate for m_t^{pred} using all events is 75.2 %. This means that Δm_H^{vis} and Δm_t^{pred} are correlated somehow. Selecting events for increasing the assignment rate of Δm_H^{vis} can decrease the assignment rate of Δm_t^{pred}. So the maximal assignment rate for the combination of Δm_H^{vis} and Δm_t^{pred} may be reached with another threshold than $|\Delta m_H^{\text{vis}}| = 100\,\text{GeV}$. Indeed, running the

assignment strategy for varying threshold between 0 and 150 GeV finds an optimum at a threshold of $|\Delta m_H^{\text{vis}}| = 57\,\text{GeV}$ with a total assignment rate of 80.8 %.

6.4.2 Outlook to further improvements

The optimizations described in the last subsection are not all that can be done.

Optimizing the assignment rate First, there may be other kinematic variables with a higher assignment rate. Variables like ϕ and p^{T} or combinations of those like $E = p^{\text{T}} \cosh\eta$ and $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ have not been tested yet. Second, as already hinted at in the last section, when combining variables, the correlation of those matters. There may be pairs of variables, which by themselves, do not lead to a large assignment rate, but lead to a large assignment rate when used in combination because they are anticorrelated. Third, not only two but three or more variables can be combined. The order of those and the thresholds for cutting lead to a lot of free parameters to be optimized. For this optimization method, it is possible to fix the order of the variables and use an analytical optimization method, like the gradient descent method. But it is also thinkable to use machine learning techniques, like artificial neural networks (ANN) or boosted decision trees (BDT).

Finding events with the hadronic tau originating from the top quark For the analysis in this thesis, the hadronic tau was assumed to always originate from the Higgs boson. Using the fractions for the event categories "TauFromTop (OS)", *OSDF* and *OSSF* in figure 6.5, where only $H \to \tau\tau$-events are considered and $H \to WW$ and $H \to ZZ$-events are ignored, the fraction of the *OS* events for which the assumption is not true is

$$\frac{4.40\,\%}{4.40\,\% + 14.14\,\% + 14.44\,\%} = 13.34\,\%.$$

In other words, for 86.66 % of events the assumption is correct. This fraction is higher than the optimized assignment rate for *SS* events of 80.8 %. However, it might be possible to increase this fraction of 86.66 % with a strategy similar to the assigning strategy described in this chapter. Therefore, kinematic properties of the hadronic tau relative to the b-jet, e.g. $m^2_{b\tau} = \left(P_b + P_{\tau_{\text{had}}}\right)^2$ or $\Delta\eta_{b\tau} = |\eta_b - \eta_{\tau_{\text{had}}}|$ could be exploited, in order to apply them on *OS* events and identify events, where the hadronic tau originates from the top quark. Since for *OS* events the assignment of the light leptons to their respective mother particles is already done at this step (using the electric charge), it is also possible to investigate kinematic properties of the hadronic tau relative to the light lepton assigned to the Higgs boson, e.g. $m^2_{l_H\tau} = \left(P_{l_H} + P_{\tau_{\text{had}}}\right)^2$ or $\Delta\eta_{l_H\tau} = |\eta_{l_H} - \eta_{\tau_{\text{had}}}|$. Those variables, or combinations of those, might help to determine if the hadronic tau originates from the top quark or from the Higgs boson. This should also lead to a better mass reconstruction.

CHAPTER 7

Conclusion

The Standard Model (SM) of elementary particle physics is very successful. Therefore, ongoing research deals with testing the SM. One way is investigating y_t, the complex coupling constant between the Higgs boson and the top quark. It is an open question if its magnitude, $y_t = 0.995$, is close to 1 by coincidence. Furthermore, the SM predicts a relative sign of -1 between y_t and g_{HVV}, which is not tested experimentally yet. Deviations in the relative sign would be hints to physics beyond the Standard Model (BSM).

While the search for Higgs-boson production in association with a $t\bar{t}$-pair allows for probing the magnitude of y_t, searching for tHq events is essential for probing its phase. tHq is very rare. In the $2\,e/\mu + 1\tau_{\text{had}}$ decay channel, a subchannel of $H \to \tau\tau$, it suffers from a bad signal-to-background ratio, especially for events where the two light leptons have the opposite electric charge (*OS* events).

In order to improve the separation from background, this work has been dedicated to the mass reconstruction of the $\tau\tau$-state. This is especially important for the separation from tZq. Pioneer work on this field has been performed in a former master thesis [77]. The overall strategy, predicting $\overrightarrow{\not{E}}_{\text{T}}$ (top) in order to use established methods for the mass reconstruction of the $\tau\tau$-state, has been adopted.

Correlation studies performed on *truth-level* revealed that the basic assumption of the old approach, $p_\nu^z = \alpha p_l^z$, is not true. Moreover, correlations between p_ν^{T} and p_l^{T} and between ϕ_ν and ϕ_l have been found. Exploiting these correlations allows for the prediction of p_ν^x and p_ν^y in a much easier way than the old approach. With the new approach, in addition, no events have to be discarded, which is especially important for the analysis of such a rare process as tHq for statistical reasons. Using both the old and the new approach for mass reconstruction, the direct comparison showed that the mass reconstruction with the new approach is clearly more accurate. This could not only be seen by direct comparison but also by comparing them to the distribution using information on *truth-level*, which gives a benchmark for the prediction in addition.

