Study Towards Analysing the Higgs Self-Coupling in the $hh \to 4\tau$ Channel at the ATLAS Detector

Henrik Junkerkalefeld

Masterarbeit in Physik
angefertigt im Physikalischen Institut

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

Juni 2019

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

Angenommen am: 07.06.2019
1. Gutachter: Prof. Dr. Klaus Desch
2. Gutachter: Prof. Dr. Jochen Dingfelder

Abstract

The Higgs boson discovery at the LHC in 2012 completes the predicted particle content of the Standard Model (SM), leaving not much space for unknown physics in the accessibly TeV scale. This heralds the precision measurement era of SM parameters of which the Higgs self-coupling is one of the last unmeasured ones. Its high sensitivity to new physics scenarios further increases its relevance. Due to the very small cross sections, the power of the LHC to measure the self-coupling is estimated to be limited, which motivates the investigation of this process in as many channels as possible.

Apart of more intensively studied decay channels, the $hh \to 4\tau$ channel shapes up as a very promising additional decay, because a very limited amount of other processes is expected to produce four truth taus. However, a large background containing several sources of QCD jets can fake hadronically decaying τ leptons and has to be fought.

This work studies the conspicuity of di-Higgs production within multi-tau events. As the signal efficiency is expected to be low due to the high amount of demanded τ leptons, a novel reconstruction and identification method is developed. Instead of reconstructing every τ lepton candidate on its own, the whole event topology is studied to separate processes with many truth taus from fake enriched ones. It is shown that the new method enables either a doubling of the signal significance or of the signal efficiency, while the corresponding other value remains at its value when using classical reconstruction methods.

All relevant background processes are included in this thesis. To that effort, a sophisticated data-driven fake estimation is developed for four objects. This method is aligned to the used tau combination method, although both approaches are based on the lepton's charges. Finally, three different cut-based signal regions are designed by demanding either none, one or two light leptons together with the corresponding amount of hadronically decaying taus ($4h0\ell$, $3h1\ell$ and $2h2\ell$). The signal selection is further improved by exploiting boosted decision trees.

Overall, a combined significance of $\sigma = 0.89$ for an integrated luminosity of $3\,\mathrm{ab}^{-1}$ is achieved. This value is comparable to the expected significance of the widely studied $hh \to 4b$ channel of $\sigma = 0.6$, although for the $hh \to b\bar{b}\tau\tau$ ($\sigma = 2.1$) and $hh \to b\bar{b}\gamma\gamma$ ($\sigma = 2.0$) channels even higher values are expected. Eventually, an upper limit of the production cross section of $\sigma_{\mathrm{ggF}}/\sigma_{\mathrm{ggF}}^{\mathrm{SM}} = 2.40^{+0.95}_{-0.61}$ is predicted for a dataset of $3\,\mathrm{ab}^{-1}$ at 95 % confidence level. To compare this result to published channels, the limit is recalculated to be $\sigma_{\mathrm{ggF}}/\sigma_{\mathrm{ggF}}^{\mathrm{SM}} = 36.9^{+22.5}_{-13.5}$ for $36.1\,\mathrm{fb}^{-1}$. This value outperforms the upper limits of published decay channels containing W bosons ($\sigma_{\mathrm{ggF}}/\sigma_{\mathrm{ggF}}^{\mathrm{SM}} = 120 - 300$) by far. Instead, it is more comparable to the exclusion capability of the previously named most promising decay channels ($\sigma_{\mathrm{ggF}}/\sigma_{\mathrm{ggF}}^{\mathrm{SM}} = 12 - 25$).

Hence, it is expected that the studied $hh \to 4\tau$ channel is capable to valuably contribute to the overall measurement of the Higgs self-coupling at the ATLAS detector and it should be investigated further in future.

Acknowledgements

I would like to thank Klaus Desch both for the opportunity to write my Master thesis in his research group and for all his helpful ideas and motivating comments. Moreover, I would like to thank Philip Bechtle and Christian Grefe for their continuous support, the inspiring discussions and their expertise in general and on specific problems of this thesis.

The whole research group provided a warm-hearted atmosphere and a lot of daily support. Especially Michael Hübner, Lara Schildgen and Tobias Klingl have to be named, but also Peter Wagner, Alina Manthei, Philipp König and Martin Werres were always helpful and supportive. Thank you.

Additionally to the named people, I would like to thank Christian Hölzer, Philipp Höllmer and Dominik Köhler for proof-reading my thesis. Last but not least, my biggest thanks go to my family and to Rebekka for their invaluable support and encouragement in any situation of my life.

Contents

CHAPTER 1

Introduction

In particle physics, the question of the origin of the universe and its content is investigated at the microscopic scale. By performing experiments at very high energies, the fundamental building blocks of the universe are studied in detail and conditions temporarily close to the big bang are created and probed.

In the 1960s, the Standard Model (SM) of particle physics was developed, which summarises the knowledge of this field in three fundamental symmetries (cf. chapter 2). Most phenomena and particle physics processes are described correctly over several orders of magnitude and it has been tested to very high precision. Several particles were predicted successfully. In 2012 the discovery of the Higgs boson completed the predicted particle content of the SM, leaving not much space for unknown physics at the accessible TeV scale of the Large Hadron Collider (LHC).

Nevertheless, the SM cannot serve as a final theory of the whole universe. Gravitation, which dominates the universe at large scales, is not included in the theory. Moreover, not even 5 % of the universe's energy density were found to be covered by the SM while for the remaining 95 %, divided into Dark Matter and Dark Energy, no explanation is given.

Hence, the completion of the SM's particle content without further clear hints to new physics heralds the precision measurement era of SM parameters, of which the Higgs self-coupling is one of the last unmeasured ones. Its high sensitivity to new physics scenarios further increases its relevance. Particle physics phenomena at the high energy end are probed by the LHC, which is the largest and most powerful particle accelerator worldwide. Protons are collided at centre-of-mass energies of 13 TeV and their collision products are recorded by the ATLAS detector (cf. chapter 3).

Due to the very small cross section of di-Higgs events, the power of the LHC to measure the self-coupling is estimated to be limited. This motivates the investigation of the named process in as many channels as possible. Apart from more intensively studied decay channels, the $hh \to 4\tau$ channel shapes up as a very promising additional decay because a very limited amount of other processes is expected to produce four truth taus (cf. chapter 4). However, a large background containing several sources of QCD jets can fake hadronically decaying τ leptons and must be fought.

This work studies the conspicuity of di-Higgs production within multi-tau events. For this purpose, a novel method to increase the efficiency of detecting events with many truth taus and separating them from fake taus is developed (cf. chapter 5). Moreover, the dominant background processes are studied in detail and their separation from signal processes is investigated (cf. chapter 6).

An essential task is the estimation of QCD multijet processes with a data-driven method for four objects. The $hh \rightarrow 4\tau$ process is separated from background with significance scans and boosted decision trees are utilised to further increase the separation. The contribution of the $hh \rightarrow 4\tau$ process to a combined search for di-Higgs processes after the whole data-taking period of the LHC up to 2038 is estimated. Moreover, an upper limit on the di-Higgs cross section is estimated and compared with published results of other channels.

CHAPTER 2

The Standard Model of Particle Physics

The theoretical framework of every particle physicist is the Standard Model of Particle Physics (SM) [1–5]. An outlook of its basic concept and the incorporated particles and interactions is given in section 2.1, before its mathematical fundament as a quantum field theory is briefly introduced in section 2.2. Very important for the consistency of the SM as well as for this thesis is the Higgs mechanism which is explained in section 2.3. The main properties of the emerging Higgs boson are described in this section and the Higgs self-coupling is introduced. Last but not least, shortcomings of the SM as a motivation for this work are summarised in section 2.4.

2.1 Particles and Forces

The Standard Model integrates the current knowledge of particle physics into one theory and was developed in the 1960s. It contains all known elementary particles, which make up the matter of the universe and describes their dominant interactions via so called gauge bosons. Apart of the gravitational force, which is rather weak on microscopic scale, the SM contains three of the four known interactions of the universe. Namely, these are the electromagnetic force, the weak force and the strong force.

It was successfully tested numerous times with very high precision and predicted the existence of several particles like gluons, the charm and the top quark as well as the W and Z boson before their discovery [6]. In 2012, the SM was completed by the discovery of the Higgs boson at the LHC [7–9]. The elegance of the SM is indicated by the possibility to give a graphical overview of its particle content, their interactions and and their main properties within one picture, which is presented in figure 2.1.

The matter particles are coloured in purple and green. They share the same spin $S = \frac{1}{2}$ which makes them fermions. Each fermion has a corresponding anti partner which has exactly opposite quantum numbers (e.g. like the electric charge) but shares the mass of the fermion. These antiparticles theoretically form antimatter. Due to energy conservation, matter can only be created if the same amount of antimatter is generated within the SM. Moreover in figure 2.1, the gauge bosons are presented in red. They mediate the different forces of the SM (cf. section 2.2) and feature a whole-number-spin of $S = 1$. Last but not least, the Higgs boson ($S = 0$) is depicted in yellow. Its underlying mechanism is responsible for the particles masses and is explained in section 2.3 in more detail.

The fermions are organised in three generations of four particles each (one column). Between the generations, the properties of each particle type are replicated. Only the mass of the charged fermions increases in each generation. Hereby, the masses are distributed over several orders of magnitude from the electron mass of 511 keV to the top mass of 173 GeV [11]. As it is energetically favoured for heavy particles to decay into lighter ones, only the first generation of particles is stable. Thus, in today's cooled

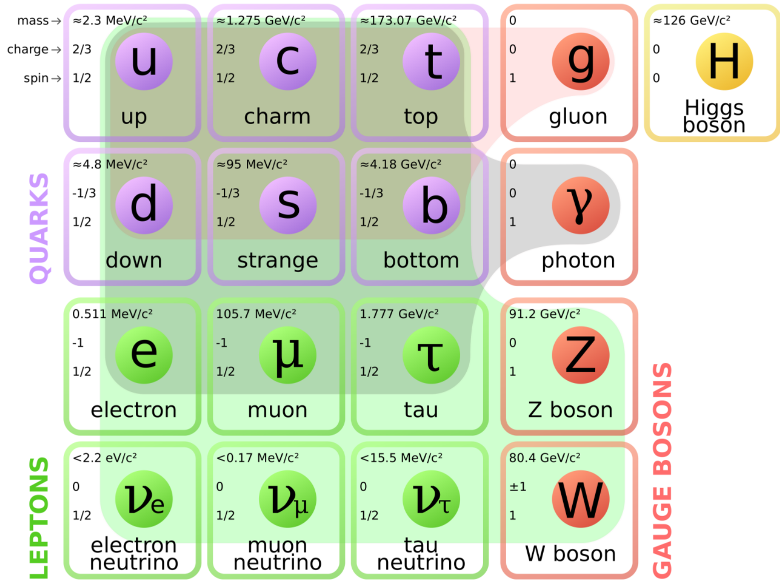


Figure 2.1: The elementary particles of the Standard model are depicted. Each particle type has its own colour and their background is dyed depending on the fundamental forces which are able to act on the specific particle. Further explanations are given in the text. (Figure taken from [10])

down universe only particles of the first generation (and neutrinos) exist in great numbers and all existing matter is built out of them. Particles of the second and third generation are only produced in high energy processes and experiments and are short-lived.

Each generation contains two leptons which are arranged in a doublet. They can be divided into the uncharged neutrino (ν_e, ν_μ and ν_τ), which only interacts via the weak force, and the charged lepton (electron, muon and tau), which also interacts electromagnetically.

The weak force is mediated by massive gauge bosons, namely the W boson and the Z boson and was originally first observed in radioactive decays. Due to the high mass of its mediator bosons, the distance of the weak force is roughly limited to $\mathcal{O}(1\text{ fm})$. The mediator of the electromagnetic force is the photon γ. It is massless which explains the long-range effect of this force. Due to its range and its relative strength (e.g. compared to the weak force) the electromagnetic force forms bound states of electrons and nuclei (as known as atoms) and of multiple atoms (which then form molecules). Moreover, electromagnetic fields are described by this force. Hence, it is the most apparent force of the SM in everyday life.

Additionally, each generation contains a quark doublet in which the up type quark (up, charm and top quark) has an electric charge of $q = +\frac{2}{3}$ while the down type quark (down, strange and bottom quark) has a charge of $q = -\frac{1}{3}$. All quarks also carry a so called colour charge and thus interact via the strong force. Due to their additional electric and weak charge, they also interact via the weak and the electromagnetic force. The strong interaction is mediated by the gluons g, which are colour charged on their own and thus interact with each other. This explains the short range of the strong force in contrast to the masslessness of the gluons.

Due to the properties of the strong force, which are explained in more detail in section 2.2, only colour-neutral particles (hadrons) are found in nature. These colour-neutral hadrons can further be divided into two groups depending on their ingredients. First, mesons contain a colour-charged quark and an anti-quark with corresponding anti-colour. Second, a colour-neutral state can be formed by adding all three different colour charges (referred to as red, blue and green). This analogy to primary colours actually is the reason for the strong charge to be named "colour" charge. Such three-quark objects are called baryons. The most prominent examples of baryons are protons and neutrons which form atomic nuclei and agglutinate despite the accumulating positive charge of the protons due to long range effects of the strong force.[1]

2.2 Theoretical Description of the Standard Model via Quantum Field Theories

Quantum field theories were developed as a solution of connecting quantum mechanics and special relativity. As a first step, this connection lead to the famous Dirac equation in which the free motion of a fermionic particle is described (and the Klein-Gordon equation for the motion of scalar particles (spin $S = 0$) respectively). From here, by demanding Lorentz invariance and local U(1) gauge invariance, the Lagrangian of **Quantum Electrodynamics (QED)** could be constructed which is given by

$$\mathcal{L}_{\mathrm{QED}} = \overline{\Psi}_f(i\gamma^\mu\partial_\mu - m_f)\Psi_f + q_f\overline{\Psi}_f\gamma^\mu\Psi_f A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.1}$$

Physical differential equations of motion of an arbitrary field Φ_i (like Ψ_f or A_μ) can be deduced by using the principle of least action leading to the Euler-Lagrange equation

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_i)}\right) - \frac{\partial\mathcal{L}}{\partial\Phi_i} = 0 \tag{2.2}$$

Hereby, in equation (2.1) the first term leads to the Dirac equation and describes a freely propagating fermionic field Ψ_f like the electron or the muon with a mass m_f. The third term describes a free, massless vector field A^μ with $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$, which can be identified as the photon field. The second term describes the interaction between photons and fermions. This interaction can be drafted in a so called Feynman diagram consisting of the photon-fermion vertex as shown in figure 2.2. Interestingly, the second term had to be added to fulfil local U(1) gauge invariance, thus electromagnetic interactions are a direct consequence of this fundamental symmetry [12, 13].

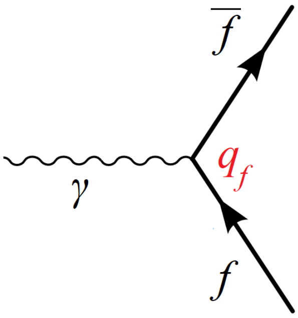


Figure 2.2: The photon-fermion vertex emerging from the term $\mathcal{L}_{\mathrm{QED}} \supset q_f\overline{\Psi}_f\gamma^\mu\Psi_f A_\mu$ is shown. The vertex consists of three interacting particles: The fermion f and antifermion $\bar{f}$ take part due to the $\overline{\Psi}_f\gamma^\mu\Psi$ part of the term and the photon γ is represented in the term by A_μ. q_f denotes the strength of the coupling.

In front of the interaction term, the pre-factor q_f denotes the strength of the coupling of each fermion

[1] Nevertheless, these “long range effects” are limited to the distance of atomic nuclei due to the mentioned self interaction of gluons. The attraction of protons and neutrons to each other is often called the “van-der-Waals force” of the strong force to underline the analogy to the eponymous force as a long distance effect of electrically neutral molecules and atoms.

to the photon and can directly be identified as the electric charge of the particular fermion. Usually, the coupling constant $\alpha_{\text{em}} = \frac{e^2}{4\pi}$, with e being the elementary electric charge, is used to quantify the strength of the electromagnetic force. In opposition to its name, however, the coupling "constant" α_{em} is not constant but depends on the momentum transfer q^μ of a process. By using the Callen-Symanzik equation of the renormalisation group [14, 15], the so called running coupling of QED can be calculated by

$$\alpha_{\text{em}}(q^2) = \frac{\alpha_{\text{em}}(\mu^2)}{1 - \alpha_{\text{em}}(\mu^2)\frac{1}{3\pi}\ln\frac{q^2}{\mu^2}} \tag{2.3}$$

Hereby, μ^2 denotes the renormalisation scale. To keep corrections to that equation small, μ^2 should be chosen of the order of the external momentum scale q^2. For example, for small μ^2 of the order of atomic energy levels ($O(10\,\text{eV})$), α_{em} is approximately $\frac{1}{137}$. For high energies, i.e. high q^2 values or very short distances, the electromagnetic coupling strength grows. Phenomenologically, this can be explained by vacuum polarisation, i.e. virtual electron-positron pairs which shield the particle's electric charge. For higher energies and equivalently shorter distances this shielding effect becomes weaker and the particle's effective charge grows. For example, at the energy scale of the Z boson mass (91.19 GeV), α_{em} becomes approximately $\frac{1}{129}$.

Quantum Chromodynamics (QCD) [5] is the theoretical framework of the strong force and can be constructed similarly to QED. The corresponding gauge transformation of QCD is a SU(3) gauge group and the QCD Lagrangian reads as

$$\mathcal{L}_{\text{QCD}} = \overline{\Psi}_q(i\gamma^\mu\partial_\mu - m_q)\Psi_q + g_s\overline{\Psi}_q\gamma^\mu T_a\Psi_q G^a_\mu - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a \tag{2.4}$$

$\overline{\Psi}_q$ and Ψ_q present the quark fields, g_s is the coupling strength of the strong force, T_a is the generator of the SU(3) group and G^a_μ can be identified as the massless gluon field, the mediator of the strong force. The similarity to the QED Lagrangian (eq. (2.1)) is glaring. In fact, any SU(N) gauge theory can be expressed with such a so called Yang-Mills Lagrangian [16] with more general couplings, generators and gauge bosons.