It was also possible to compare three different mass reconstruction methods: the transverse mass method and two different formulas using the collinear approximation. Among those three methods, the collinear approximation using formula 2 [80] has proven to be the best.

The correlations exploited for the prediction could also be explained. Essential is that the top quark is heavier than the W boson, which allows the W boson to be on-shell. The found kinematic pattern is possible because the lepton and the neutrino from the decay of the W boson are much lighter than the W boson. It can be explained with the functional dependence of the W-boson mass on the kinematic variables. This is interesting because it means that the correlations should be present at any leptonic

top-quark decay and could be exploited in a variety of other analyses.

Another challenge in the mass reconstruction of tHq in the $2\,e/\mu + 1\tau_{\text{had}}$ decay channel is the assignment of the two light leptons to their respective mother particle (Higgs boson or top quark) for *SS* events. The correlations exploited for the prediction could also be exploited for the assignment, by creating a variable, m_t^{pred}, with a high discrimination power. A baseline method using only this variable was able to achieve already 75.2 % of correctly assigned events. The combination of m_t^{pred} with another variable achieved an assignment rate of 80.8 %.

All in all, the analysis presented in this thesis is a great success. The mass reconstruction could be improved by replacing the existing approach with a new one, which is handier, more accurate and applicable for all events. Moreover, also the *SS* events can now be included, which is especially important since the signal-to-background-ratio is more advantageous for those events.

With the following steps, the analysis can be even extended. Another mass reconstruction method, the missing mass calculator (MMC), is very promising and can be implemented in the analysis. The correlations between b-quark and lepton and between b-quark and neutrino from the top-quark decay have not been investigated thouroughly yet. Maybe the prediction of p_ν^x and p_ν^y can still be optimized. The assignment strategy for events where the two light leptons have the same electric charge (*SS* events) presented in this thesis only exploits two kinematic variables so far. Adding more variables and combinining them with more sophisticated techniques, e.g. machine learing techniques, should increase the assignment rate.

Bibliography

[1] A. M. Sirunyan et al., *Search for associated production of a Higgs boson and a single top quark in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Rev. D **99** (2019) 092005, arXiv: 1811.09696 [hep-ex] (cit. on pp. 1, 16–18).

[2] [Online; accessed 11-February-2023], 2008, URL: https://plato.stanford.edu/archives/fall2008/entries/atomism-ancient/ (cit. on p. 3).

[3] M. Born, W. Heisenberg and P. Jordan, *Zur Quantenmechanik. II.*, Zeitschrift für Physik **35** (1926) 557, URL: https://doi.org/10.1007/bf01379806 (cit. on p. 3).

[4] R. Shankar, *Principles of Quantum Mechanics*, Springer US, 1994, URL: https://doi.org/10.1007/978-1-4757-0576-8 (cit. on p. 3).

[5] J. Faye, "Copenhagen Interpretation of Quantum Mechanics", *The Stanford Encyclopedia of Philosophy*, ed. by E. N. Zalta, Winter 2019, Metaphysics Research Lab, Stanford University, 2019 (cit. on p. 4).

[6] M. Thomson, *Modern particle physics*, New York: Cambridge University Press, 2013, ISBN: 978-1-107-03426-6, URL: https://doi.org/10.1017/cbo9781139525367 (cit. on pp. 5, 12, 13, 15, 57).

[7] W. Nolting, *Grundkurs Theoretische Physik 4*, Springer Berlin Heidelberg, 2010, URL: https://doi.org/10.1007/978-3-642-01604-2 (cit. on p. 6).

[8] JabberWok, *A graphic showing the relationship between angle and pseudorapidity*, [Online; accessed 12-November-2021], 2017, URL: https://commons.wikimedia.org/wiki/File:Pseudorapidity2.png (cit. on p. 8).

[9] C. MissMJ, *Standard model of elementary particles: the 12 fundamental fermions and 5 fundamental bosons. Brown loops indicate which bosons (red) couple to which fermions (purple and green). Please note that the masses of certain particles are subject to periodic reevaluation by the scientific community. The values currently reflected in this graphic are as of 2019 and may have been adjusted since. For the latest consensus, please visit the Particle Data Group website linked below.* [Online; accessed 12-November-2021], 2019, URL: https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg (cit. on p. 9).

[10] P. A. Zyla et al., *Review of Particle Physics*, 2020, URL: https://pdglive.lbl.gov/Particle.action?node=S043 (cit. on p. 8).

[11] P. A. Zyla et al., *Review of Particle Physics*, 2020, URL: https://pdglive.lbl.gov/Particle.action?node=S044 (cit. on p. 9).

[12] M. Tanabashi et al., *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001 (cit. on p. 10).

[13] P. A. Zyla et al., *Review of Particle Physics*, 2020, URL: https://pdglive.lbl.gov/Particle.action?node=S035&init=0 (cit. on p. 11).

[14] 2015, URL: https://www.nobelprize.org/prizes/physics/2015/summary/ (cit. on p. 11).

[15] T. Plehn, "LHC Phenomenology for Physics Hunters", *Theoretical Advanced Study Institute in Elementary Particle Physics: The Dawn of the LHC Era*, 2010 125, arXiv: 0810.2281 [hep-ph] (cit. on p. 12).