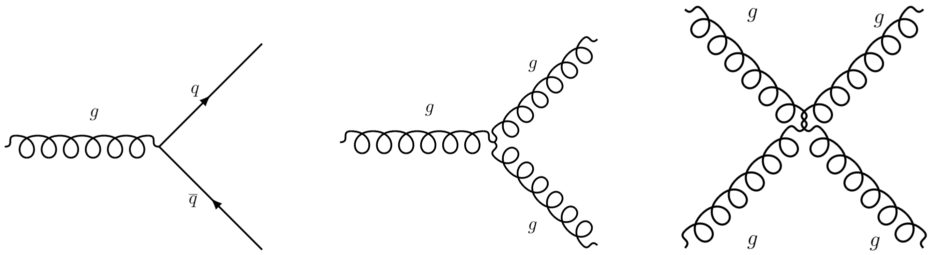
The very different phenomenological behaviour of the strong and the electromagnetic interaction can be traced back to the fact, that the SU(3) group is non-abelian. Hence, the bosonic part of the Lagrangian has to be generalised such that the Lagrangian stays gauge invariant and still transforms covariantly. The generalised bosonic field-strength tensor is $G^{\mu\nu}_a = \partial^\mu G^\nu_a - \partial^\nu G^\mu_a - g_s f_{abc} G^\mu_b G^\nu_c$.[2] f_{abc} is the completely antisymmetric structure constant of the corresponding symmetry group and equal to zero for any abelian U(1) group, like QED. In the SU(3) case of QCD, however, the term does not vanish and induces self interacting vertices of the gluons. In summary, in a very simplified way, the Lagrangian can be describes as

$$\mathcal{L}_{\text{QCD}} = \text{``}\overline{\Psi}_q\Psi_q\text{''} + \text{``}g_s\overline{\Psi}_q\Psi_q G\text{''} + \text{``}GG\text{''} + \text{``}g_s GGG\text{''} + \text{``}g_s^2 GGGG\text{''} \tag{2.5}$$

which additionally to the Feynman diagrams of QED gives rise to vertices with three or four gluons (cf. figure 2.3).

A direct consequence of these self interacting vertices is an altered running of the strong coupling. Precisely, the three-gluon vertex causes a shift in sign in the Callen-Symanzik equation such that the

[2] This holds for any SU(N) gauge theory. The structure constant is defined with the commutator of the group's generators: $[T_a, T_b] = if\text{abc}T_c$.



(a) "$g_s\overline{\Psi}_q\Psi_q G$" (b) "$g_s GGG$" (c) "$g_s^2 GGGG$"

Figure 2.3: The QCD interaction vertices are shown. They are directly represented in the Lagrangian by the underlying terms.

running of the QCD coupling α_s can be described as

$$\alpha_s(q^2) = \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2)\frac{11N_C - 2N_f}{12\pi}\ln\frac{q^2}{\mu^2}} \tag{2.6}$$

in which N_C is the number of colours and N_f is the number of strongly interacting fermions (quarks in the SM framework) which are accessible at energies of q^2. As a result of this altered coupling, α_s decreases with increasing momentum scale or shorter length scales (cf. figure 2.4). This is referred to as the asymptotic freedom of quarks at high energies. Additionally, this means that for low momenta or larger distances, the coupling strength grows significantly (hence the name "strong" interaction). Thus, when separating two colour-charged quarks, the strong coupling creates an increasingly strong force between these quarks. This binding energy eventually surpasses the threshold to produce additional quark-antiquark pairs which then form colour-neutral binding states with the primary quarks. This process is referred to as hadronisation in collider experiments. Consequently, no coloured particles can be observed freely in nature but only in colour-neutral binding states ("confinement") and the range of the strong interaction is very limited [12, 17].

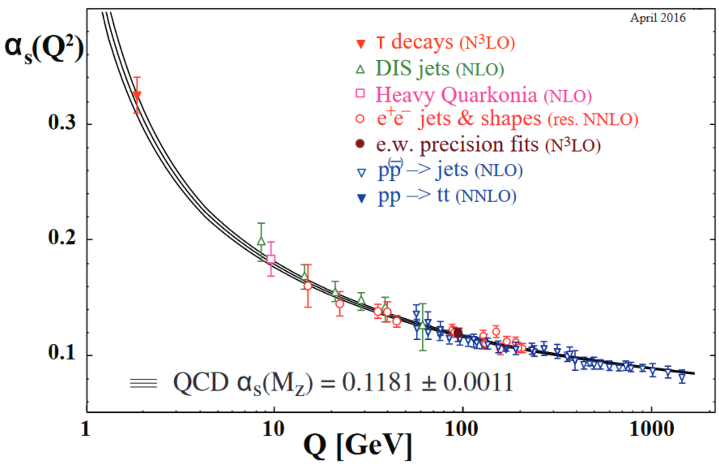


Figure 2.4: The measurement of the strong coupling strength α_s by various experiments is shown. Its running from very high values at low energies (confinement) to comparably lower values for high energies (asymptotic freedom) is clearly observable. Source: [11]

The third and last fundamental interaction within the SM is the **weak interaction**. For several reasons, this interaction is unique in the SM. As the name already suggests, the effective coupling strength of this force is much weaker than the ones of the electromagnetic and strong interaction. Hence, it is the only interaction which does not form binding states in nature and mainly shows up in radioactive β-decays in everyday life.

Additionally, it is the only force which is maximally parity ($\mathcal{P}$) violating [18, 19] and couples only to left-handed particles (or right-handed antiparticles). Moreover, it is the only interaction with massive force-mediating gauge bosons, the W and the Z boson. Actually, the weakness of the interaction is a direct consequence of the gauge boson masses as the effective coupling scales with $\frac{g_{\mathrm{w}}^2}{m_{W/Z}^2}$. The weak coupling constant $\alpha_{\mathrm{w}} = \frac{g_{\mathrm{w}}^2}{4\pi}$ itself amounts to roughly $\frac{1}{29.5}$ at low energies and hence would be larger than the electromagnetic coupling. The resulting weakness of the force leads to much higher lifetimes of particles that decay only via the weak interaction compared to particles that decay electromagnetically or strongly.

Furthermore, transitions between fermion generations are only possible via the charged current (exchange of a W boson) of the weak force. Neutral currents (exchange of a Z boson) are the only possibility of neutrinos to participate in scattering processes.

In the 1960s, Glashow, Weinberg and Salam were able to combine the electromagnetic and the weak interaction into an overlaying electroweak interaction [1, 3, 4]. The symmetry group of this electroweak unification is represented by $\mathrm{SU}(2)_{\mathrm{L}} \times \mathrm{U}(1)_{\mathrm{Y}}$ in which all left-handed particles are grouped into doublets $(u_{\mathrm{L}}, d_{\mathrm{L}})$ or $(\nu_{\mathrm{L}}, \ell_{\mathrm{L}})$ and carry a weak isospin $T = \frac{1}{2}$ due to the local $\mathrm{SU}(2)_{\mathrm{L}}$ symmetry. Additionally, every particle carries a so called hypercharge Y due to the local $\mathrm{U}(1)_{\mathrm{Y}}$ symmetry. In this picture, the electric charge q can be deduced by $q = T_3 + \frac{Y}{2}$ with T_3 being the third component of the weak isospin.

The observed gauge bosons γ, Z and $W^{\pm}$ are mixtures of the B^{μ} (the gauge boson of the $\mathrm{U}(1)_{\mathrm{Y}}$ symmetry) and the three gauge bosons of the $\mathrm{SU}(2)_{\mathrm{L}}$ symmetry, the W_1^{μ}, W_2^{μ} and W_3^{μ} via the weak mixing angle θ_{w} [12, 17]:

$$\begin{pmatrix} A^{\mu} \\ Z^{\mu} \end{pmatrix} = \begin{pmatrix} \cos\theta_{\mathrm{w}} & \sin\theta_{\mathrm{w}} \\ -\sin\theta_{\mathrm{w}} & \cos\theta_{\mathrm{w}} \end{pmatrix} \cdot \begin{pmatrix} B^{\mu} \\ W_3^{\mu} \end{pmatrix} \qquad \text{and} \qquad W^{\mu,\pm} = \frac{1}{\sqrt{2}} \left(W_1^{\mu} \mp i W_2^{\mu} \right) \tag{2.7}$$

Now, nifty readers might object that massive gauge bosons contradict the required gauge invariance of our Yang-Mills theory as e.g. a term like $\frac{1}{2} m_{\gamma}^2 A_{\mu} A^{\mu}$ clearly is not invariant under the local U(1) gauge transformation $A_{\mu} \to A'_{\mu} = A_{\mu} - \partial_{\mu}\chi$. The same statement holds for local $\mathrm{SU}(2)_{\mathrm{L}}$ and $\mathrm{SU}(3)_{\mathrm{C}}$ transformations which is in strong tension to the unambiguously observed masses of the W and Z boson. Moreover, as stated before, left-handed particles transform as weak isospin doublets and right-handed particles as singlets under the local $\mathrm{SU}(2)_{\mathrm{L}}$ gauge transformation. However, the fermionic mass term $-m_{\Psi}\overline{\Psi}\Psi$ (see eq. (2.1) or eq. (2.4)) can also be expressed by $-m_{\Psi}(\overline{\Psi}_{\mathrm{R}}\Psi_{\mathrm{L}} + \overline{\Psi}_{\mathrm{L}}\Psi_{\mathrm{R}})$ and thus couples left-handed and right-handed particle states. Hence, this term also breaks the required gauge invariance and the electroweak theory would not be renormalisable as shown by t'Hooft [20–22].

As a solution of this issue, the Higgs mechanism was developed and is introduced in the following section.

2.3 The Higgs Mechanism and Electroweak Symmetry Breaking

In the Standard Model, the massive gauge bosons as well as the fundamental fermions acquire their mass without breaking the local gauge symmetries of the SM due to the Higgs mechanism [23, 24]. Here, the corresponding mass terms are not added directly to the Lagrangian but they are generated through the interaction of the particles with the Higgs field and a spontaneous symmetry breaking of the $SU(2)_L \times U(1)_Y$ symmetry. A brief overview of the most important steps is given in this section. More detailed calculations can be found in refs. [12, 17].

The simplest choice to gain as many degrees of freedom as needed to give mass to three gauge bosons is to introduce two complex scalar fields embedded in a weak isospin doublet:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.8}$$

The possible additional terms arising in the Lagrangian for the scalar fields, which are invariant under $SU(2)_L$ and $U(1)_Y$ gauge transformations are

$$\mathcal{L}_{\text{Higgs}} \supset (D_\mu\Phi)^\dagger(D^\mu\Phi) - g_f\left(\bar{L}\Phi R + \bar{R}\Phi^\dagger L\right) - V_{\text{Higgs}}(\Phi) \quad . \tag{2.9}$$

Here, D_μ is the covariant derivative $D_\mu = \partial_\mu + i\frac{g_w}{2}\vec{\sigma} \cdot \vec{W}_\mu + i\frac{g'}{2}Y_\Phi B_\mu$ (note that it actually is a 2×2 matrix) with the hypercharge of the Higgs doublet of $Y_\Phi = 1$. L is a left-handed lepton or quark doublet $L = (f_L^u, f_L^d)^\dagger$ and R are the corresponding right-handed singlets.

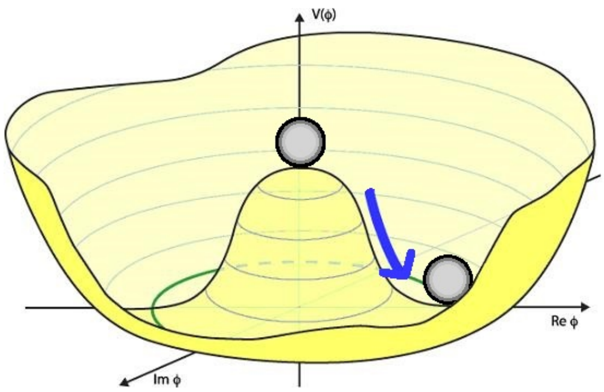


Figure 2.5: The Higgs potential is sketched. Apparently, its minima are unequal to zero. Source: [25]

The Higgs potential is given by

$$V_{\text{Higgs}}(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2 \tag{2.10}$$

which has very interesting properties if the μ^2 term becomes negative. In that case, the vacuum state of the field Φ or correspondingly the minimum of $V_{\text{Higgs}}(\Phi)$ is no longer at $|\Phi| = 0$ but satisfies the equation $|\Phi| = \sqrt{-\frac{\mu^2}{2\lambda}} := \frac{v}{\sqrt{2}}$. The field moves to its minimal value (cf. figure 2.5) and hereby gains a so called vacuum expectation value v. The original symmetry of the Lagrangian is broken and thus this process is referred to as *spontaneous symmetry breaking*. Quantum fluctuations of the field around the new minimal value describe a particle state which can be identified as the SM Higgs boson. Without loss of generality, the vacuum state can be chosen to be in the real direction and directly choosing the *Unitary gauge* for the excited field, after symmetry breaking the Higgs doublet can be written as

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad . \tag{2.11}$$

The Higgs boson is hereby represented by $h(x)$. Inserting this into the first term of equation (2.9), the term reads as follows:

$$(D_\mu \Phi)^\dagger (D^\mu \Phi) = \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{1}{4} g_\mathrm{w}^2 W_\mu^- W^{\mu,+} (v + h)^2 \tag{2.12}$$

$$+ \frac{1}{8} \begin{pmatrix} B_\mu & W_{\mu,3} \end{pmatrix} \begin{pmatrix} g'^2 & -g_\mathrm{w} g' \\ -g_\mathrm{w} g' & g_\mathrm{w}^2 \end{pmatrix} \begin{pmatrix} B^\mu \\ W_3^\mu \end{pmatrix} (v + h)^2 \quad . \tag{2.13}$$

The terms multiplied by v^2 are responsible for the **gauge boson masses** while the terms multiplied by $2vh$ or h^2 induce gauge boson couplings to the Higgs boson. The W mass term appears as $m_W^2 W_\mu^- W^{\mu,+}$ in the Lagrangian and thus, the W mass can directly be read off to be $m_W = \frac{1}{2} g_\mathrm{w} v$. Moreover, The mixing of W^3 and B appears naturally due to the off-diagonal terms in their mass matrix in term (2.13). After diagonalising the matrix, the term can be written as

$$\frac{1}{8} v^2 \begin{pmatrix} A_\mu & Z_\mu \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & g_\mathrm{w}^2 + g'^2 \end{pmatrix} \begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} \quad . \tag{2.14}$$

By defining the weak mixing angle θ_w to fulfil $\tan \theta_\mathrm{w} = \frac{g'}{g_\mathrm{w}}$, the gauge-boson-mixing matrix as stated in equation (2.7) results. Furthermore, this induces the massless photon and the Z boson mass of $m_Z = \frac{1}{2} \frac{g_\mathrm{w}}{\cos \theta_\mathrm{w}} v$. It is worth to note that a massless vector boson has two degrees of freedom in contrast to a massive vector boson with three degrees of freedom due to the additional longitudinal polarisation. These additional degrees of freedom can be associated to three Goldstone modes of the Higgs which are a consequence of the symmetry breaking. Instead of manifesting physically, they are "absorbed" by the W and Z bosons such that the latter ones gain their masses. After measuring m_W, m_Z and g_w, the vacuum expectation value of the Higgs field is found to be $v = 246\,\mathrm{GeV}$.

For the **fermion masses**, equation (2.11) can be inserted into the second term of equation (2.9). As an example, for the first quark doublet this leads to

$$\mathcal{L} \supset -\frac{g_d}{\sqrt{2}} (\bar{d}_\mathrm{L} d_\mathrm{R} + \bar{d}_\mathrm{R} d_\mathrm{L})(v + h), \tag{2.15}$$

resulting in a down quark mass of $m_d = g_d \frac{v}{\sqrt{2}}$. The up-type quarks gain their masses through the Hermitian conjugate of the Lagrangian.

A few interesting remarks can be made. To begin with, the second term of equation (2.15) induces a coupling of the specific fermion to the Higgs boson. Hence, the coupling of each fermion to the Higgs is expected to be proportional to the measured fermion mass, which can be tested experimentally. Moreover, the Yukawa couplings g_f are not predicted by the Higgs mechanism and have to be measured. Actually, they are the only property which separates the different fermion generations from each other. The reason for their huge variation, which becomes manifest in the very different fermion masses, remains unanswered and somehow unsatisfactory from a theoretical point of view. Especially the very small neutrino masses would need Yukawa couplings of $\lesssim 10^{-12}$, if they are produced via the Higgs mechanism. Also due to this reason, many theoreticians prefer other hypothetical mechanisms like the see-saw mechanism to generate the neutrino masses.[3]

Last but not least, after the Higgs field gained its vacuum expectation value (eq. (2.11)), the Higgs potential (eq. (2.10)) is given by

$$V_{\text{Higgs}}(\Phi) = \lambda v^2 h^2 + \lambda v h^3 + \frac{1}{4}\lambda h^4 - \frac{1}{4}\lambda v^4 \quad . \tag{2.16}$$

Thus, looking at the first term of this equation, a famous prediction of the Higgs mechanism is a massive scalar particle, the Higgs boson. Its discovery in 2012 [7–9] is a huge success of the whole Standard Model and its measured mass of roughly $m_h = 125\,\text{GeV}$ suggests a value of $\lambda = \frac{1}{2}\frac{m_h^2}{v^2} \approx 0.129 \approx \frac{1}{8}$. Hence, all values of the Higgs potential, as predicted by the Standard Model, are known and can be tested experimentally.

Apart of the unphysical constant term, the remaining terms of equation (2.16) induce **Higgs self-couplings**. They are the main motivation of this thesis and thus are rewarded with their own section (2.3.2) after general properties of the SM Higgs boson are summarised in section 2.3.1.

2.3.1 Higgs Production and Decay

As stated in the previous section, the Higgs boson coupling strength to fermions is proportional to their mass. At the LHC, the colliding protons would thus implicate extremely tiny Higgs production rates if the protons solely consisted of their valence quarks (*uud*). Fortunately from that point of view, due to the nature of the strong force, the proton has a much more complex structure and consists of a wealth of quarks, antiquarks and gluons. Their appearance in the proton is described by parton distribution functions (PDFs) which are shown in figure 2.6.

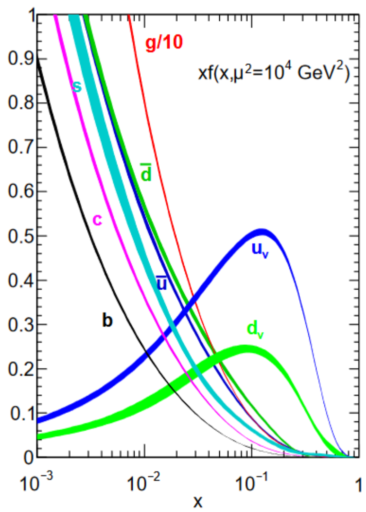
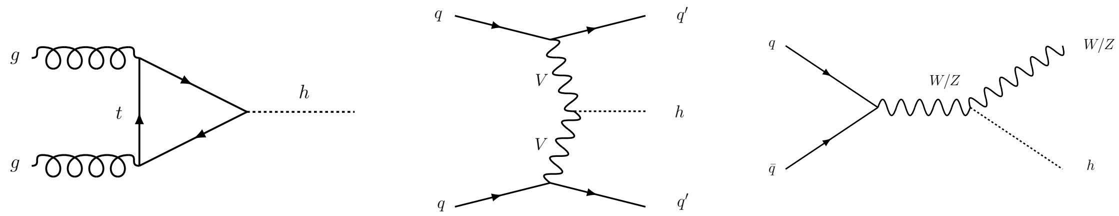


Figure 2.6: The parton distribution functions of the proton as obtained in a NNLO NNPDF3.0 global analysis [29]. The Bjorken x scaling variable can be interpreted as the fraction of the proton momentum which is carried by the particular parton (figure from [11]).

[3] In the Standard Model, the neutrinos do not have a right-handed component as this would be a singlet under all SM gauge transformations. Hence, neutrinos are massless in the SM. In contradiction to that, neutrinos are experimentally favoured to have tiny masses [26–28].

For small fractions of the proton momentum, the gluon PDF dominates by far. Although massless gluons do not couple to the Higgs boson directly, this results in the fact that they participate in the most dominant Higgs production process at the LHC, the so called gluon fusion. This is a one-loop process, in which the gluons couple indirectly via a loop to the Higgs. As top quarks are by far the heaviest particles in the SM, their loop contribution dominates this process (cf. figure 2.7(a)). Other noticeable processes are the Higgs production via vector boson fusion of the electroweak bosons which is depicted in figure 2.7(b) or Higgs radiation (cf. figure 2.7(c)). The cross sections of the latter two processes are more than one order of magnitude smaller such that they will not be taken into account for the di-Higgs signal in this thesis. However, the additional jets might fake τ leptons (cf. chapter 5, 6) such that they might be a scalable background process of this work.



(a) gluon fusion (b) vector boson fusion (c) Higgs radiation

Figure 2.7: The three dominant Higgs production processes at the LHC are shown.

The Higgs decay is also prescribed by its coupling to heavy particles. As a result the most dominant decay is $h \to b\bar{b}$ with 58.2 % as b quarks are the heaviest particles which can be produced pairwise by the Higgs. Although $m_h < 2m_W$, the decay $h \to WW^*$ has the second largest branching fraction of 21.4 %. Here, one W boson is produced off-mass-shell, which suppresses the matrix element. However, the special coupling of the Higgs boson to the massive W leads to the relatively large branching ratio. Most importantly for this thesis, Higgs bosons decay into τ leptons in roughly 6.3 % of all cases (cf. figure 2.8).

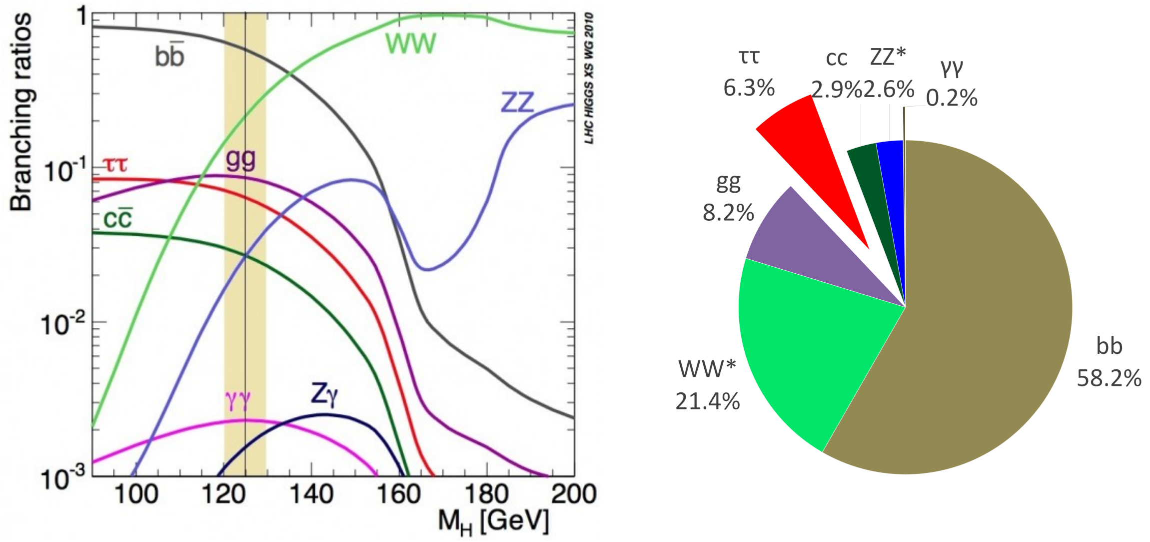
2.3.2 Probing the Higgs Potential with the $hh \to 4\tau$ Channel

With the discovery of the Higgs boson, the Higgs mechanism as described in this section can be called one of the greatest successes of the SM. It sets the course for the theory of electroweak unification with the prediction of the massive W and Z boson. It is not only consistent with electroweak precision fits made by the LEP experiments, also the couplings of fermions to the Higgs seem to agree with their prediction [33–38]. Thus, it is safe to state that the discovered Higgs boson is at least Standard Model like.

Nevertheless, it is worth to pause for a moment and reflect what we really know about the Higgs potential at the moment. The vacuum expectation value v is measured indirectly by different measurements. Moreover, the Higgs mass induces information about the direct slope of the potential around the minimum. But being disarmingly honest, this knowledge is visualised in figure 2.9(a) while further properties of the potential remain unmeasured (cf. figure 2.9(b)).

On top of that, many theories beyond the Standard Model predict a significant alteration of the potential (cf. section 2.4). Hence, it should be probed as precisely as possible.

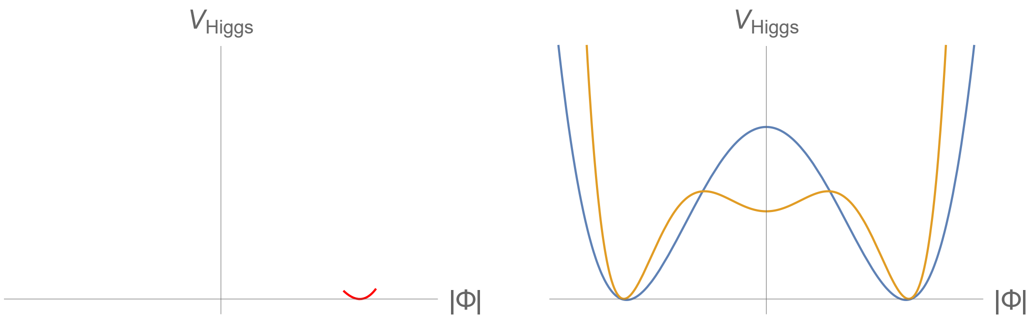
After symmetry breaking as predicted by the Standard Model (eq. (2.16)), the potential contains a cubic and a quartic term with respect to the Higgs field. This induces the phenomenological effect



(a) Figure inspired from [30, 31]

(b) Values taken from [32]

Figure 2.8: Higgs branching ratios in dependence of the Higgs mass (a) and as a pie chart for $m_h = 125.09\,\text{GeV}$ (b). The left plot is historically motivated without the proper knowledge of the Higgs mass. Nevertheless, for educational reasons, this diagram is presented here.



(a)

(b)

Figure 2.9: The current knowledge of the Higgs potential (a) and valid possible extrapolations (b) in a provocative sketch. The possible extrapolations shown here are meant to be educational and should not be taken too seriously. In fact, the slope around the minimum is also known to a certain extend and would have to be fulfilled by the extrapolation. However, the global behaviour of the Higgs potential remains unmeasured.

of either trilinear or quadrilinear Higgs boson self-coupling. While the latter remains inaccessible at the LHC or future linear colliders [39–41], the trilinear Higgs coupling can be probed in multi-Higgs final states. Moreover, its effects on precision observables [42] or loop corrections to single Higgs production [43, 44] can be studied.

This thesis focusses on the multi-Higgs approach and its purpose is to participate in the ongoing efford of measuring di-Higgs production for the first time. Within the SM, di-Higgs production is several orders of magnitude lower than single Higgs production [45] due to additional vertices and the additional on-shell Higgs boson which significantly reduces the kinematic phase space. Moreover, the production suffers from a strong destructive interference of the two leading-order production processes [46], which are presented in figure 2.10.

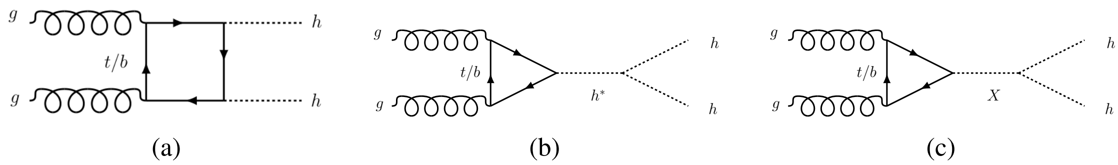


Figure 2.10: The dominating leading order di-Higgs production diagrams are shown. Within the SM, di-Higgs production is possible via a heavy-quark loop (a) and the trilinear Higgs self-coupling (b). Theories beyond the SM further predict possible high mass resonances (c) which decay into two Higgs bosons.

The most interesting diagram for this thesis is the so called triangle diagram (cf. figure 2.10(b)) in which two Higgs bosons are produced via the exchange of another virtual Higgs boson. This is the only tree-level diagram which is sensitive to the Higgs boson self-coupling constant g_{hhh}. Moreover, two Higgs bosons can be produced via a heavy quark loop (cf. figure 2.10(a)) which scales with the tth and bbh coupling strength. Due to their negative interference, the SM prediction of the overall di-Higgs production cross section only accounts for 33.4 fb [32].

This cross section can be enhanced by several scenarios beyond the SM (BSM) which are partly motivated in section 2.4. In these scenarios, either the destructive interference is reduced by a shifted tth, bbh or g_{hhh} coupling, or new hypothetical particles allow for additional production processes. These could either interact via loop corrections such as light coloured scalars [47] or via resonant production if they decay into two Higgs bosons (cf. figure 2.10(c)). For the latter case, several possible particles serve as candidates for this resonant intermediate boson and are mentioned in the following section.

By confining ourselves to the SM-prediction, in summary the overall expected cross section of 33.4 fb causes the observation of di-Higgs production to be a huge challenge as only in roughly one of two trillion (!) proton-proton collisions [48], two Higgs bosons are produced at the same time. Thus, the overall amount of these events is extremely limited and as many further decay channels as possible should be taken into account to detect them.

The most promising decay channels [49, 50] to measure di-Higgs production are thought to be $hh \to b\bar{b}\tau^-\tau^+$ [51], $hh \to b\bar{b}b\bar{b}$ [52] and $hh \to b\bar{b}\gamma\gamma$ [53]. To further improve the detection significance at the ATLAS detector, also studies for the less prominent production channels $hh \to \gamma\gamma WW^*$, $hh \to b\bar{b}WW^*$ and $hh \to WW^*WW^*$ were published [54–56]. They exploit the high branching ratio of Higgs decaying into b quarks or W bosons and partly try to connect that to channels which leave a very unique structure in the detector such as two τ leptons or two photons. However, they all suffer from an extremely large QCD background produced in the hadron collider which mimics the signal of b quarks or hadronically decaying W bosons. Thus, this thesis focusses on an alternative approach and serves as a first estimate of the capability of the **$hh \to \tau^-\tau^+\tau^-\tau^+$ channel**. The drawback of an extremely low branching fraction of roughly 0.4 % might well be overcome by the fact that hardly any other SM process produces four

τ leptons at the same time. Hence, a very low SM background might be possible (cf. section 4.1). A summary of the expected amount of events in each channel after the full LHC data taking period[4] (which is planned to amount to an integrated luminosity of 3 ab^{-1}) is shown in table 2.1.

Channel	Branching Ratio	Expected Events
total	100 %	100 000
$b\bar{b}b\bar{b}$	35.38 %	35 380
$b\bar{b}WW^*$	24.91 %	24.91
$b\bar{b}\tau\tau$	7.28 %	7 280
WW^*WW^*	4.58 %	4 580
$b\bar{b}\gamma\gamma$	0.24 %	240
$\gamma\gamma WW^*$	0.08 %	80
$\tau\tau\tau\tau$	0.38 %	380

Table 2.1: Amount of di-Higgs events expected in selected channels for an integrated luminosity of 3 000 fb^{-1} at the LHC. The branching ratio is calculated by considering the single Higgs BRs (cf. figure 2.8).

Furthermore, it should be emphasised that even by combining the mentioned channels above, only a detection significance corresponding to 3σ is expected. This results in a trilinear Higgs coupling uncertainty of roughly 50 % [49]. Indeed, many BSM theories allow for a strong modification of this coupling and theoretically, even a six times enhanced trilinear coupling agrees well with all measured electroweak properties at the moment [57]. As a result, the highest achievable accuracy is desirable to exploit the self-coupling's full probing power of BSM physics. In summary, every appreciable contribution of the $hh \to 4\tau$ channel to the overall di-Higgs coupling sensitivity could be assessed as a success.

2.4 Shortcomings of the Standard Model and Their Connection to Scalar Fields

Despite to the tremendous successes of the Standard Model and its enormous precision in correctly describing microscopic processes, there are still open questions which cannot be answered by the SM alone.

First of all, gravity is not included in the SM. Instead, Einstein's theory of General Relativity from 1915 [58] is still the most recent and successful theory to describe macroscopic phenomena which are completely gravity-dominated. Since then, several attempts have been made to integrate General Relativity in an overlying quantum field theory. None of them succeeded and in fact, gravity seems to be non-renormalisable and thus incompatible with the QFT framework of the SM.

Although these two theories seem to be well separated into the microscopic and macroscopic cosmos, there are scenarios in the universe, in which the energy-momentum-density is high enough for gravity to compete with the other fundamental forces with respect to the coupling strength. Examples for this are the very early universe or black holes which both need a unified theory of QFTs and gravity, referred to as quantum-gravity [59].[5]

Moreover, the probably most glaring caveat of the SM also has to do with macroscopic scales. Astronomical observations discovered that only 4.96 % of the universe are made out of baryonic matter which is

[4] this includes the high-luminosity upgrade of the LHC which is planned to start taking data in 2026

[5] In fact, these are also the two examples in which General Relativity on its own runs into mathematical singularities. Thus, the examples can also be seen as general problems of the gravitational theory.

described by the SM. Another 26.54 % are made out of so called Dark Matter [60]. This arcane substance only reveals itself through gravitational interaction like in rotational curves of galaxies, in gravitational lensing observations or in its role in the structure forming of the early universe. Possible candidates for Dark Matter might be unknown particles which only interact weakly (i.e. with the weak interaction or a yet unknown new interaction) with the SM. The remaining 68.5 % of the universe are made out of the even more mysterious Dark Energy. The energy is responsible for the observed accelerated expansion of the universe and manifests itself as a cosmological constant in Einstein's equations of gravity. Its physical nature is completely unknown.

Astronomical observations also indicate that the baryonic matter of the universe seems to be only build out of matter and not antimatter. However, within the SM, it is ensured that matter can only be formed by creating an equal amount of antimatter. Therefore, the question arises why all antimatter disappeared which is addressed as the matter-antimatter asymmetry problem. It might be possible that such an asymmetry occurs under specific circumstances. They include deviations from thermal equilibrium such as the moment when the Higgs field developed its vacuum expectation value and broke electroweak symmetry. This phase transition, though, needs to be of first order to serve as a sufficiently strong deviation from equilibrium. However, this could only be incorporated into the SM Higgs mechanism if the Higgs mass would be $m_h \lesssim 70\,\mathrm{GeV}$ which is clearly excluded.

Moreover, a $\mathcal{CP}$-violating process is needed. In the SM, the CKM matrix of quarks provides such a $\mathcal{CP}$-violating phase. However, the observed resulting $\mathcal{CP}$-violation is several orders of magnitude too small to explain the observed matter-antimatter asymmetry in the universe [61].