[16] P. A. Zyla et al., *Review of Particle Physics*, 2020, URL: https://pdglive.lbl.gov/Particle.action?node=Q007 (cit. on p. 12).

[17] G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex] (cit. on p. 13).

[18] P. A. Zyla et al., *Review of Particle Physics*, 2020, URL: https://pdglive.lbl.gov/Particle.action?node=S126&init=0 (cit. on p. 13).

[19] G. Aad et al., *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*, Eur. Phys. J. C **76** (2016) 6, arXiv: 1507.04548 [hep-ex] (cit. on p. 13).

[20] J. R. Andersen et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, ed. by S. Heinemeyer, C. Mariotti, G. Passarino and R. Tanaka, 2013, arXiv: 1307.1347 [hep-ph] (cit. on pp. 13, 17).

[21] J. at English Wikipedia, *A Feynman diagram of an s-channel process*, [Online; accessed 12-November-2021], 2006, URL: https://en.wikipedia.org/wiki/File:S-channel.svg (cit. on p. 14).

[22] J. at English Wikipedia, *A Feynman diagram of an t-channel process*, [Online; accessed 12-November-2021], 2006, URL: https://en.wikipedia.org/wiki/File:T-channel.svg (cit. on p. 14).

[23] J. at English Wikipedia, *A Feynman diagram of an u-channel process*, [Online; accessed 12-November-2021], 2006, URL: https://en.wikipedia.org/wiki/File:U-channel.svg (cit. on p. 14).

[24] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Reading, USA: Addison-Wesley, 1995, ISBN: 978-0-201-50397-5 (cit. on p. 16).

[25] K. Flöh, *Investigation of the Top Quark Yukawa Coupling in Higgs Boson Production in Association with Top Quarks at 13 TeV with the CMS Experiment*, PhD thesis: KIT, Karlsruhe, 2020 (cit. on p. 16).

[26] A. M. Sirunyan et al., *Observation of $t\bar{t}H$ production*, Phys. Rev. Lett. **120** (2018) 231801, arXiv: 1804.02610 [hep-ex] (cit. on p. 16).

[27] M. Aaboud et al., *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173, arXiv: 1806.00425 [hep-ex] (cit. on p. 16).

[28] D. de Florian et al.,
Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector, vol. 2/2017, 2016, arXiv: 1610.07922 [hep-ph] (cit. on p. 17).

[29] F. Demartin, F. Maltoni, K. Mawatari and M. Zaro,
Higgs production in association with a single top quark at the LHC, Eur. Phys. J. C **75** (2015) 267, arXiv: 1504.00611 [hep-ph] (cit. on p. 17).

[30] DESY, *Der LHC am CERN liegt in Genf im Grenzgebiet zwischen Frankreich und der Schweiz. In der Grafik sind auch die vier großen Detektoren ATLAS, CMS, ALICE und LHCb eingezeichnet.* [Online; accessed 09-December-2021], 2008,
URL: https://www.weltmaschine.de/service__material/mediathek/lhc
(cit. on p. 20).

[31] C. CERN Education and O. Group, *CERN-Brochure-2021-004-Eng*, 2021, URL:
https://cds.cern.ch/record/2809109/files/CERN-Brochure-2021-004-Eng.pdf
(cit. on p. 19).

[32] *LHC Machine*, JINST **3** (2008) S08001, ed. by L. Evans and P. Bryant (cit. on p. 19).

[33] CERN, *News*, 2022, URL: https://home.cern/news/news/physics/lhc-run-3-physics-record-energy-starts-tomorrow (cit. on p. 20).

[34] E. Ward, "LHC and HL-LHC timeline for ATLAS website", General Photo, 2018,
URL: http://cds.cern.ch/record/2652466 (cit. on p. 20).

[35] H. Wiedemann, *Particle Accelerator Physics*, Graduate Texts in Physics,
Berlin, Germany: Springer, 2015, ISBN: 978-3-319-18316-9 (cit. on p. 19).

[36] H. Kolanoski and N. Wermes, *Particle Detectors*, Oxford University Press, 2020,
ISBN: 978-0-19-885836-2 (cit. on pp. 19, 23).

[37] L. Evans and L. Evans, *The Large Hadron Collider: a marvel of technology; 2nd ed.*
Physics (EPFL Press), On the cover : Including the discovery of the higgs boson,
Lausanne: EPFL Press, 2018, ISBN: 978-2-88914-482-2,
URL: http://cds.cern.ch/record/2645935 (cit. on p. 21).

[38] J. Wenninger, *LHC status and performance*, PoS **CHARGED2018** (2019) 001,
URL: https://cds.cern.ch/record/2710042 (cit. on p. 21).

[39] G. Aad et al., *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det] (cit. on pp. 21, 37).

[40] G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*,
JINST **3** (2008) S08003, URL: http://cds.cern.ch/record/1129811 (cit. on pp. 22, 26).

[41] L. Carminati and G. Marchiori,
Physics with Photons Using the ATLAS Run 2 Data: Calibration and Identification, Measurement of the Higgs Boson Mass and Search for Supersymmetry in Di-Photon Final State,
Springer Theses, Springer, 2019, ISBN: 978-3-030-24369-2, 978-3-030-24370-8 (cit. on p. 22).