More technical caveats of the SM are the following. First, after the successful unification of the electromagnetic and the weak force, it is tempting to suppose that all forces including the strong and the gravitational force emerge from a single symmetry. This is referred to as a theory of everything. To start more humble, the possible unification of the electroweak and the strong force to a so called grand unified theory (GUT) is widely studied. Promising candidates for such a fundamental symmetry could be e.g. $\mathrm{SO}(10) \rightarrow \mathrm{SU}(5) \rightarrow \mathrm{SU}(3)_\mathrm{C} \times \mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$, although most of the hypothetical theories come with some serious drawbacks. Moreover, an interesting fact of the SM is, that quantum anomalies are only cancelled if the following relation holds:

$$0 \overset{!}{=} N_\mathrm{C} \cdot q^\mathrm{quarks} + q^\mathrm{leptons} = 3 \cdot \left(\frac{2}{3} - \frac{1}{3}\right) + (0 - 1) \tag{2.17}$$

This induces an interesting correlation of the electromagnetic charge and the number of colour charges which can be interpreted as a hint of an underlying theory.

Indeed, the running of all four interaction couplings points into the same direction. It seems that the three couplings of the SM reach comparable values at roughly 10^{16} GeV which is referred to as the GUT scale. However, in the SM these couplings do not meet exactly, thus additional tuning like new processes or new particles would be needed. This might emerge somewhere within the untested 12 orders of magnitude up to the GUT scale. It is not too far-fetched to notice that the gravitational coupling which reaches a strength of $\mathcal{O}(1)$ at the Planck scale of 10^{19} GeV might have similar values to the SM couplings somewhere in between the GUT scale and the Planck scale. Eventually, it is desirable that the 27 free parameters of the SM (including neutrino masses) emerge from such a unified theory. At the moment, they are not predicted by theory and can only be tested experimentally.

There are several proposed extensions of the Standard Model which might explain a few of the listed caveats. Many of them come with an extended Higgs sector such as a two-Higgs-doublet model (2HDM) [62]. It is worth to note that the introduced Higgs mechanism in section 2.3 is only the simplest possible solution of electroweak symmetry breaking but apart of that, there is no real reason why this mechanism is chosen over more complicated ones.

The best studied BSM theory is supersymmetry [63, 64]. Originally developed to solve specific problems, it appears to provide solutions for a broad range of the SM's shortcomings and is essential for UV complete theories like string theory. In supersymmetry, among other things the scalar potential must be holomorphic to provide a Lagrangian which is invariant under the theory's specific symmetries. This directly implies that two different Higgs doublets are needed such that masses for up-type and down-type fermions can be provided.

Together with twin Higgs models [65] or composite Higgs models [66, 67], the extended Higgs sector provides an additional heavy scalar which decays into two Higgs bosons and thus could be detected with resonant production in a process as depicted in figure 2.10(c). Other heavy resonances might occur due to gravitons or radions [68].

The non-resonant coupling might be altered by BSM theories such as the following very selective example found in [69]. By expanding the Higgs sector to a $\mathcal{CP}$-violating two Higgs-doublet model (2HDM) it is claimed that the needed first-order electroweak phase transition as well as additional sources of $\mathcal{CP}$-violation are achieved. Most importantly for this thesis, the introduced new model would significantly alter the trilinear higgs coupling and enhance it up to a factor of $\kappa = g_{hhh}^{\text{model}}/g_{hhh}^{\text{SM}} = 2.5$.

CHAPTER 3

The Experimental Setup

This work is based on data taken at the ATLAS experiment at CERN. In this chapter, a brief overview of the underlying particle accelerator, the LHC, is given in section 3.1. In section 3.2, the ATLAS detector and its sub-systems are introduced and the used reconstruction techniques of physical objects are briefly presented (cf. section 3.3). The computing techniques to simulate and analyse physical processes are explained in section 3.4.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [70–72] is the world's largest and most powerful accelerator with a circumference of 27 km and is located and maintained at CERN, Geneva, near the Swiss-French border (cf. figure 3.1). It is a circular accelerator and is designed to produce proton-proton collisions with a centre-of-mass energy of 14 TeV. Moreover, it also accelerates and collides heavy ions like lead nuclei. As the latter program focusses on different physics, it is not relevant for this thesis and will not be described in much detail.

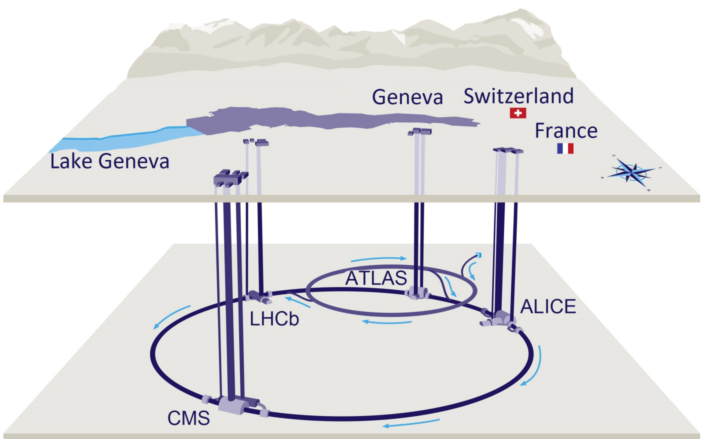


Figure 3.1: The LHC and its four large experiments are shown. Their surrounding landscape at the Swiss-French border is indicated (figure from [73]).

The LHC hosts four large experiments, namely ATLAS, CMS, LHCb and ALICE. Each experiment is specially adjusted for different exploratory focusses and purposes. ALICE mainly studies the collision of heavy ions to widen the knowledge of fundamental properties of the strong force and to probe conditions comparable to the early universe. LHCb focusses on flavour physics with the b quark and is designed to precisely record the boosted region close to the beam pipe. Finally, ATLAS and CMS are designed to probe a broad range of high-energy-particle-physics processes. Their main purpose was the experimental discovery of the Higgs boson in particular and the search for several BSM physics scenarios in general. The former detector recorded the data used in this analysis and is introduced in detail in the following sections.

For these purposes, a very high amount of particle collisions is needed which scales as $N_{\text{event}} = \mathcal{L} \cdot \sigma_{\text{event}}$. Hereby, the cross section σ_{event} can be interpreted as a probability of each specific process to happen and reaches from $\sigma = 7.81 \times 10^{-2}$ b for proton-proton collisions in general [48] to $\sigma = 1.3 \times 10^{-16}$ b for the signal process $gg \to hh \to 4\tau$ of this thesis.

The luminosity $\mathcal{L}$ is the amount of events per time and area and depends on machine parameters of the accelerator like the number of particles in each particle bunch, the number of bunches in the accelerator and the comprehensiveness of the bunches at the collision point. At the LHC, luminosities of $2 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ are achieved [74]. The required number of particles in each bunch for such high luminosities is impossible to achieve with antiprotons. Accordingly, two separate acceleration rings are used in which counter-rotating proton bunches are accelerated.

Several pre-accelerators at CERN serve as an injection chain for the LHC and ramp up the proton energy to 450 GeV. At the LHC, 1 232 dipole magnets with magnetic fields up to 8.36 T force the protons into a circular movement. Additionally, 858 quadrupole magnets are used to focus the beam. Eventually, the particles are brought to collision at four interaction points. Here, the named four experiments are located which are designed to study the collision products of the particles.

Due to the large proton-proton cross section, one bunch crossing contains several proton interactions, referred to as *pile-up* depending on the focussing magnets in front of each experiment. At the ATLAS detector, the exact number of interactions varies roughly between 10 and 70 (cf. figure 3.2).

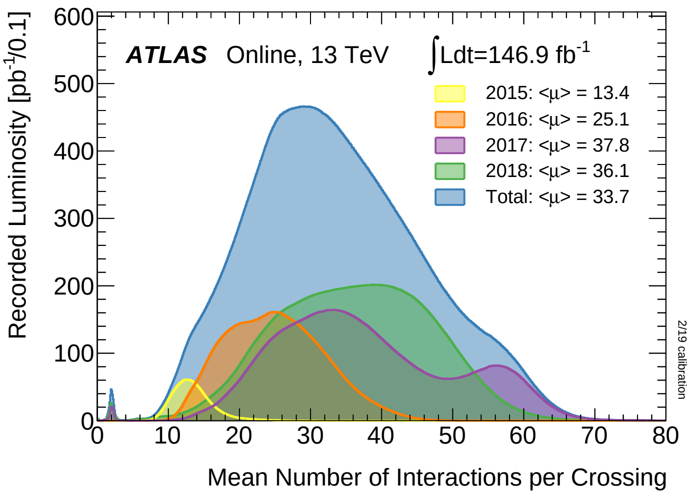


Figure 3.2: Pile-up at the ATLAS detector for different data taking periods (figure from [75]).

The amount of data taken by the experiments is measured in integrated luminosity and the ATLAS detector recorded a total of $\int \mathcal{L}\,\mathrm{d}t = 147\,\text{fb}^{-1}$ of usable data with a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ [75]. This thesis, however, focusses on the data taken in 2017 which adds up to $46.9\,\text{fb}^{-1}$. In this data taking period, the average amount of pile-up events was 37.8 with a broad distribution (cf.

figure 3.2). In general, the amount of pile-up events might not be negligible for this thesis as the signal region will suffer from many fake-tau jets (cf. chapters 5, 6). More pile-up events might increase the number of fakes and they pollute the low-momentum region which impedes the tau reconstruction. As an outlook, another 150 fb^{-1} of data are planned in Run 3 of the LHC which starts in 2021 before the LHC will shut down for major upgrades.

The following so called **high-luminosity LHC** (HL-LHC) will use upgraded technology such as magnets up to 12 T and is planned to reach peak luminosities of $5 \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$. This is partly done by increasing the average pile-up to values of 140, which will lead to much more challenging tasks in terms of tracking, vertexing and particle identifications. Starting in 2026, the HL-LHC is planned to increase the total amount of collected data to 3 000 fb^{-1} until 2037 [76].

This thesis will study the significance of the $hh \rightarrow 4\tau$ channel after full data taking. Hence, if not stated differently, an integrated luminosity of 3 000 fb^{-1} will be assumed. Hereby, the possible difficulties of the much higher pile-up are neglected. However, the HL-LHC might run at centre-of-mass energies of 14 TeV which increases the signal cross section compared to many background cross sections [32]. Moreover, the ATLAS detector will be upgraded to higher solid angle coverage and thus might lose less signal events (cf. section 3.2.2). Hence, positive and negative influences of the next upgrade might at least equal out (in a pessimistic scenario) such that the data of 2017 is thought to serve as a good extrapolation dataset for the full 3 000 fb^{-1} in future.

3.2 The ATLAS Detector

The ATLAS (**A** **T**orodial **L**HC **A**pparatu**S**) experiment [77] (cf. figure 3.3) is one of the two multi-purpose detectors at the LHC. As well as the similar experiment CMS, it is designed to detect the decay products of particles produced in high-energy-particle-physics processes of broad phenomenology such as in Higgs boson processes or hypothetical events beyond the Standard Model.

It covers the interaction point with a solid angle of almost 4π such that as many emerging particles as possible are detected. For this purpose, the detector surrounds the named point with several layers of specific sub-detectors and forms a cylindrical shape. In the barrel region, the sub-detectors encase the beam pipe and in the end-cap region at each end, they are arranged perpendicular to the beam in a wheel-like structure. In total, the detector spans a volume of $44\,\text{m} \times \pi(12.5\,\text{m})^2$ which makes it the largest experiment at CERN.

Concerning their purposes, the subsystems can roughly be arranged in three categories. They are briefly discussed in the following sections. The Inner Detector (section 3.2.2) reconstructs the tracks and momenta of charged particles. It is surrounded by different calorimeters (section 3.2.3) which measure the energies of different particle types. The outermost part of the ATLAS detector consists of the muon spectrometer (section 3.2.4) which detects and measures weakly interacting particles such as muons [77].

3.2.1 Used Coordinate System at ATLAS

Before dedicating the following sections to the subsystems, a short introduction to the used coordinate system at the ATLAS detector is useful.

Its origin lies in the nominal interaction point in the centre of the detector and the z-axis is pointing along the beam axis. Its x-axis points towards the centre of the LHC accelerator ring and the y-axis points upwards which makes it a right-handed coordinate system. The usual spherical coordinate system might be used. Here, $\phi \in [-\pi, \pi]$ defines the azimuthal angle which spans the area of the x-y-plane and $\theta \in [0, \pi]$ defines the polar angle between the z-axis and the selected direction.

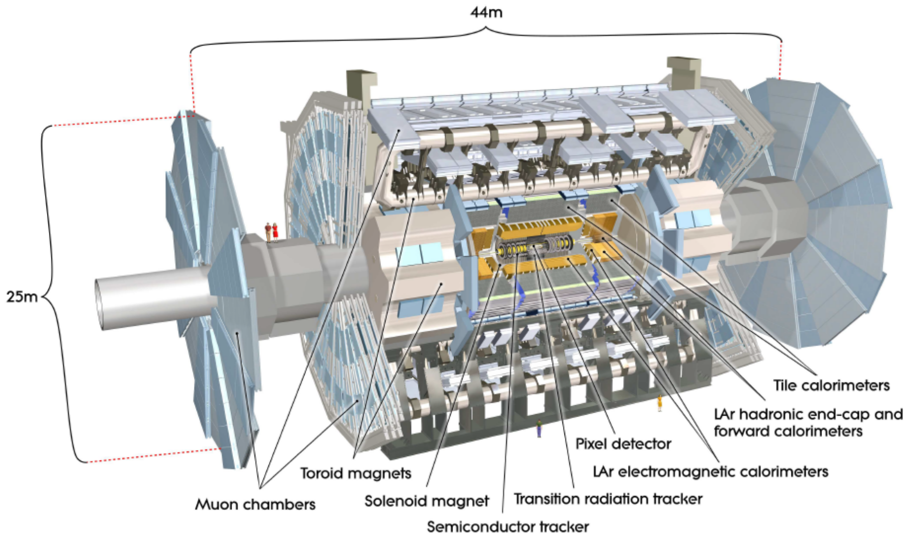


Figure 3.3: A schematical view of the ATLAS detector and all its components is presented. Its dimension is illustrated by the comparison to several dummy-people standing in front of the detector (figure from [77]).

However, it is more convenient to introduce the pseudorapidity η:

$$\eta = -\ln\tan\left(\frac{\theta}{2}\right) \quad . \tag{3.1}$$

An advantage of η is that the fraction of emerging particles at a hadron collider is roughly equal in each $\Delta\eta$-interval. Moreover, the difference between two pseudorapidities is Lorentz invariant. η runs from 0 at the perpendicular direction to the z-axis ($\theta = 90°$) to infinity (parallel to the z-axis, $\theta = 0°$). Nevertheless, due to detector properties, this thesis will be limited to $\eta = 2.5$ which corresponds to $\theta = 9.4°$.

As already stated in section 2.3.1, the colliding protons are composite particles. The interaction is carried out by partons whose original momentum fraction of the proton is described by the PDFs and remains unknown for a specific event. Hence, the system is very likely to be boosted in one direction as the momenta of participating partons are usually not balanced. As a result, at hadron colliders, it is useful to use quantities which are easily transformed and the distance between two objects will be calculated as follows:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad . \tag{3.2}$$

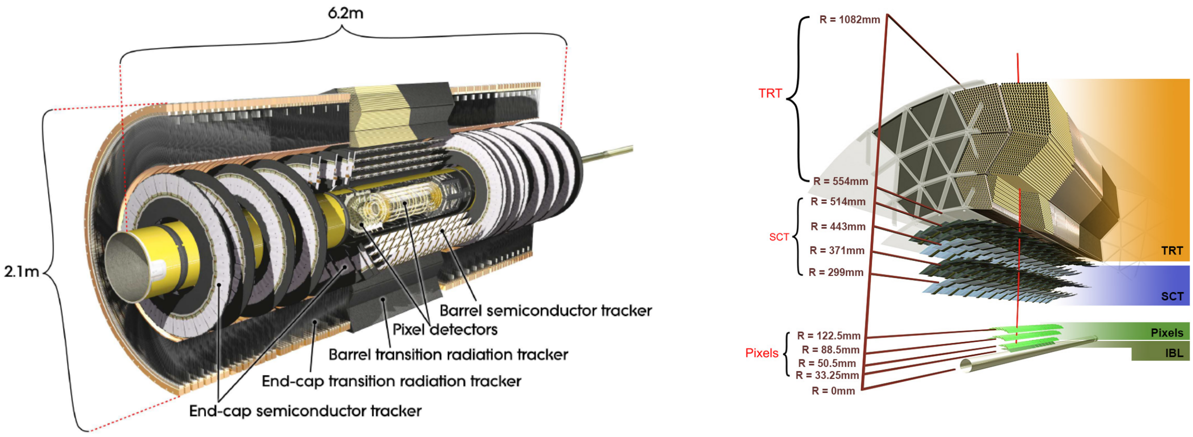
Furthermore, the transverse plane becomes of enhanced interest as only here, momentum conservation can be assumed. The transverse momentum is given by

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2} \tag{3.3}$$

and can be interpreted as the projection of the total momentum to the named plane.

3.2.2 Tracking and Vertexing

The innermost detector of ATLAS [78, 79] is located directly around the beam pipe and is designed to reconstruct the tracks and vertices of electrically charged particles (cf. figure 3.4). For this purpose, the Inner Detector (ID) is surrounded by a solenoid which creates an axial magnetic field of 2 T such that the charged particle's tracks are bent depending on their momentum.



(a) Longitudinal profile [77]

(b) Radial profile [80]

Figure 3.4: The Inner Detector is represented with two different lateral cuts. In (a), the proportions of each detector component are visualised and in (b) the barrel part is shown in more detail.