[42] N. Ilic, *Observation of the Higgs Boson Decaying to $WW^* \to \ell\nu\,\ell\nu$*, PhD thesis: Toronto U., 2015 (cit. on p. 23).

[43] J. Pequenao, "Event Cross Section in a computer generated image of the ATLAS detector.", 2008, URL: https://cds.cern.ch/record/1096081 (cit. on p. 24).

[44] C. Grupen and B. Schwartz, *Particle detectors*, Cambridge, UK: Cambridge Univ. Pr., 2008 (cit. on p. 23).

[45] C. AC, "Layout of ATLAS", 1998, URL: https://cds.cern.ch/record/39038 (cit. on p. 25).

[46] J. Pequenao, "Computer generated image of the ATLAS inner detector", 2008, URL: https://cds.cern.ch/record/1095926 (cit. on p. 26).

[47] J. Montejo Berlingen, *Triggering in the ATLAS Experiment*, PoS **ICHEP2020** (2021) 756, URL: http://cds.cern.ch/record/2742661 (cit. on p. 26).

[48] M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, (2010), URL: https://cds.cern.ch/record/1291633 (cit. on p. 27).

[49] A. La Rosa, *The ATLAS Insertable B-Layer: from construction to operation*, JINST **11** (2016) C12036, ed. by C. Gemme and L. Rossi, arXiv: 1610.01994 [physics.ins-det], URL: https://arxiv.org/pdf/1610.01994.pdf (cit. on p. 27).

[50] M. Aaboud et al., *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv: 1704.07983 [hep-ex], URL: https://arxiv.org/abs/1704.07983 (cit. on pp. 28, 32).

[51] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444 (cit. on p. 29).

[52] G. Aad et al., *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex] (cit. on pp. 29, 30).

[53] G. Aad et al., *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **81** (2021) 578, arXiv: 2012.00578 [hep-ex], URL: https://arxiv.org/abs/2012.00578 (cit. on pp. 29, 34).

[54] M. Aaboud et al., *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex], URL: https://arxiv.org/pdf/1703.10485.pdf (cit. on p. 30).

[55] G. Aad et al., *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 689, arXiv: 2007.02645 [hep-ex], URL: https://arxiv.org/abs/2007.02645 (cit. on p. 31).

[56] G. Aad et al., *New techniques for jet calibration with the ATLAS detector*, (2023), arXiv: 2303.17312 [hep-ex], URL: https://arxiv.org/pdf/2303.17312.pdf (cit. on p. 31).

[57] L. Scodellaro, "b tagging in ATLAS and CMS", *5th Large Hadron Collider Physics Conference*, 2017, arXiv: 1709.01290 [hep-ex] (cit. on p. 31).

[58] *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, (2022), arXiv: 2211.16345 [physics.data-an], URL: https://arxiv.org/abs/2211.16345 (cit. on p. 31).

[59] *Calibration of the light-flavour jet mistagging efficiency of the b-tagging algorithms with Z+jets events using 139* fb^{-1} *of ATLAS proton-proton collision data at* $\sqrt{s} = 13$ *TeV*, (2023), arXiv: 2301.06319 [hep-ex], URL: https://cds.cern.ch/record/2846342 (cit. on p. 31).

[60] *Identification of the Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector*, (2013), URL: https://cds.cern.ch/record/1562839 (cit. on p. 32).

[61] *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at* $\sqrt{s} = 13\,TeV$, (2017), URL: https://cds.cern.ch/record/2261772 (cit. on p. 32).

[62] *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, (2015), URL: https://cds.cern.ch/record/2064383 (cit. on p. 32).

[63] *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, (2019), URL: http://cds.cern.ch/record/2688062 (cit. on p. 32).

[64] G. Aad et al., *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex], URL: https://arxiv.org/abs/1908.00005 (cit. on p. 33).

[65] M. Aaboud et al., *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv: 1902.04655 [physics.ins-det] (cit. on p. 33).

[66] URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronChargeFlipTaggerTool (cit. on p. 34).

[67] T. I. Carlson, *Pre-HCW2022: Missing transverse momentum and pile-up tagging*, [Online; accessed 17-March-2023], 2022, URL: https://indico.cern.ch/event/1188525/contributions/4994884/attachments/2496681/4288314/pre_HCW_MET_PUTagging_2022.pdf (cit. on p. 35).

[68] E_T^{miss} *performance in the ATLAS detector using 2015-2016 LHC pp collisions*, (2018), URL: http://cds.cern.ch/record/2625233 (cit. on p. 35).

[69] A. Buckley et al., *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011) 145, arXiv: 1101.2599 [hep-ph] (cit. on p. 37).

[70] S. Weinzierl, *Introduction to Monte Carlo methods*, (2000), arXiv: hep-ph/0006269 (cit. on p. 37).

[71] G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det] (cit. on p. 37).

[72] S. Agostinelli et al., *GEANT4–a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 37).

[73] URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/SingleTopProductionV32ll (cit. on p. 38).

[74] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph] (cit. on p. 38).

[75] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph] (cit. on p. 38).

[76] URL: https://root.cern/ (cit. on p. 38).