Its innermost ingredient is a four-layer **Pixel Detector** (PD) which surrounds the detector centre with a radius spanning from 33.25 mm to 122.5 mm. At the end-caps, the PD is completed by three disks and in total, it covers a pseudorapidity region up to $|\eta| < 2.5$. Due to its vicinity to the interaction point, it is exposed to a very high track density and hence a large radiation dose. To allow for a precise vertex and momentum reconstruction anyhow, it is highly granulated and consists of more than 90 million silicon pixel detectors. Each pixel has a size of $50\,\mu\text{m} \times 400\,\mu\text{m}$ or $50\,\mu\text{m} \times 250\,\mu\text{m}$ for the innermost layer. With this setup, a resolution of 10 µm in the transverse plane and 115 µm in z direction is achieved [77, 80].

The PD is surrounded by the **Semiconductor Tracker** (SCT). It consists of four cylindrical layers of silicon stripes in the barrel region. In total nine wheels of silicon strip detectors are arranged in the end-caps to enable at least four space point measurements up to $|\eta| = 2.5$. Each strip is between 5 cm and 12 cm long and has a width of 80 µm. The layers are arranged with a small angle with respect to each other to enable information about the z direction [81]. In total, the detector achieves a resolution of 17 µm in the transverse plane and 580 µm in z.

The ID is completed by the **Transition Radiation Tracker** (TRT). In contrast to the other two tracking detectors, it is a gaseous detector which consists of 351 000 mainly xenon filled straw tubes with an detection wire along their central axis. Its detection principle is analogous to the one of a drift chamber in which particles ionise the gas mixture along their trajectory. The ionised gas is accelerated towards the detection wires creating an avalanche of ions which then can be detected.

Additionally, the TRT is also used for electron identification. For this purpose, special materials are places in between the straw tubes such that high energy particles emit transition radiation due to quickly changing refractive indices. Light particles like electrons are much more likely to produce transition radiation than heavier ones such as pions. In the end-cap region, the barrels are arranged radially and the TRT covers a range of $|\eta| = 2.0$ [82].

The ID limits the solid angle coverage of the ATLAS detector. As the two most sensitive tracking detectors (the PD and SCT) only cover the detector up to $|\eta| = 2.5$, no particles outside of this region are considered. After the HL- LHC upgrade, however, it is planned to achieve a full coverage up to $|\eta| = 4.0$.

3.2.3 Calorimetry

The calorimeter system (cf. figure 3.5) is dedicated to measure the energy of emerging particles and covers a range up to $|\eta| = 4.9$. It directly encases the Inner Detector.

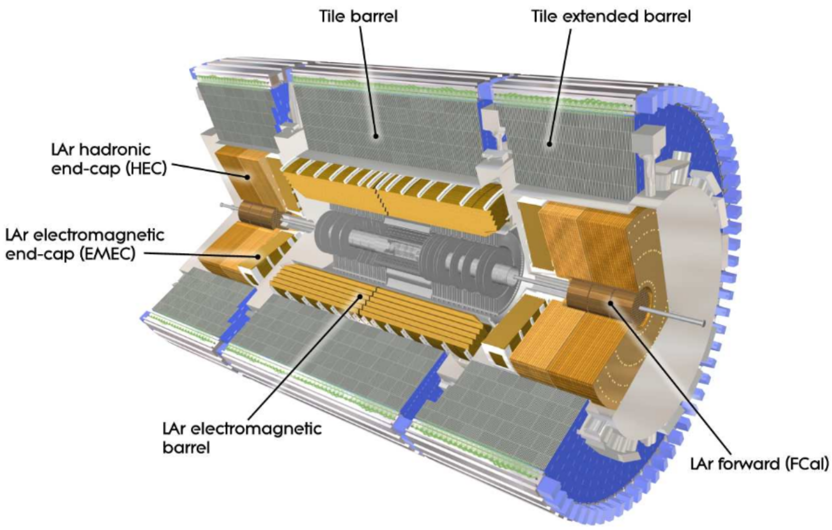


Figure 3.5: Cut-away view of the ATLAS calorimeter system (taken from [77]).

In opposition to the ID, which is designed such that it is transited by particles with the lowest possible energy loss, the calorimeter system is constructed with the aim that particles deposit their whole energy in this system. Especially neutral particles which do not leave a track in the ID can only be detected by the calorimeters. Thus, it is designed sufficiently large and divided into two subsystems due to the diverse interaction behaviour of different particles.

The **electromagnetic calorimeter** (eCAL) is designed to stop electromagnetically interacting particles like electrons and photons such that their whole energy can be measured. Two different alternating layers with separate tasks are arranged in an accordion shape for this purpose. On the one hand, the particles deposit their energy in the passive absorber material made out of lead plates by creating relatively narrow electromagnetic showers. On the other hand, the active absorber material liquid Argon (LAr) gets ionised by the particle showers and creates a measurable signal proportional to the particles energy.

To ensure that most mainly electromagnetically interacting particles deposit their full energy in this calorimeter, it has a thickness of more than 22 times the radiation length X_0 in which particles lose all but $1/e$ of their original energy.

Particles that pass the eCAL are most likely hadrons which then are stopped in the **hadronic calorimeter** (hCAL). Such as the eCAL, the hCAL consists of alternating layers of active and passive absorber materials. Hereby, the passive material mainly consists of steel in the barrel region while scintillating tiles form the active layers. The latter ones are read out by photomultipliers. In the end-caps, liquid argon serves as the active material and copper is used as the passive absorber. In the very forward region, the forward calorimeter allows energy measurements up to $|\eta| = 4.9$. It is exposed to very high occupancy and radiation doses and uses copper and tungsten as passive material. Again, liquid argon is used as the active material. The hCAL is designed to stop strongly interaction particles like jets or the daughter particles of hadronically decaying τ leptons. They initiate hadronic showers that show a much larger structure than the electromagnetic ones. Thus, the hCAL is even larger than the eCAL and reaches a thickness of roughly 9.7λ where λ is the interaction length defined as the mean free path of a particle before it scatters inelastically [77].

Unfortunately, between the barrel and the end-cap region, a small region is occupied by cables and cryostats which leads to a bad energy resolution. To account for this, usually particles within $1.37 \leq |\eta| \leq 1.52$ are discarded in ATLAS analyses.

3.2.4 Muon System

Muons are the only detectable particles of the SM which cannot be stopped by the calorimeters. They do not interact via the strong force and due to their higher mass they by far do not lose as much energy as electrons via bremsstrahlung. To still measure the trajectory and momentum of minimally ionising particles like the eponymous muon, the ATLAS detector is completed by the muon system (depicted in light blue in figure 3.3).

It is pervaded by a toroidal magnetic field which forces particles on a curved track and thus enables to measure information about their momenta. To achieve good resolutions, the muon system is large and gives the ATLAS detector its size. Four different subsystems are used. Hereby, the **Monitored Drift Tubes** as well as in the **Cathode Strip Chambers**, both gaseous detectors, provide precise spatial measurements perpendicular to the beam axis.

Moreover, the muon system provides a very good time resolution and a fast responses and hence can be used for fast trigger decisions. The **Resistive Plate Chambers** as well as the **Thin Gap Chambers** are designed for this task and achieve time resolutions of 4 ns. Additionally, they measure the spatial observable parallel to the beam axis [77].

3.2.5 Trigger System

At the ATLAS detector, every 25 ns the proton beams collide which is equivalent to a frequency of 40 MHz. Moreover, each collision produces roughly 10 to 70 interactions (cf. figure 3.2). Considering an average event size of 1.5 MB, a storage of all these events would require 60 TB of data each second, which is not only strongly limited by the overall required storage facilities but also technically in terms of data transfer time. Fortunately, most events are physically not very interesting and contain only soft QCD interactions with low transverse momenta. Therefore, a fast decision system is implemented to select events with potentially interesting properties. For example, high transverse momenta indicate the production of a heavy (and thus potentially interesting) particle. All other events are then discarded and cannot be reproduced if in future some properties might become interesting which nobody thought of today.

This delicate task is fulfilled by the trigger system [83]. In the version used in Run II a hardware-based level 1 trigger (L1) reduces the data rate to 100 kHz in a latency of 2.5 µs. Hereby, compressed information of the calorimeters and the muon system is available. Events which were accepted by the L1 trigger are further processed by the software-based High-Level Trigger (HLT), which reduces the data rate further to 1 kHz. The HLT bases its decisions on full detector information and first reconstructions which are available for 0.2 s. Each event passing the HLT is stored permanently and accordingly, a storing capability of the order $\mathcal{O}(1\,\mathrm{GB/s})$ is needed [83, 84].

A serious issue in future might be to actually trigger on a multi-tau system as it is demanded in this thesis for the $4h0\ell$ channel. Especially, when all taus are soft, it will be very challenging for the triggers to distinguish them from jets at the first place. This problem will intensify at the HL-LHC where much more pile-up is expected. Thus, a neat trigger solution has to be found once the possibilities at the HL-LHC are foreseeable.

In this work, a general combination of any trigger is used which then demands e.g. at least two high energetic taus or similar properties. Especially at the HL-LHC data-taking period, the trigger system will significantly be altered and further development of tau triggers is expected. Therefore, it is relinquished from selecting one possible trigger scenario for current data which has to be completely reconsidered for most of the expected $3\,000\,\text{fb}^{-1}$ anyway. As all signal selections defined in section 6.4 include high requirements on the sum of all four lepton momenta, it is imaginable that a high-p_{T} cut on at least one or two of the τ leptons due to a future trigger selection might not reject further signal events.

For leptonically enriched final states, the most promising approach probably will be to trigger on the light lepton as these triggers achieve much higher reconstruction efficiencies.

3.3 Reconstruction and Identification of Physical Objects at the ATLAS Detector

Proton-proton collisions produce various particles which decay further or directly burst apart of each other. The detection and identification of these particles is the main purpose of the ATLAS detector. Each particle type leaves characteristic traces in the detector and information of several sub-detectors is combined to identify them. Only neutrino interactions are orders of magnitude too weak such that they leave the detector without any direct signal. A short overview of the basic methods to reconstruct jets, leptons and missing transverse energy is given in this section.

3.3.1 Jet Reconstruction

As already described in section 2.2, the strong coupling grows for low energies or correspondingly low transferred momenta. As a result, in strong interactions the relative momenta of emerging quarks and gluons are mostly small such that they are usually collimated to the direction of the original particle. Moreover, due to the confinement, coloured particles cannot exist on their own but form colour neutral states in the process of hadronisation. Therefore strongly interacting particles or hadronically decaying τ leptons usually result in a jet of hadrons pointing into similar directions [17].

At the ATLAS detector, jets are typically reconstructed by using the anti-k_t-algorithm [85], which uses three-dimensional, togological clusters as input. In these clusters, the signals of neighbouring cells in the calorimeters are merged. Here, cells with a high energy deposit serve as seed cells from which neighbouring cells are iteratively added depending on their signal properties. The anti-k_t-algorithm then defines the distance of the clusters with respect to each other and to the beam and combining the closest objects. If the closest distance is the one to the beam, the object is considered to be a jet and is removed from the list of unidentified objects. By repeating this step several times, eventually all objects are assigned to jets. In this process, the size of the jets can be governed by adjusting corresponding parameters in the distance measure. Typically, the jets are formed in a radius of $\Delta R < 0.4$.

Finally, to determine the jet's energy and momentum several corrections apply. For instance, pile-up effects have to be considered. Further calibrations are done by comparing the reconstructed jets to Monte Carlo simulations.

Among the various amount of QCD jets, the ones containing b quarks, so called "***b*-jets**" can be singled out. Here, one exploits the fact that hadrons containing a b quark only decay under the weak interaction and thus (cf. section 2.2) have a relatively long lifetime of the order of $O(1.5\,\text{ps})$. Accordingly, due to their additional high energies, their travelled distance before the decay into a jet becomes measurable which results in a displaced vertex. Moreover, the usual decay modes of b-hadrons are used such that quite effective b-tagging algorithms have been developed [86]. In this analysis, b-tagging is mainly used to separate QCD and t quark background from the signal.

3.3.2 Lepton Reconstruction and Light Lepton Identification

The outstanding property of the signal channel of this analysis is that it contains solely four τ leptons in its final state. With a mass of $m_\tau = 1.777\,\text{GeV}$, the τ lepton is the heaviest of all leptons and due to its small lifetime of 290 fs, τ leptons have to be reconstructed from their decay products. Its high mass enables the decay into hadrons in addition to the emerging τ-neutrino, which occurs with a probability of 65 % [11].

An essential task, while working with τ leptons, is to distinguish hadronically decaying τ leptons (further referred to as τ_{had}) from QCD jets and other fakes. This is addressed phenomenologically in this section and in more technical detail in chapter 5.

Leptonically decaying τ leptons (τ_{lep}, occuring with a probability of 35 %) suffer from an additional undetectable neutrino and cannot be distinguished from directly emerging electrons or muons in the ATLAS detector. Instead, these light leptons are directly reconstructed as such and their reconstruction and identification techniques are very important in the signal channels allowing leptonic decays of one or two of the signal taus. Hence, the characteristic properties of each lepton are shortly explained in this section.

Electrons

Electrons leave signals in the Inner Detector and are stopped by the eCAL. Therefore, information of these two systems is combined for electron reconstruction and identification. There are algorithm starting the other way around, but usually, the electron identification algorithm starts with the search for an characteristic eCAL cluster with $p_{\text{T}} > 2.5\,\text{GeV}$. This cluster is then matched to a track. In this process tracks with many hits in the silicon tracker are preferred. Moreover, electrons differ to other particles by their strong transition radiation in the TRT as described in section 3.2.2. Usually, it is demanded that they appear isolated and carry at leasts 90 % of the detected energy in their surrounding. Depending on several criteria like the track quality, their shower shape in the eCAL or the ratio of their measured energy in the calorimeter and momentum measured in the tracker, several likelihood-based working points like *Loose*, *Medium* or *Tight* are defined in ATLAS analyses. They differ in the achieved electron-efficiencies compared to background rejection. For example, for *Medium* electrons, efficiencies of 88 % up to 95 % are achieved depending on the electron p_{T} [87].

The electron signal in the eCAL can be mimicked by photons. Indeed, the photon algorithm also starts by searching for eCAL clusters. However, photons are neutral and do not leave a track in the Inner Detector accordingly. Even if the photons already decay via pair production into e^+e^--pairs in the ID, the emerging electrons can be distinguished from original ones by the two clear trajectories of different charge emerging from the same vertex [88, 89].

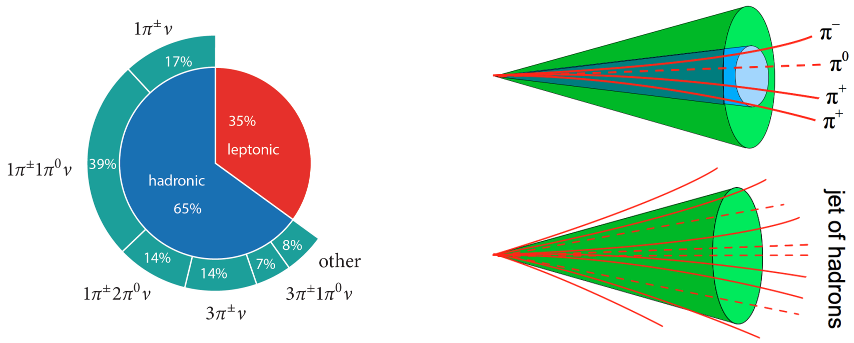
Muons

For muon reconstruction and identification, signals of the Inner Detector, the calorimeter and the muon system are combined. As muons are the only particles, which manage to leave the detector's calorimetry, they leave a very characteristic signature in the muon system. At first, particle trajectories are reconstructed independently in the ID and in the muon system. The tracks are then extrapolated and compared with the reconstructed tracks in the other system. Muons are called *combined muons* if such a track agreement is found. *Segment-tagged muons*, however, are muons whose track in the ID can only be associated to at least one track segment in the MDT or CSC. Typically, these are low p_T muons or they traverse a low-acceptance region of the muon system. Muons which leave no information in the muon chambers but can be identified due to their track and energy deposition in the calorimeters (which has to match an expected signal of a minimal ionising particle) are called *calorimeter-tagged muons*. Contrary to that, muons which are reconstructed solely with muon spectrometer information are called *stand-alone muons*.

Apart of muons, only highly energetic particles partly manage to escape the calorimeters. However, in opposition to the muons, they deposit huge amounts of energy in the calorimeters before and thus by capitalising on the calorimeter's information, effective methods to identify muons were developed [90]. Depending on the transverse momentum of the muon, reconstruction efficiencies of up to 97 % are achieved [91].

The high reconstruction and identification efficiencies for light leptons lead to the fact that in this work, pure QCD fake objects will be considerably less important in leptonic decay channels than in the $4h0\ell$ channel, in which all taus decay hadronically (cf. section 6.2).

Hadronically decaying τ leptons



(a) Figure from [92]

(b) Figure from [93]

Figure 3.6: The tau decay modes with the corresponding branching ratios and the number of emerging charged and neutral particles are shown in (a). The leptonic decay mode is further divided into electron and muon decays which occur nearly equally often. In figure (b) a typical particle shower of a 3 prong tau decay is compared to a shower emerging from a hadron.

Roughly three out of four hadronically decaying τ leptons decay into one charged particle ("1 prong"), while the other one decays into three charged decay products ("3 prong") [11].[1] These particles are mostly pions (cf. figure 3.6(a)). Apart of that, one or more neutral particles such as π^0 can be produced and for every decaying tau, one τ neutrino emerges. Since the neutrino leaves no trace in the detector, the overall tau energy cannot be detected directly.

As stated in section 3.3.1, hadrons are reconstructed as jets, hence also hadronically decaying taus are first reconstructed as jets with the anti-k_t algorithm. If they satisfy $p_\mathrm{T} > 10\,\mathrm{GeV}$ and $|\eta| < 2.5$, they serve

[1] The amount of hadronically decaying taus which produce five or more charged particles is less than 0.2 % and therefore can be neglected.

as seeds for the τ_{had} reconstruction. The challenge is to distinguish hadrons emerging from a tau to them emerging from a quark or gluon.