[77] H. N. We, *Associated Production of a Top Quark and a Higgs Boson in pp Collisions at 13 Tev Using the ATLAS Detector*, Master Thesis: University of Bonn, 2020, URL: https://www.pi.uni-bonn.de/brock/en/results/home#theses (cit. on pp. 39, 41, 42, 45–47, 83, 93–96).

[78] A. Elagin, P. Murat, A. Pranko and A. Safonov, *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, Nucl. Instrum. Meth. A **654** (2011) 481, arXiv: 1012.4686 [hep-ex], URL: https://arxiv.org/abs/1012.4686.

[79] A. Elagin, P. Murat, A. Pranko and A. Safonov, *A new mass reconstruction technique for resonances decaying to di-tau*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **654** (2011) 481, ISSN: 0168-9002, URL: http://www.sciencedirect.com/science/article/pii/S0168900211014112 (cit. on pp. 42–44, 91).

[80] P. Konar and A. K. Swain, *Reconstructing semi-invisible events in resonant tau pair production from Higgs*, Physics Letters B **757** (2016) 211, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/S0370269316300417 (cit. on pp. 43, 83, 92).

APPENDIX A

Collinear approximation

A.1 Formula 1

The first approach [79] defines x_1 and x_2 as fractions of the magnitudes of the momenta:

$$x_1 = \frac{|\vec{p}_{\text{invis, 1}}|}{|\vec{p}_{\text{invis, 1}}| + |\vec{p}_{\text{visib, 1}}|} \qquad x_2 = \frac{|\vec{p}_{\text{invis, 2}}|}{|\vec{p}_{\text{invis, 2}}| + |\vec{p}_{\text{visib, 2}}|} \tag{A.1}$$

$|\vec{p}_{\text{visib, 1}}|$ and $|\vec{p}_{\text{visib, 2}}|$ are known. $|\vec{p}_{\text{invis, 1}}|$ and $|\vec{p}_{\text{invis, 2}}|$ are determined by writing equation 4.5 in spherical coordinates

$$\begin{aligned} \not{E}^x &= |\vec{p}_{\text{invis, 1}}| \sin\theta_{\text{invis, 1}} \cos\phi_{\text{invis, 1}} + |\vec{p}_{\text{invis, 2}}| \sin\theta_{\text{invis, 2}} \cos\phi_{\text{invis, 2}} \\ \not{E}^y &= |\vec{p}_{\text{invis, 1}}| \sin\theta_{\text{invis, 1}} \sin\phi_{\text{invis, 1}} + |\vec{p}_{\text{invis, 2}}| \sin\theta_{\text{invis, 2}} \sin\phi_{\text{invis, 2}} \end{aligned} \tag{A.2}$$

and, using that - following the idea of collinear approximation - all angles of the invisible decay products should be the same as the angles of the visible decay products,

$$\begin{aligned} \not{E}^x &= |\vec{p}_{\text{invis, 1}}| \sin\theta_{\text{visib, 1}} \cos\phi_{\text{visib, 1}} + |\vec{p}_{\text{invis, 2}}| \sin\theta_{\text{visib, 2}} \cos\phi_{\text{visib, 2}} \\ \not{E}^y &= |\vec{p}_{\text{invis, 1}}| \sin\theta_{\text{visib, 1}} \sin\phi_{\text{visib, 1}} + |\vec{p}_{\text{invis, 2}}| \sin\theta_{\text{visib, 2}} \sin\phi_{\text{visib, 2}} \end{aligned} \tag{A.3}$$

the two equations for two unknowns can be solved:

$$\begin{aligned} |\vec{p}_{\text{invis, 1}}| &= \frac{\not{E}^y \cos\phi_{\text{visib, 2}} - \not{E}^x \sin\phi_{\text{visib, 2}}}{\sin\theta_{\text{visib, 1}}(\sin\phi_{\text{visib, 2}} \cos\phi_{\text{visib, 1}} - \cos\phi_{\text{visib, 1}} \sin\phi_{\text{visib, 2}})} \\ |\vec{p}_{\text{invis, 2}}| &= \frac{\not{E}^x \sin\phi_{\text{visib, 1}} - \not{E}^y \cos\phi_{\text{visib, 1}}}{\sin\theta_{\text{visib, 2}}(\sin\phi_{\text{visib, 2}} \cos\phi_{\text{visib, 1}} - \cos\phi_{\text{visib, 1}} \sin\phi_{\text{visib, 2}})} \end{aligned} \tag{A.4}$$

The mass of the $\tau\tau$-state $m_{\tau\tau}$ is calculated via

$$m_{\tau\tau} = \frac{m_{\text{vis}}}{\sqrt{x_1 x_2}} \tag{A.5}$$

A.2 Formula 2

The second approach [80] applies the idea of collinear approximation by using proportional factors F_1 and F_2, which are defined for the vectors:

$$\begin{aligned} \vec{p}_{\text{invis, 1}} &= F_1 \, \vec{p}_{\text{visib, 1}} \\ \vec{p}_{\text{invis, 2}} &= F_2 \, \vec{p}_{\text{visib, 2}} \end{aligned} \tag{A.6}$$

F_1 and F_2 can be determined by applying the definition to equation 4.5

$$\begin{aligned} \not{E}^x &= F_1 \, p^x_{\text{visib, 1}} + F_2 \, p^x_{\text{visib, 2}} \\ \not{E}^y &= F_1 \, p^y_{\text{visib, 1}} + F_2 \, p^y_{\text{visib, 2}} \end{aligned} \tag{A.7}$$

and solving it for F_1 and F_2:

$$\begin{aligned} F_1 &= \frac{\not{E}^y p^x_{\text{visib, 1}} - \not{E}^x p^y_{\text{visib, 1}}}{p^y_{\text{visib, 1}} p^x_{\text{visib, 2}} - p^x_{\text{visib, 1}} p^y_{\text{visib, 2}}} \\ F_2 &= \frac{\not{E}^y p^x_{\text{visib, 2}} - \not{E}^x p^y_{\text{visib, 2}}}{p^y_{\text{visib, 1}} p^x_{\text{visib, 2}} - p^x_{\text{visib, 1}} p^y_{\text{visib, 2}}} \end{aligned} \tag{A.8}$$

F_1 and F_2 are related to f_1 and f_2 via

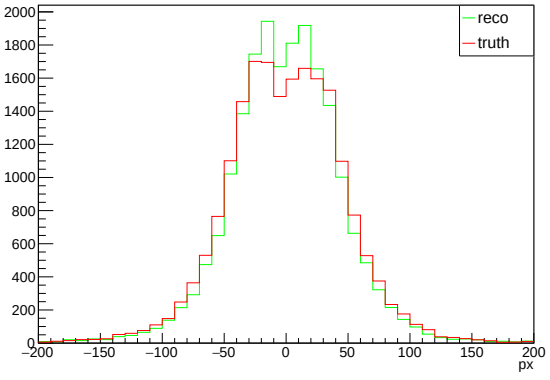
$$f_1 = \frac{1}{1 + |F_1|} \qquad f_2 = \frac{1}{1 + |F_2|} \tag{A.9}$$

Following the same idea as approach 1, but with different factors f_1 and f_2, $m_{\tau\tau}$ is calculated via

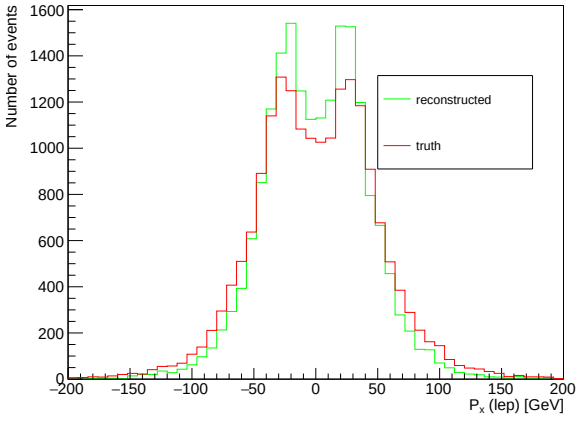
$$m_{\tau\tau} = \frac{m_{\text{vis}}}{\sqrt{f_1 f_2}} \tag{A.10}$$

APPENDIX B

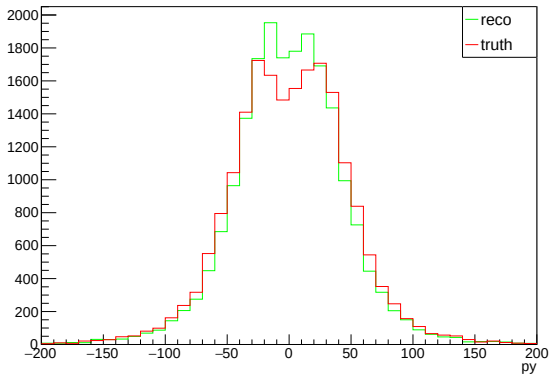
Truth studies



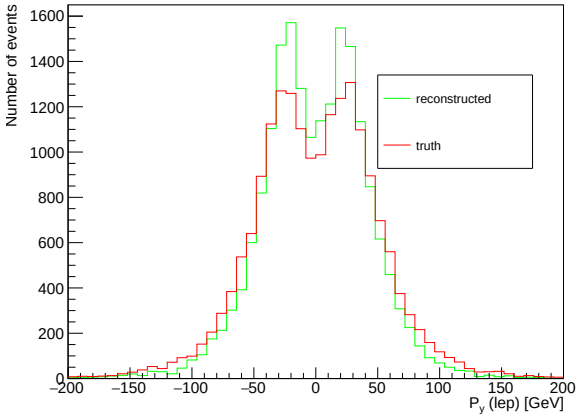
(a) Number of tHq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 32 (all campaigns) have been used.



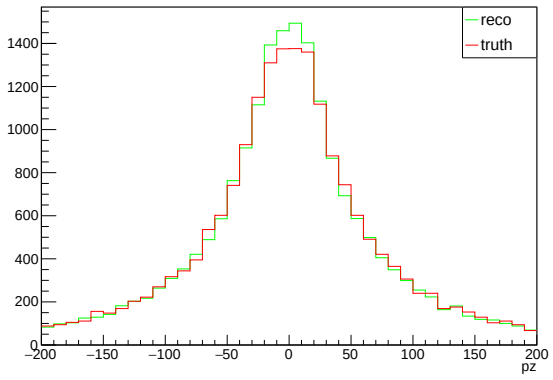
(b) Number of tZq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 29 (campaign mc16a) have been used. Events with anti-top quarks in the final state insteaf of top quarks have been removed. Taken from [77].