Accordingly, the tau decay properties are fully exploited. First, jets tend to produce much more particles and thus charged tracks than τ leptons. Moreover, the overall charge of a jet does not need to be limited to ± 1 and jets tend to be broader. In general, jets emerging from τ leptons are more collimated than others because they are already colour neutral (cf. figure 3.6(b)). Consequently, the tau candidates are divided into a *core* ($0 < \Delta R < 0.2$) and an *isolation* ($0.2 < \Delta R < 0.4$) region. The tracks of a tau candidate are required to lie in the core region. They have to fulfil $p_{\text{T}} > 1\,\text{GeV}$ and a sufficient amount of hits in the SCT and pixel detector. The tracks within the isolation region are further used for tau identification variables (cf. section 5.2). The tau vertex is determined by choosing the vertex with the highest contribution to the transverse momentum p_{T} of the τ lepton. [94–96].

In this thesis, the exact tau identification method differs to the approach which is usually used in ATLAS analyses. This novel approach as well as the detailed and technical usual reconstruction is presented in chapter 5.

3.3.3 Reconstruction of Missing Transverse Energy

Although the overall momentum of the fundamental parton-parton collision remains unknown in a hadron collider (cf. sections 2.3.1, 3.2.1), statements about the transverse plane are possible. Compared to the usual energies considered at the LHC, the partons can be assumed to have no initial momentum component perpendicular to the beam axis. Thus, energy and momentum conservation can be applied in the transverse plane and a balanced distribution is expected. This balance can be spoilt by particles which leave the detector undetected and thus carry away a certain amount of energy and momentum. In the SM, only neutrinos serve as such particles. Summing up, by considering all measured energies and momenta, the missing transverse energy $E_{\text{T}}^{\text{miss}}$ can be reconstructed.

As at least four tau neutrinos are expected to be produced in the signal region of this analysis, $E_{\text{T}}^{\text{miss}}$ is an interesting quantity to look at even though is seems not to provide the best separation power. This can be explained if the neutrinos leave the detector in different directions. In this case the effect of a considerable $E_{\text{T}}^{\text{miss}}$ in one particular direction is balanced out.

3.4 Computational Methods, Data Simulation and Software

3.4.1 Event Generation and Detector Simulation

To draw meaningful conclusions in science, observations have to be compared to predictions from (possible) underlying theories. In particle physics, however, the data as well as the predictions from the Standard Model quickly become very complex as a huge multitude of effects and probabilities provide a rich field of possible observables and phenomena. Thus, precise simulations of particle physics events and the resulting detector response are crucial for the interpretation of ATLAS data.

Starting with the proton-proton collision, the produced particles and their properties as well as their decay products have to be modelled. Eventually, their interaction with the ATLAS detector and its response have to be simulated. To account for the probabilistic quantum nature of particle physics processes, Monte Carlo (MC) simulations are used for the event generation. Here, the actual physical process is divided into several steps which are neatly visualised in figure 3.7.

At the primary interaction of each proton-proton collision process in this analysis, a **hard parton-parton interaction** (signalled in red in figure 3.7), the momentum transfer is large. This results in a coupling of the strong force which is weak enough to be described with perturbation theory (cf.

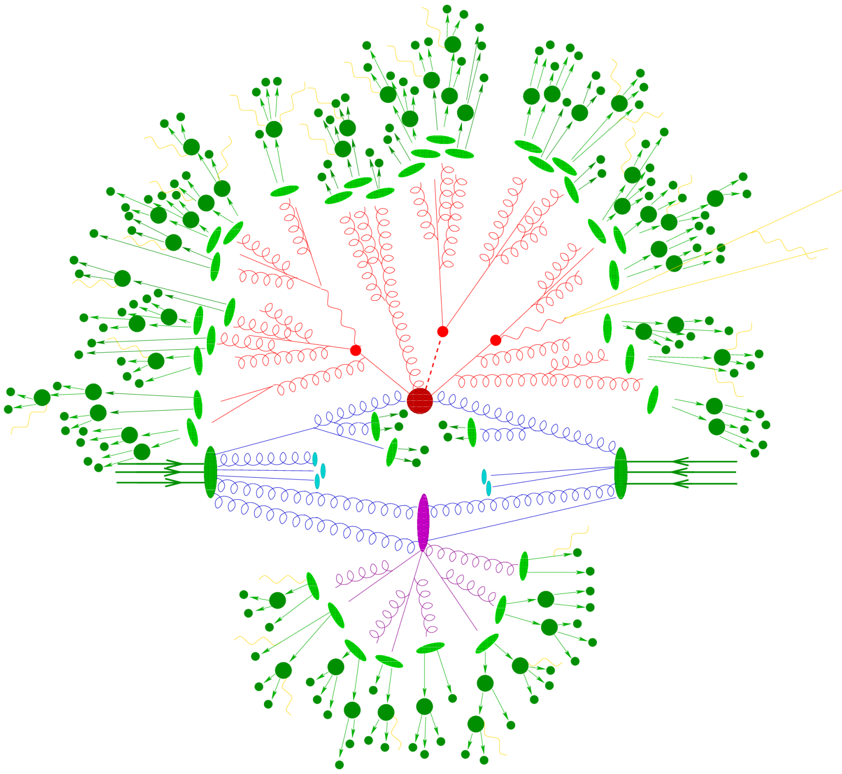


Figure 3.7: The sketch of a hadron-hadron collision as simulated by a Monte-Carlo event generator is shown. In this figure, the dark red blob represents the actual hard collision and in red its collision products and bremsstrahlung processes are depicted. The purple blob signals the event from the proton remnants and in light green, the hadronisation process, i.e. the transition of partons to hadrons (dark green) is shown. Yellow lines denote soft photon radiation (taken from [97]).

figure 2.4). Hence, matrix elements can be calculated with Feynman diagrams. Moreover, the parton distribution functions of the proton (cf. figure 2.6) are used. Additionally, the interactions of the proton constituents which do not participate in the hard scattering have to be modelled (purple). This is typically a non-perturbative process due to its low-momentum transfer and corresponding strong coupling.

Next, initial (marked in blue) and final (again in red) state radiation in which partons radiate of a gluon, cause additional **parton showers**. These showers are simulated as an extrapolation from the scale of the hard process towards the non-perturbative confinement scale. However, parton shower algorithms are limited to a certain cut-off scale. In the infrared scale, where QCD becomes strongly interacting, sophisticated non-perturbative models of **hadronisation** (green) have to be taken into account. This is done on a purely phenomenological level. The two most successful models of hadronisation are the string and the cluster model. Here, the partons and gluons are either modelled as a connected string or as parts of a colour-singlet cluster [98].

The event generators provide the generated particles and their four-momenta. Now, their interaction with the detector has to be modelled which is done by the program `GEANT4` [99]. This includes effects of their deflection by the detector's magnetic fields up to the release of additional particles as showers in the calorimeters and the corresponding energy loss. Next, the detector response is mimicked by simulating the electrical readout signals caused by the primary and secondary particles. This output is analysed by the same reconstruction algorithms as explained in section 3.3 such that the simulated events are comparable to actual data.

The search for rare processes usually takes place in the tail region of various background processes. For a proper estimation of these tail regions, many events have to be simulated. Moreover, the simulated events have to be validated and optimised by comparing them to data in well understood control regions.

In this analysis, the events involving vector bosons and leptons were fully simulated by `Sherpa` [100], which is a general purpose leading order event generator that simulates the full event from the hard interaction to hadronisation. For processes involving t quarks or Higgs bosons, `PowHeg` [101] was used as the event generator and `Pythia8` [102] simulated the parton showering and hadronisation by using the Lund string model. For the $hh \to 4\tau$ signal sample, `MC@NLO` [103] was used as the event generator and the particle showering was modelled with `Herwig` [104]. A full table of the used samples and their generators is given in the appendix in table A.1 to A.6.

3.4.2 Boosted Decision Trees

The boosted decision trees used in this thesis are binary classification methods to discriminate signal from background events. Here, signal events can either be truth taus (cf. chapter 5) or $hh \to 4\tau$ events (cf. chapter 6). In general, a boosted decision tree (BDT) consists of a variable amount of decision trees which are weighted depending on their separation power between the classes.

In a decision tree, data is split into several subsets with respect to specific properties. These properties are defined by the input variables of the tree. Each tree consists of a number of binary decisions which each divide the dataset further. In figure 3.8 a binary tree structure is sketched.

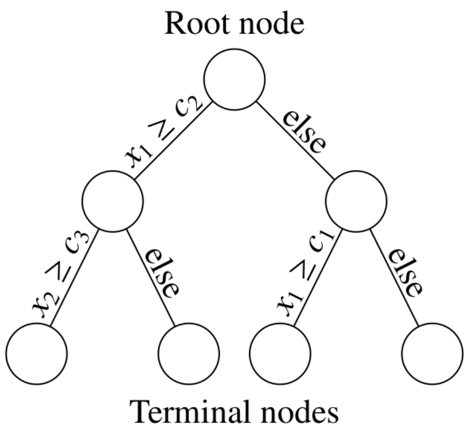


Figure 3.8: A pictorial decision tree is visualised. The tree has a depth of two and utilises two input variables x_1 and x_2 (Figure taken from [105]).

While the root node contains the full dataset, it is split into several disjoint subsets which are represented by the terminal nodes in figure 3.8. The splitting depends on the defined maximum depth of the tree and further stopping criteria like a minimal number of training events in each node. Commonly used for creating decision trees is the Gini impurity $I_G(p) = 2p(1 - p)$. Here, p is the signal purity in each node. Usually, the best split is chosen such that the mean of the Gini impurities of the resulting nodes is minimised. In this process, each node is weighted with its sum of event weights [105]. In the program `TMVA` [106], which was used in chapter 6, a further parameter is defined which governs the number of grid points in each input variable to find the optimal cut in node splitting.

For the used BDTs, hundreds of such decision trees are produced. To improve the overall separation power of the two input classes (signal and background), these trees are now *boosted* (hence the straightforward name *boosted decision tree*). In the boosting process, algorithms are used which in the end assign weights to each decision tree depending on its separation power. The boosting algorithm with best performance was found to be Adaptive Boosting (AdaBoost) [107] in this analysis. Here, the decision trees are trained while assigning higher weights to events which were previously classified incorrectly. Next, a weighted sum of these trees is created and a final BDT score is assigned to each event. For instance, signal-like events might achieve a score close to 1 while background-like events are close to -1. The exact value of the maximal score, however, is arbitrary.

3.4.3 Analysis Software

In this thesis, mainly the `xTauFramework` [108] and `HAPPy` [109] were used and adapted. Both utilize the `ROOT` software package [110]. The former one provides tools to process MC or data samples and was used to create N-tuples from the more detailed `DAOD` files. In this step, the framework provides all common recommendations for object definitions (although the tau object definition was studied in more detail in this analysis and accordingly was altered several times). Also event cleaning and the standard ATLAS object calibration is applied. `HAPPy` was used to further study the N-tuples, to compute results and to visualise them in histograms and graphs.

More suited to specific tasks is the `TMVA` [106] package which is integrated in the `ROOT` software. `TMVA` was used to train and study BDTs (cf. section 6.5). Last but not least, a shape fit of the signal region and a corresponding calculation of an upper limit on the di-Higgs cross section was performed by `HistFitter` [111]. This framework was originally developed to search for SUSY processes but can be used for several signals with low cross sections.

CHAPTER 4

Setting Up the Scene

In this chapter a general guideline for approaching the di-Higgs signal is presented. For this purpose, first it is motivated on truth level[1] that the $hh \to 4\tau$ channel indeed protrude from possible processes with four taus in the SM (section 4.1). On reconstruction level, several technical drawbacks such as fake taus have to be dealt with. A general strategy to do so is presented in section 4.2. Subsequently, in section 4.3 a preselection is introduced which utilises the expected event topology while keeping much signal as possible.

4.1 Truth Level Considerations

As already stated in section 2.3.2, there are several di-Higgs decay channels such as $hh \to b\bar{b}b\bar{b}$ or $hh \to b\bar{b}\tau\tau$ with higher branching ratios than $hh \to 4\tau$ (cf. table 2.1) and hence more signal events. However, they all either contain b quarks or W bosons and thus suffer from a large hadronic background. The $hh \to 4\tau$ channel, though, might single out as one of the very few SM production channels of four τ leptons. Hence, this channel might serve a very clear signal. Indeed, by studying several physical processes on truth level, a promising signal to background ratio of up to 1.5 or 2 is observed (cf. table 4.1).

Cut	$hh \to 4\tau$	Higgs	Multiboson	V + Jets	$Z \to \tau\tau$	Top	$t\bar{t}$	$S/\sqrt{S+B}$
$N_{\tau^{\text{truth}}} \geq 4$	268	6 849	52 672	160	202 903	8 906	96 506	0.44
$q_{\tau_{1-4}}$: 2+2−	268	6 656	50 689	131	151 079	5 668	61 318	0.51
$N_{b\text{-Jets}} = 0$	222	5 605	43 916	113	124 499	2 165	16 022	0.51
$\sum p_{\mathrm{T}}^{\tau^{\text{truth}}_{i=1-4}} > 250\,\text{GeV}$	202	2 349	6 095	9	1 085	208	919	1.94
$\sum p_{\mathrm{T,\,vis}}^{\tau^{\text{truth}}_{i=1-4}} > 200\,\text{GeV}$	107	1 058	2 326	3	258	93	214	1.68
$\Delta R_{\tau^{\text{truth}}_{1/3}\tau^{\text{truth}}_{2/4}} < 2.0$	64	691	876	0.0	115	0	11	1.53
$m^{\text{vis}}_{\tau^{\text{truth}}_{1/3}\tau^{\text{truth}}_{2/4}} > 50\,\text{GeV}$	37	292	258	0	0	0	0	1.52

Table 4.1: Cutflow of several SM processes by assuming an integrated luminosity of $3\,000\,\text{fb}^{-1}$. The cuts are meant to add up from top to bottom. Hereby, the $\sum p_{\mathrm{T}}^{\tau^{\text{truth}}_{i=1-4}} > 250\,\text{GeV}$ is an exception as this cut is only applied in the forth row and then replaced by the $\sum p_{\mathrm{T,\,vis}}^{\tau^{\text{truth}}_{i=1-4}} > 200\,\text{GeV}$ cut. Further explanations are given in the text.

[1] "Truth level" refers to the exploitation of underlying truth information in Monte Carlo generated events.

Unfortunately, a full truth level study is not possible without much more technical effort, as some cuts on reconstruction level have already been placed on the used DAOD[2] datasets. These DAOD cuts are summarized in table 4.2. Only events which fulfil these requirements are available for this analysis. Therefore, at least one (reconstructed) tau candidate and either a light lepton (HIGG4D2) or another tau candidate (HIGG4D3) are demanded. They further have to fulfil some momentum requirements and other properties to ensure their reconstruction quality. The ratio of events which fulfils these requirements spans from 73.37 % for the di-Higgs signal to 8.46 % for associated single vector boson and jet production (cf. table 4.4 in section 4.3). The amount of expected taus to be lost in the $hh \to 4\tau$ channel is cross checked with the acceptance study in section 5.1 and is in agreement with that.

	HIGG4D2	HIGG4D3
object related	$N_\tau \geq 1, N_e + N_\mu \geq 1$	$N_\tau \geq 2$, $N_e + N_\mu = 0$
momentum related (cuts in GeV)	$p_T^\tau \geq 23$ and ($p_T^\mu \geq 12$ or $p_T^e \geq 15$) or $p_T^\tau \geq 18$ and ($p_T^\mu \geq 18$ or $p_T^e \geq 22$)	$p_T^{\tau_1} \geq 33$ and $p_T^{\tau_2} \geq 23$
object identification related	e = Medium or μ = Good	τ_1 = Loose or τ_2 = Loose
tau related	$N^\tau_{\text{tracks}} = 1$ or 3, $\lvert q_\tau \rvert = 1$	$N^{\tau_{1,2}}_{\text{tracks}} \in [1, 8]$, $\lvert q_{\tau_{1,2}} \rvert = 1$

Table 4.2: The two different DAOD types are presented which are used in this thesis. They differ in the different requirements which have to be fulfilled for events to be written in the DAOD sample. In order to ensure orthogonality between both samples, in this work the light lepton veto for HIGG4D3 samples (marked in red) was added subsequently. The effect of this selection on each considered process is summarised in table 4.4.

To summarize the partly-truth-level study of table 4.1, in the first row, four or more truth τ leptons are demanded. Next, additionally to that, the tau charges are required to equalize which means that all four taus should be produced in correlated events such as Higgs or Z boson decays into $\tau^+\tau^-$. It can be observed that especially the amount of $Z \to \tau\tau$ and t quark events drops noticeably. This can be explained, as in both processes, only the correlated production of two taus is expected either via $Z \to \tau^+\tau^-$ or via $t\bar{t} \to b\bar{b}\bar{\nu}_\tau\nu_\tau\tau^+\tau^-$. Here, the t quark decays via $t \to b\tau\nu_\tau$ in 7.1 % of all cases [11]. The two additional truth taus probably occur as decay products of additional hadronic jets in the event and do not necessarily have to be charge correlated. As a result, many of these events are discarded by the charge cut.

The veto of b-jets again refers to reconstruction level due to the information which was stored in the used N-tuples. This explains why so many t quark events survive despite the nearly exclusive branching ratio of t quarks into b. Theoretically, on truth level almost all top events should be excluded by this cut.