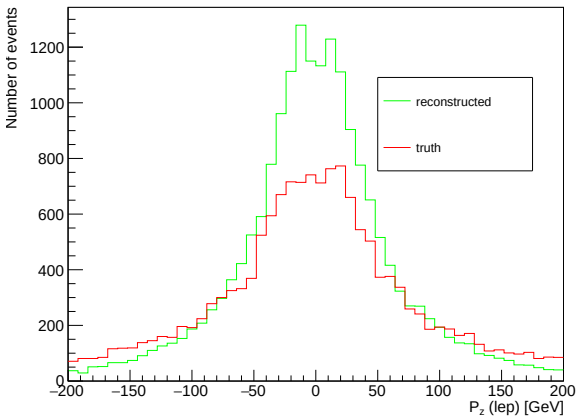
(a) Number of tHq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 32 (all campaigns) have been used.



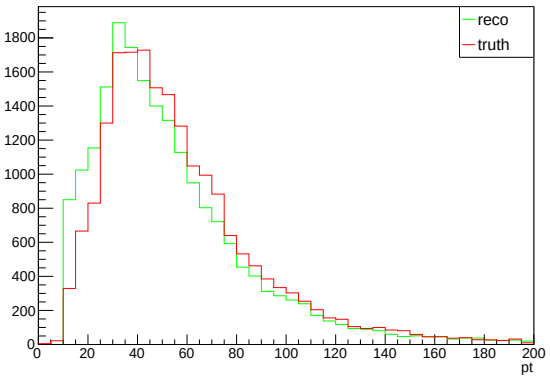
(b) Number of tZq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 29 (campaign mc16a) have been used. Events with anti-top quarks in the final state insteaf of top quarks have been removed. Taken from [77].



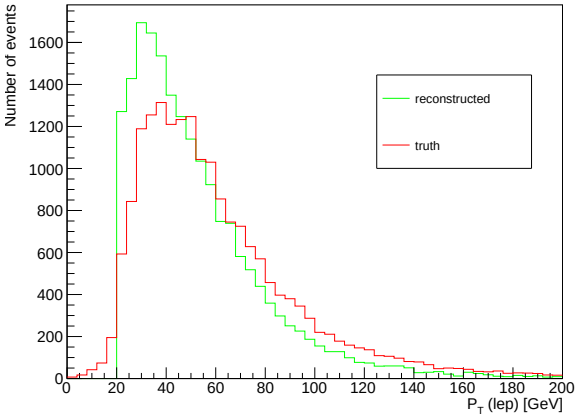
(a) Number of tHq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 32 (all campaigns) have been used.



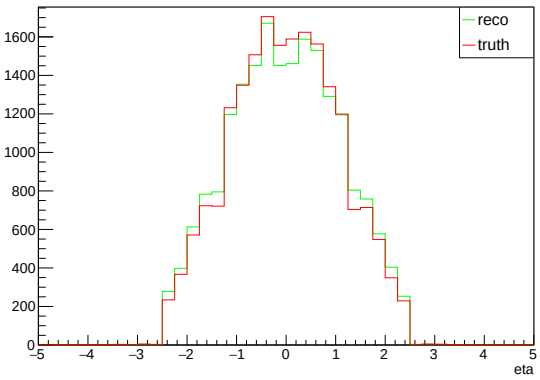
(b) Number of tZq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 29 (campaign mc16a) have been used. Events with anti-top quarks in the final state insteaf of top quarks have been removed. Taken from [77].



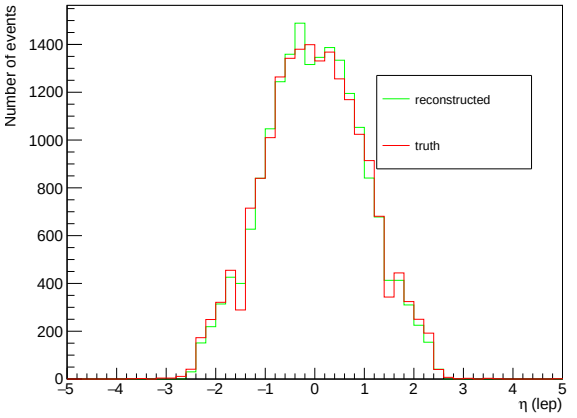
(a) Number of tHq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 32 (all campaigns) have been used.



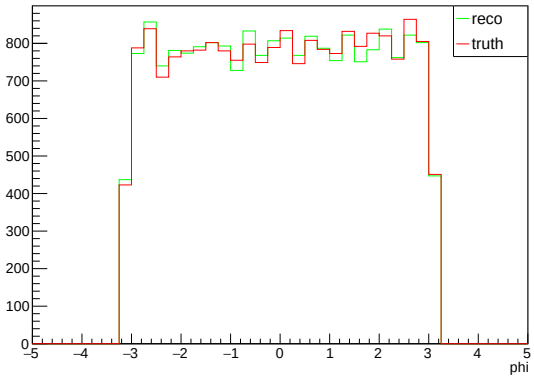
(b) Number of tZq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 29 (campaign mc16a) have been used. Events with anti-top quarks in the final state insteaf of top quarks have been removed. Taken from [77].