The further two cuts examine the scalar sum of all four tau p_T. In general the taus emerging from an Higgs boson should carry a higher momentum (at least compared to taus from Z bosons or other lighter particles) due to the high mass of the Higgs. First the truth tau momentum is considered (forth row), afterwards this cut is replaced with the more applicable visible tau p_T. The latter cut is much harder as a significant ratio of the momentum is carried away by the ν_τ. As a result, the significance drops due to the loss of signal events and further decreases after applying additional cuts on either the ΔR of each tau pair or their invariant mass. However, these cuts are thought to play an important role in the analysis on

[2] The data recorded at the ATLAS detector is stored in data files of the *Analysis object data* (AOD) type. Their content is referred to as the xAOD data format. As the original data files are too big to analyse directly (usually several Petabytes), they are reduced according to the needs of each physics group. The resulting *Derived AODs* (DAODs) still contain information in the xAOD data format, but they are much smaller (Terrabyte scale). They are further processed by the xTauFramework in this thesis which creates N-tuples of the order of Gigabyte in which a very selective amount of information is stored.

reconstruction level and (at least the scalar sum of all tau momenta) ensures that the event has a higher probability to be successfully identified.[3]

Furthermore, this quick study is not thought to examine the highest possible significance which might be achieved on truth level. Instead, it is proven that a few very straight-forward cuts already lead to significances which are well suited to significantly improve the overall measurement of di-Higgs production at the ATLAS detector. Especially after applying all cuts (last row), it is interesting to note that nearly all background can be reduced to zero. The remaining events are single Higgs events or part of "Multiboson". The latter one includes the production of two or three vector bosons (W or Z) or the more general Sherpa 4ℓ and Sherpa $3\ell 1\nu$ samples (cf. table A.6 in the appendix). It is natural to expect many truth taus if multiple vector bosons are involved. However, the vector boson masses fall below the Higgs mass and thus it is much more likely for at least one of their decay products to undershoot the momentum threshold for identification. Indeed, the $ZZ \to 4\tau$ process was investigated quite detailed in this analysis[4] and it was found that di-Z events are only roughly ten times more frequent than di-Higgs events after complete identification of four leptons even though the cross section of $ZZ \to 4\tau$ is 161 times higher.

To sum up this section, it is proven that the $hh \to 4\tau$ process is very interesting on (partly) truth-level. Therefore, this work is dedicated to the task to examine the possible achievable significance of this signal on reconstruction level. Obviously, without having access to truth-level information, it is much harder to single out signal events from the background. The analysis strategy to do so is presented in the following section.

4.2 Analysis Strategy

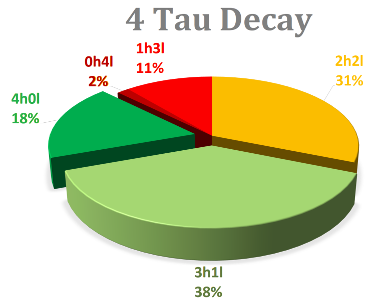
The $hh \to 4\tau$ channel can be split into several sub-channels depending on their specific tau decays. As already stated in section 3.3.2, τ leptons have a 35 % probability to decay into electrons or muons. Only in the remaining 65 % of the cases, the tau decays hadronically and appears as a reconstructed tau in the ATLAS detector.

The naming scheme of the sub-channels refers to the number of hadronically and leptonically decaying taus. For example, in $3h1\ell$, three τ leptons decay hadronically and one leptonically. Correspondingly, at least three tau candidates and exactly one light lepton will be demanded in this channel. The other channels are named analogously. In figure 4.1, the probabilities of the different channels and the expected amount of events in each channel is shown.

In most signal events, one or two light leptons are produced (38 % or 31 % of all four-tau events). The amount of four hadronically decaying taus is also considerably high (18 %). A main aim of this work is to develop an identification method for multi-tau events. To fully exploit the method's performance, in this thesis the main focus lies on events with many hadronically decaying taus, i.e. the $4h0\ell$ and the $3h1\ell$ channel, especially in chapter 5. Nevertheless, the $2h2\ell$ channel will be investigated in chapter 6 although it might suffer from different background and would need a more specific treatment. Additionally, channels with many light leptons also originate from $hh \to XXWW \to X'X'\ell\nu\ell\nu$ decays. These are already studied in more detail in separate analyses and it might become difficult to guarantee orthogonality in the $2h2\ell$ channel. In this channel, either exactly one electron and one muon or two light leptons of the same charge are demanded to reject most $Z \to \ell^+\ell^-$ events.

[3] This refers to the fact that all taus have to fulfil a certain p_T threshold to be identified.

[4] This was also due to the fact that the $hh \to 4\tau$ signal sample on reconstruction level was not available before mid-January, i.e. only during the last five month of this work.



(a)

Channel	Events
All	380
$4h0\ell$	68.4
$3h1\ell$	144.4
$2h2\ell$	117.8
$1h3\ell$	41.8
$0h4\ell$	7.6

(b)

Figure 4.1: The division of the $hh \rightarrow 4\tau$ process into its five sub-channels is shown in (a). In (b), the corresponding expected event numbers for an integrated luminosity of $3\,000\,\text{fb}^{-1}$ is presented.

Due to the extremely low cross section of $hh \rightarrow 4\tau$ processes, an essential task will be not to lose too many of the rare signal events. Therefore, every event selection has to be reviewed with respect to the amount of signal events which are discarded by each cut.

A huge drawback of the signal channel is the creation of at least four neutrinos which leave the detector undetected and carry away energy and momentum. As a result, a significant percentage of the remaining decay products of the τ leptons will fall below the momentum threshold to be successfully identified. Moreover, the solid angle coverage of the detector is limited to $|\eta| < 2.5$ (cf. section 3.2.2) such that any particle is lost that leaves the detector outside of this region. An event can already be classified as insufficient if only one of the four τ leptons could not be identified. Thus, the ratio of taus which is lost during the identification process scales with the power of four. E.g. if the overall reconstruction and identification efficiency of a single τ lepton is 70 %, only in 24 % ($0.7^4 = 0.24$) of all four-tau events, every tau is identified successfully .

It is straightforward to assume that the amount of di-Higgs events which successfully undergo the full identification process is very limited (cf. section 4.3). This is especially true if usual tau identification approaches of ATLAS analyses (cf. section 5.2) are used. The expected tau acceptance as well as the reconstruction and identification efficiencies are studied and possible improvements are outlined.

As the signal events' unique signature can only be fully exploited if all four leptons are successfully reconstructed and identified, a major part of this analysis is the development of a new tau identification method. This methods exploits the shape of multi-tau events instead of identifying each tau on its own and is presented in section 5.3. Its aim is to lose as few signal events as possible while rejecting most of the background without four truth taus.

In chapter 6, further properties of the signal are studied to separate it from background events. For this purpose, the background has to be controlled first. Although the amount of events which produces four truth taus is limited, hadronically decaying τ leptons can be faked by hadronic jets. They can either emerge from other heavy particle decays of the SM or from pure QCD interactions which are very frequent in a hadron-hadron collider like the LHC.

Heavy particles result from hard scattering processes and can thus acceptably be modelled with Monte Carlo simulations (cf. section 3.4.1). As this work can be seen to be the first attempt on four-tau structures, a proper estimation of all SM events which might mimic four taus is crucial. This is outlined in section 6.1. In pure QCD interactions, however, the hadronisation process of all involved particles cannot be treated perturbatively and is extremely difficult to model properly. Unfortunately, this is exactly the process which determines weather a jet is identified as a tau or not. To effectively estimate the extreme case of four QCD jets which are misidentified as tau candidates, a huge amount of mostly uninteresting QCD events would need to be simulated and corrected if their model assumptions are found to be not ideal.

Instead, a data driven fake estimation of four tau-like objects is developed and studied in section 6.2.. Here, phase space regions of the recorded data are utilised in which it is ensured that they almost exclusively contain pure QCD events. With appropriate scaling, they can serve as estimations of these processes in the other phase space regions.

Eventually, a proper signal region is defined by studying and developing a vast amount of observables which are suitable to separate signal from background (cf. section 6.4). In section 6.5 boosted decision trees are trained to improve the significance further. Furthermore, the expected shape of signal and background in the developed signal regions is fitted and an upper limit on the di-Higgs cross section is calculated in section 6.6.

It should be emphasized that this work is meant to study the capability of the four-tau channel to participate in the measurement of di-Higgs production at all. Thus, the analysis is pushed to final significances and statistical interpretations while detailed background studies unfortunately have to be left aside at some points. The drawback of this approach is that several usual tasks of a complete analysis are postponed such as proper estimations of control regions of different background processes or a decent estimation of systematic uncertainties. In the limited time of a Master thesis, the focus of this work is a final estimation of the potential improvement of the measurement of the di-Higgs self-coupling due to the $hh \to 4\tau$ channel. If this is found to be interesting, further detailed studies are motivated.

4.3 Preselection

As described in more detail in section 6.1, many processes are studied (cf. section A.1) with respect to their probability to produce four taus or four tau-like objects (so called fake-taus). As it was already motivated in the previous section, usual tau identification approaches might cut away almost all of the signal. Thus, a very general preselection is designed for the $4h0\ell$ and $3h1\ell$ channel such that the unique event topology can be exploited but as few signal events as possible are lost.

Therefore, the tau identification requirements are generalised as much as possible. In the following, every jet will be considered as a tau candidate, that exceeds a transverse momentum of $p_\mathrm{T} = 20\,\mathrm{GeV}$. Moreover, the jet or tau candidate is required to consist of one or three charged tracks and an overall absolute value of the charge of one. Additionally, only tau candidates are considered which lie in high performance regions of the detector. The tau candidate requirements as well as additional preselection cuts are summarised in table 4.3.

	Selection
tau candidate properties	$p_\mathrm{T} > 20\,\mathrm{GeV}$, $\lvert q\rvert^{\mathrm{total}} = 1$, $N_{\mathrm{tracks}} = 1$ or 3 $\lvert\eta\rvert < 2.5$, $\lvert\eta\rvert \notin [1.37, 1.52]$
required objects	$N_{\tau^{\mathrm{cand}}} \geq 3$, $(N_{\tau^{\mathrm{cand}}} + N_e + N_\mu) \geq 4$
object properties	q_i: 2+2−, all τ^{cand} from same vertex, $\lvert z_\ell^{\mathrm{vertex}} - z_{\tau^{\mathrm{cand}}}^{\mathrm{vertex}}\rvert < 1\,\mathrm{mm}$

Table 4.3: The preselection cuts are presented. If not stated differently, they are applied to all samples.

In summary, three or more tau candidates are required. Moreover, at least four leptons in total, i.e.

either a fourth tau or one light lepton are demanded and their charges have to equalize to ensure their physical pair production. Additionally, in the $hh \to 4\tau$ process, all four leptons are produced in the same vertex. Thus, it is ensured that the tau candidates share the same vertex. As the lepton is thought to originate from a leptonically decaying tau, its z component is demanded to be close to the multi-tau vertex. Although these cuts are designed to be as general as possible, mainly due to tau acceptance problems (cf. section 5.1 for more details), a significant ratio of the signal is already lost after these requirements. In table 4.4, the effect of the preselection cuts as well as the `DAOD` cuts (cf. table 4.2) on different processes is summarized.

	$hh \to 4\tau$	Higgs	Multiboson	V + Jets	$Z \to \tau\tau$	Top	$t\bar{t}$
			Events after `DAOD` cuts				
Ratio	73.37 %	22.97 %	19.21 %	8.46 %	8.64 %	25.37 %	30.54 %
Events	268	$4.1 \cdot 10^6$	$6.0 \cdot 10^7$	$1.1 \cdot 10^9$	$5.4 \cdot 10^8$	$1.2 \cdot 10^8$	$7.4 \cdot 10^8$
			Preselection cuts				
$N_\tau \geq 2$	71.99 %	21.33 %	29.29 %	10.74 %	50.76 %	41.01 %	45.09 %
$N_\tau \geq 3$	29.55 %	3.48 %	5.10 %	1.00 %	6.57 %	9.80 %	12.81 %
$N_{\text{lep}} \geq 4$	16.86 %	2.00 %	2.98 %	0.91 %	1.31 %	5.80 %	8.98 %
Vertex	15.48 %	1.23 %	2.09 %	0.37 %	0.80 %	4.61 %	7.50 %
q_i: 2+2−	11.30 %	0.65 %	0.97 %	0.20 %	0.29 %	2.00 %	3.46 %
Events	30.2	$2.7 \cdot 10^4$	$5.8 \cdot 10^5$	$2.1 \cdot 10^6$	$1.6 \cdot 10^6$	$2.3 \cdot 10^6$	$2.6 \cdot 10^7$

Table 4.4: In the first block, the effect of the `DAOD` cuts (cf. table 4.2) are summarised. The ratios of events that survive the `DAOD` selection are presented with respect to the overall amount of expected events in each process due to their cross sections. The total amount of the surviving events in the `DAOD` files is presented in the second row. This amount is available for further studies. Hence, the ratios of surviving events due to preselection cuts (cf. table 4.3) refer to that number. They are presented in the second block. For example, after demanding $N_\tau \geq 2$, 71.99 % of the 268 $hh \to 4\tau$-events which passed the `DAOD` selection is left. To get the total ratio of accepted events in each process, the ratios of the second block thus have to be multiplied with their value in the first row (in the named example, this would be $0.7199 \cdot 0.7337 = 0.5282$).
In this table, N_{lep} summarises the total amount of leptons, i.e. tau candidates and light leptons. The requirement to have at least four leptons which share the same vertex (or light leptons close to it) is abbreviated with "Vertex". In all cases, an integrated luminosity of $3\,000\,\text{fb}^{-1}$ is assumed.

As expected, only a tiny fraction of background events fulfils the combined requirements. Unfortunately, also the amount of signal events scales down to $0.73 \cdot 0.11 = 0.08$, i.e. only 8 % of all $hh \to 4\tau$ processes are successfully reconstructed in the $4h0\ell$ or $3h1\ell$ channel. Furthermore, the amount of background events still exceeds the signal by far. Sophisticated methods to separate these processes are thus needed and are developed in the following chapters.

CHAPTER 5

Tau Identification

The exceptional property of the $hh \rightarrow 4\tau$ signal channel is the production of four τ leptons. Thus, it is obvious that a decent tau identification is crucial to adequately separate the named signal from possible background. The ratio of taus which are accepted and successfully reconstructed is estimated in section 5.1. The classical tau identification approach, which is introduced in section 5.2 rejects an even higher ratio of truth τ leptons. As every tau loss scales with the power of three or four in the signal regions of this analysis, a significant amount of signal would already be lost before any other selection cut. Thus, a novel tau identification method is developed and introduced in section 5.3. It is based on the classical tau jet BDT score (cf. section 5.2) and is further developed (sections 5.3.1, 5.3.2, 5.3.3) such that significantly improved signal-to-background ratios are achieved (section 5.3.4).

5.1 Tau Acceptance and Reconstruction

In order to enable the identification of τ leptons (cf. section 5.2), they have to fulfil several requirements. A unique detector response is needed to differentiate them from jet traces in the detector.

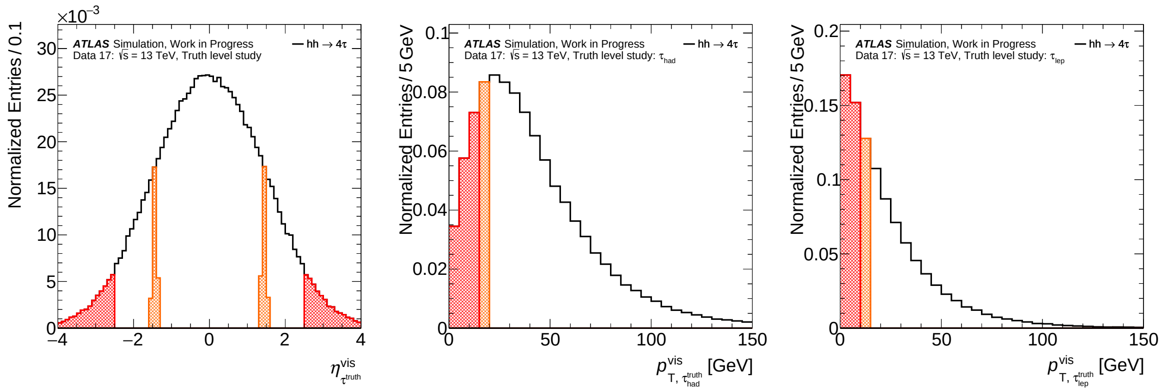
As discussed in section 3.2.2, the pseudorapidity region $|\eta| > 2.5$ is not covered by the Inner Detector. Accordingly, no track information is available and no τ leptons can be reconstructed in this region. Moreover, in the region of $|\eta| \in [1.37, 1.52]$, the calorimeter response suffers from a crack region in which cables and cryostats are placed (cf. section 3.2.3). Hence, usually particles in this region are discarded. Additionally, only taus with a sufficiently large transverse momentum can be distinguished from jets. The official recommendations demand for the provided scale factors a p_{T} higher than 20 GeV [96]. The lower p_{T} region suffers from a large pile-up background as many QCD processes take place with low transverse momentum transfer (cf. section 2.2).

Taking both into account, several taus are lost before they can be reconstructed as tau candidates. To estimate the amount of τ leptons which do not fulfil the detector requirements in the $hh \rightarrow 4\tau$ channel, the `DAOD` signal sample is only partially appropriate. As the `DAOD` cuts (cf. table 4.2) already require at least one reconstructed tau, truth taus in `DAOD` samples are enriched in the "allowed" regions of reconstructable taus. As a result, the amount of truth taus in the rejected parameter space is slightly underestimated. To avoid this disadvantage, a truth-level data sample is produced for the $hh \rightarrow 4\tau$-channel.[1] This sample

[1] For this purpose, `MC@NLO` [103] is used as the event generator and the particle showering is modelled with `Herwig` [104]. More detailed comparisons between the created truth-level sample and the more sophisticated official ATLAS reconstruction sample revealed that they do not agree well in all properties. In detail, the boosting of the di-Higgs system was modelled more properly in the ATLAS sample which also leads to higher tau p_{T} values. Accordingly, the truth-level sample slightly overestimated the amount of truth taus which do not reach the need momenta. In summary, neither the truth-level sample nor

was produced without any preselection and hence is more suitable for the following considerations.