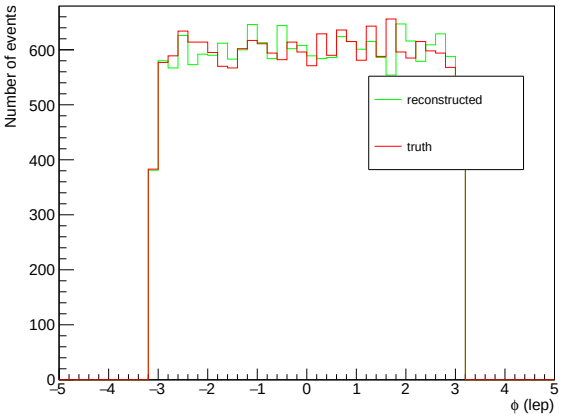
(a) Number of tHq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 32 (all campaigns) have been used.



(b) Number of tZq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 29 (campaign mc16a) have been used. Events with anti-top quarks in the final state insteaf of top quarks have been removed. Taken from [77].



(a) Number of tHq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 32 (all campaigns) have been used.



(b) Number of tZq events as a function of kinematic variables of the lepton originating from the top-quark decay, both on *reco-level* (green) and *truth-level* (red). *OS* events of `SingleTop Ntuples` of version 29 (campaign mc16a) have been used. Events with anti-top quarks in the final state insteaf of top quarks have been removed. Taken from [77].

APPENDIX C

Fit curve from *new* p^T*-fit* with the correlation of all events

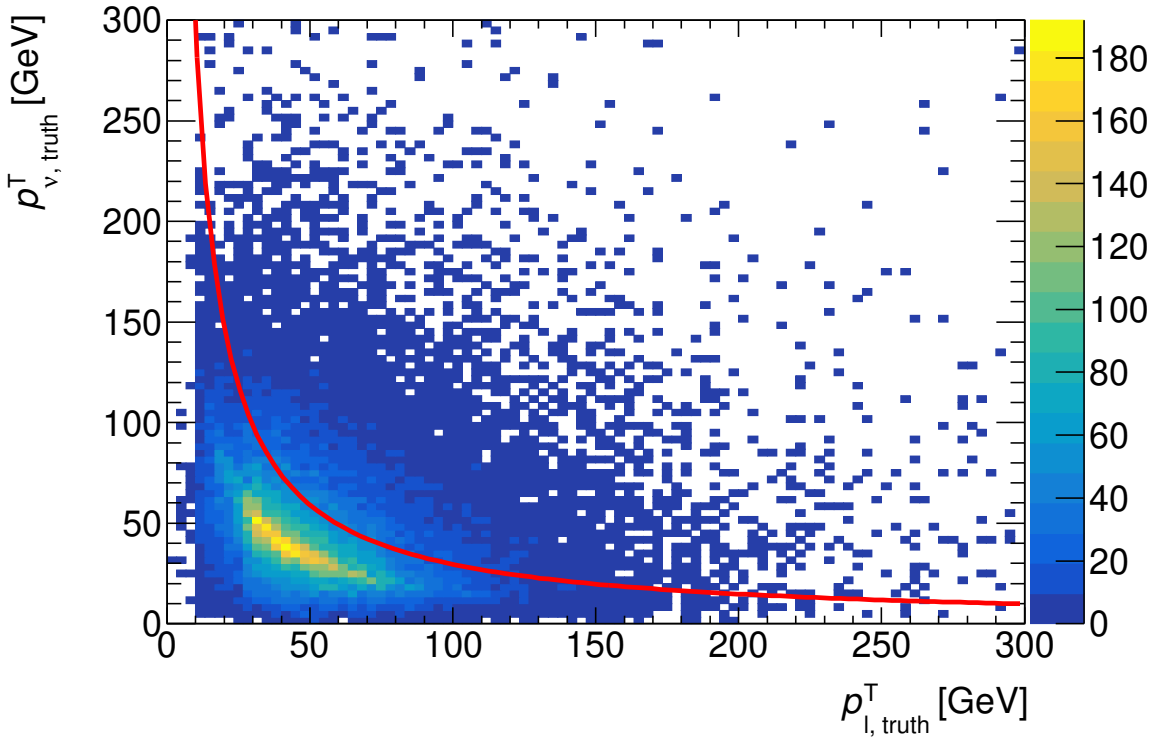


Figure C.1: Correlation of p_ν^{T} with p_l^{T} on *truth-level*. The red curve shows the function of the fit of the model $p_\nu^{\mathrm{T}} = \frac{f}{p_l^{\mathrm{T}}}$ applied to the events fulfilling $|\phi_{\nu,\,\mathrm{pred}} - \phi_{\nu,\,\mathrm{truth}}| < 0.1\,\mathrm{rad}$ with the result of the fitparameter of $f = (2\,955.14 \pm 90.47)\,\mathrm{GeV}^2$. It illustrates that the fit applied to the events fulfilling $|\phi_{\nu,\,\mathrm{pred}} - \phi_{\nu,\,\mathrm{truth}}| < 0.1\,\mathrm{rad}$ does not fit the set of all events.

List of Figures

List of Tables