In figure 5.1, the amount of taus outside the preferred $|\eta|$ or p_T region is visualised. The relative loss due to each effect is further summarised in table 5.1. The corresponding plots for the DAOD files can be found in figure A.1 in the appendix. Figure A.2 also shows the distributions for $ZZ \to 4\tau$ events.



(a) (b) (c)

Figure 5.1: Fraction of reconstructable truth taus in the $hh \to 4\tau$ channel with respect to their pseudorapidity (a) or their transverse momenta. The latter one is divided into hadronically (b) and leptonically (c) decaying taus. Taus in the red region are not reconstructable or impossible to identify. Taus in the orange region are very hard to successfully identify. However, they might be accessible in general. Further explanations are given in the text.

Cut	τ_had	τ_lep
$p_\mathrm{T}^\mathrm{vis} > 10\,\mathrm{GeV}$	-	72 %
$p_\mathrm{T}^\mathrm{vis} > 12\,\mathrm{GeV}$	-	67 %
$p_\mathrm{T}^\mathrm{vis} > 15\,\mathrm{GeV}$	86 %	60 %
$p_\mathrm{T}^\mathrm{vis} > 20\,\mathrm{GeV}$	79 %	51 %
$\lvert\eta\rvert < 2.5$	93 %	93 %
$\lvert\eta\rvert \notin [1.37, 1.52]$	95 %	95 %
$p_\mathrm{T}^\mathrm{vis}$ above threshold, valid $\lvert\eta\rvert$	70 %	53 %

Table 5.1: Amount of truth taus in the $hh \to 4\tau$ process that fulfil specific requirements which are needed for reconstruction and identification. For hadronically decaying taus a p_T threshold of 20 GeV is assumed in the last row, while for leptonically decaying taus 15 GeV are taken. An average of a truth-level sample and the DAOD sample of this process was taken. Further explanations are given in the text.

A large fraction of 21 % of the hadronically decaying taus already does not fulfil the $p_\mathrm{T} > 20\,\mathrm{GeV}$ requirement. This limit is mainly caused by pile-up corrections. Within the tau candidate's ΔR cone, it is not unusual to have an overall energy deposition of up to 5 GeV in the calorimeters due to pile-up processes. Hence, the tau candidate's energy calibration becomes very challenging for events with momenta lower than 20 GeV even with advanced reconstruction techniques like particle flow [112, 113]. Nevertheless, tau candidates with lower momenta are accessible in principle. Even though their exact

the incomplete DAOD ATLAS sample describe the correct fraction of lost taus. However, the fractions of taus in the studied regions are similar in both samples. Hence, an average of both samples is taken as it is expected that the overestimation and underestimation of the respective samples roughly equalise. The conclusions of this section do not change if the considered values fluctuate by $O(1\,\%)$.

energy and momentum comes with higher uncertainties, they might be worth to consider in the $hh \to 4\tau$-channel. Especially the special structure of the $h \to \tau\tau$-decay might be helpful as every low-energy tau usually comes with a high p_T tau (with lower relative uncertainties) which was boosted in the same direction as the Higgs. Thus, taus up to a transverse momentum of 15 GeV might be considerable in this analysis. Accordingly, this region is coloured in orange in figure 5.1.

For leptonically decaying taus the ratio of leptons exceeding high momenta is even smaller. This is natural as an additional neutrino emerging in this decay will carry a considerable amount of energy as well. Fortunately, electrons and especially muons can be reconstructed with high efficiencies up to lower momenta of usually 12 GeV for muons or 15 GeV for electrons. Going lower than these values in future might also allow a considerable gain of signal events.

Moreover, taus which decay in the eta-crack region of $|\eta| \in [1.37, 1.52]$ pass the Inner Detector and their tracks are perfectly reconstructed. Although their assigned energy information of the calorimeters will lead to wrong values, the track information of these taus might be enough to use them for this analysis. This is supported by the fact that the energy resolution due to the track curvature (assuming pion masses for each track) is higher than the one of the calorimeters up to energies of roughly 100 GeV. Hence, it might be possible to allow at least parts of the tau leptons to decay in the eta-crack region if further signal events should be saved.

These signal enhancing possibilities are further discussed in section 6.4 and 6.5. In this chapter, tau candidates with $p_T > 20$ GeV in unspoilt eta regions are considered. This is fulfilled for a fraction of 70 % of all truth τ_{had} in the signal channel of this work (cf. table 5.1).

In general, to select these taus as possible tau candidates, their tracks have to be reconstructed unambiguously. Here, the main problem is the correct track classification for each tau candidate as several pile-up tracks pollute the tracking detector. As a result, an additional *reconstruction efficiency* of roughly 85 % for 1-prong taus and 70 % for 3-prong taus was measured [114]. Hence, in total only roughly 60 % (1-prong) or 49 % (3-prong) of all hadronically decaying taus in the $hh \to 4\tau$ channel are reconstructed in such a way that they are available for further studies.

Leptonically decaying taus are reconstructed as light leptons in the ATLAS detector. These leptons are considered in this analysis in the $3h1\ell$ and $2h2\ell$ channel and due to the `HIGG4D2` derivation cuts, muons with $p_T > 12$ GeV and electrons with $p_T > 15$ GeV are demanded. In the di-Higgs process, on average 56 % of all leptonically decaying taus fulfil this requirement together with the $|\eta|$ constraint. Fortunately, the reconstruction efficiencies for light leptons are much higher (cf. section 3.3).

5.2 Classical Tau Identification Approach

The predominant difficulty of identifying τ leptons after their successful reconstruction is to distinguish them from quark or gluon initiated jets. Especially quark jets tend to be quite collimated (compared to gluon jets) and thus are likely to fake the typical τ_{had} decay properties. This can be observed in figure 5.2 in which the distribution of tau candidates is depicted after applying several preselection cuts like $p_T > 20$ GeV (for further details see [96]). Although this figure should not be taken quantitatively in a general case[2], it serves as an educative example of the distribution of jets and light leptons in tau candidates depending on their number of tracks.

A sufficient separation power is gained by using a multivariate approach based on boosted decision trees as described in section 3.4.2. In this approach, 1-prong and 3-prong tau candidates are trained in separate BDTs. $Z/\gamma^* \to \tau\tau$ serves as the signal and dijet events (selected from data) are taken as

[2] E.g. in this figure, the $Z \to \tau_\mu\tau_{had}$ is studied. Muons are less likely to fake taus than electrons and hence, the ratio of leptonic fakes is underestimated in this figure.

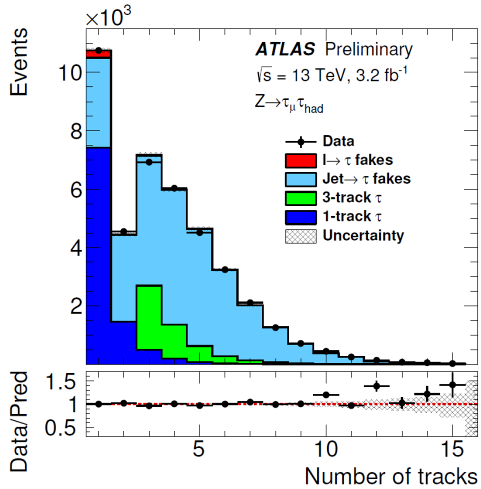


Figure 5.2: The origin of tau candidates in dependence of the track multiplicity is shown. Here, tracks in the core and isolation region are considered and several preselection cuts are applied which is explained in more detail in [96]. The number of tracks shown here is not equal to the number of tracks which will be used in the further analysis. However, this plot serves as a nice example of the contamination of jet fakes in the tau reconstruction (taken from [96]).

background events. The discriminating variables used as the BDT input are listed below and are directly taken from [95]. Three of them are exemplarily shown in figure 5.3.

- **Central energy fraction (f_{cent})**: The fraction of transverse energy deposited in the centre region $\Delta R < 0.1$ compared to all energy deposited in the core region $\Delta R < 0.2$ around the τ_{had} candidate. Pile-up contributions are corrected based on μ, the average number of pile-up interactions per bunch crossing.

- **Leading track momentum fraction ($f^{-1}_{\mathrm{leadtrack}}$)**: The transverse energy deposited in all cells belonging to topological clusters in the core region divided by the transverse momentum of the highest-p_{T} charged particle in the core region.

- **Track radius ($R^{0.2}_{\mathrm{track}}$)**: p_{T}-weighted ΔR distance of the associated tracks in the core region to the τ_{had} direction.

- **Leading track IP significance ($|S_{\mathrm{leadtrack}}|$)**: Absolute value of the transverse impact parameter of the highest-p_{T} track in the core region divided by its estimated uncertainty.

- **Fraction of tracks p_{T} in the isolation region ($f^{\mathrm{track}}_{\mathrm{iso}}$)**: Scalar sum of the p_{T} of tracks of the τ_{had} candidate in the isolation region divided by the sum of the transverse momenta of all tracks associated to the τ_{had} candidate.

- **Maximum ΔR (ΔR_{Max})**: The maximum ΔR between a track in the core region associated with the τ_{had} candidate and the τ_{had} direction.

- **Transverse flight path significance ($S^{\mathrm{flight}}_{\mathrm{T}}$)**: The x-y-plane distance of the vertex reconstructed from the tracks associated with the core region of the τ_{had} candidate (secondary vertex) with respect to the reconstructed tau vertex, divided by its estimated uncertainty. This observable is only defined for 3 prong τ_{had} candidates.

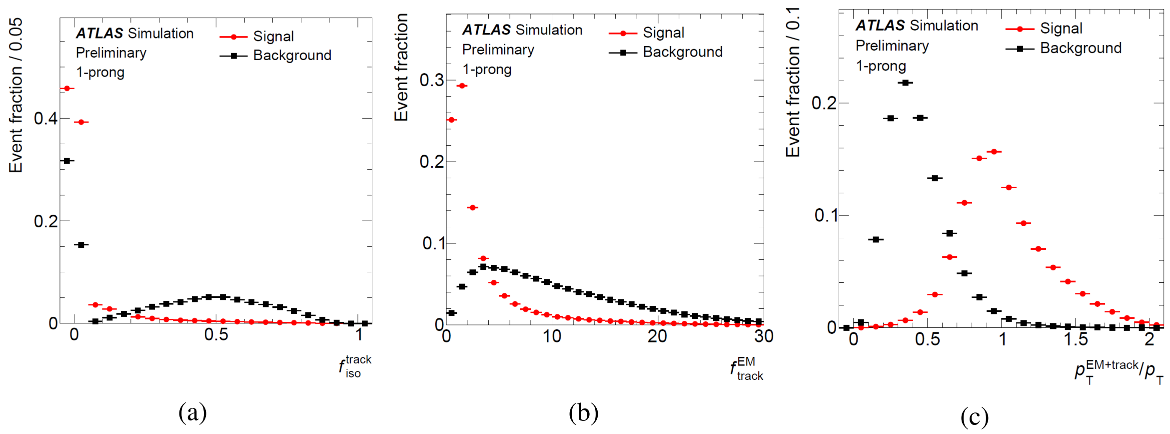
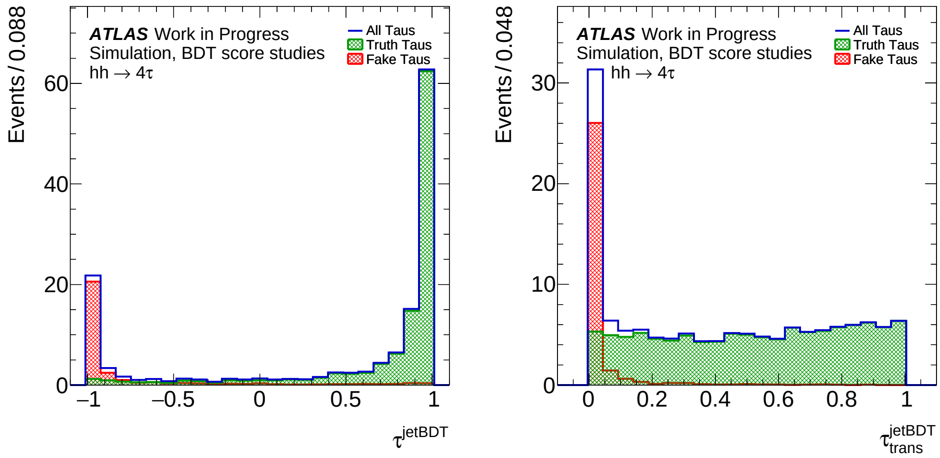


Figure 5.3: Exemplary discriminating variables for the tau jet BDT score for tau leptons (red circles) and multi-jet background (black squares). In general, their discriminating power is less for 3 prong taus (taken from [95]).

- **Track mass (m_{track})**: For each track a pion mass is assumed and the invariant mass of the sum of all four-momenta of the core- and isolation-region tracks is calculated.
- **Fraction of EM energy from charged pions ($f_{\text{EM}}^{\text{track–HAD}}$)**: The difference between the sum of the momentum tracks in the core region and the sum of cluster energies deposited in the hadronic calorimeter divided by the sum of energies deposited in the electromagnetic part of each cluster associated to the τ_{had} candidate.
- **Ratio of EM energy to track momentum ($f_{\text{track}}^{\text{EM}}$)**: The sum of energy deposited in the electromagnetic part of each cluster associated to the τ_{had} candidate divided by the sum of momenta of the core-region tracks.
- **Track-plus-EM-system mass ($m_{\text{EM+track}}$)**: The invariant mass of all tracks and up to two most energetic electromagnetic clusters in the core region. The four-momentum of an EM-cluster is calculated assuming zero mass.
- **Ratio of track-plus-EM-system to p_{T} ($p_{\text{T}}^{\text{EM+track}}/p_{\text{T}}$)**: Ratio of the vector sum of track momenta and up to two most energetic EM clusters in the core region to the calorimeter-only measurement of the momenta.

The resulting BDT score is further corrected such that its efficiency is relatively flat with respect to the transverse momenta of the tau candidates and the pile-up of the event. The resulting score τ^{jetBDT} runs from -1 for very jet-like objects to $+1$ for very tau-like objects as displayed in figure 5.4(a). Moreover, a second BDT score is created as follows. The original BDT score is transformed such that its values directly correspond to the *identification efficiency* of true τ leptons, i.e. a value of 0.3 of the transformed BDT score $\tau_{\text{trans}}^{\text{jetBDT}}$ directly implies that 30 % of the reconstructed truth taus do not exceed this value [115]. Naturally, this score runs from 0 (all truth taus included) to 1 (no truth tau above this value) and is shown in figure 5.4(b).

Three working points are provided for the taus depending on their overall reconstruction and identification efficiency. These working points *Loose*, *Medium* and *Tight* are aiming for *signal efficiencies* (the product of identification efficiency and reconstruction efficiency) of 0.6, 0.55 and 0.45 respectively for 1-prong taus and 0.5, 0.4 and 0.3 for 3-prong taus [95, 96]. Effectively however, they are defined directly by means of the $\tau_{\text{trans}}^{\text{jetBDT}}$ score. Here, every 1-prong tau candidate exceeding a value of 0.15, 0.25 or 0.4 is



(a) (b)

Figure 5.4: The tau jet BDT score τ^{jetBDT} (a) and its transformed value $\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$ (b) are shown for the $hh \to 4\tau$ process. In (a) it is nicely observable that fake taus peak at -1 while truth taus peak at $+1$. In (b), the truth tau behaviour in dependence of the transformed BDT score is flat, which is directly implied by the definition of the score.

labelled *Loose*, *Medium* and *Tight*, respectively. For 3-prong tau candidates, these values read 0.25, 0.4 and 0.5.

The ATLAS tau working group provides p_{T} correction scale factors and uncertainties for these working points. Therefore, in usual ATLAS analyses, it is recommended to demand the considered τ leptons to fulfil one of the named working points such that Monte Carlo simulations can be scaled appropriately.

However, the di-Higgs signal of this work suffers from extremely low production rates and produces a total of four τ leptons. By using the usual *Loose* working point, the probability to successfully reconstruct and identify every tau would roughly be $0.6^4 = 0.13$ (for three taus, the calculation reads $0.6^3 = 0.22$ while the efficiency loss due to the light lepton from the τ_{lep} decay is neglected).

Hence, this analysis will not demand the tau candidates to be *Loose*, *Medium* or *Tight*. Instead, the BDT values are combined to serve as a new, more powerful separation observable for multi-tau events than each BDT value on its own. This approach is outlined in the following section.

5.3 Novel Identification Method for Multi-Tau Events

In section 5.1, the acceptance and reconstruction efficiencies for 1-prong and 3-prong τ leptons are determined. Using these results, a reconstruction efficiency of about $0.70 \cdot (0.74 \cdot 0.85 + 0.23 \cdot 0.70) = 0.55$ for hadronically decaying taus is achieved.[3] Expecting 68 events in the $4h0\ell$ channel and 146 events in the $3h1\ell$ channel, a successfull reconstruction efficiency of all taus is expected to be $0.55^4 = 0.091$ or $0.55^3 \cdot 0.56 = 0.093$ respectively.[4] This leaves 6.1 reconstructable events in the $4h0\ell$ channel and 13.6

[3] Only 70 % of all hadronically decaying taus fulfil the acceptance requirements. 74 % of all hadronic tau decays are 1-prong decays and 85 % of the 1-prong taus are successfully reconstructed (cf. section 5.1). Another 23 % of all taus decay via 3-prong channels for which a reconstruction efficiency of 70 % is achieved.

[4] For leptonically decaying τ leptons, an acceptance of 56 % was deduced as they equally decay into electrons and muons and either 15 GeV or 12 GeV is demanded (cf. table 5.1). Further, an unrealistically high reconstruction efficiency of 100 % was assumed for light leptons.

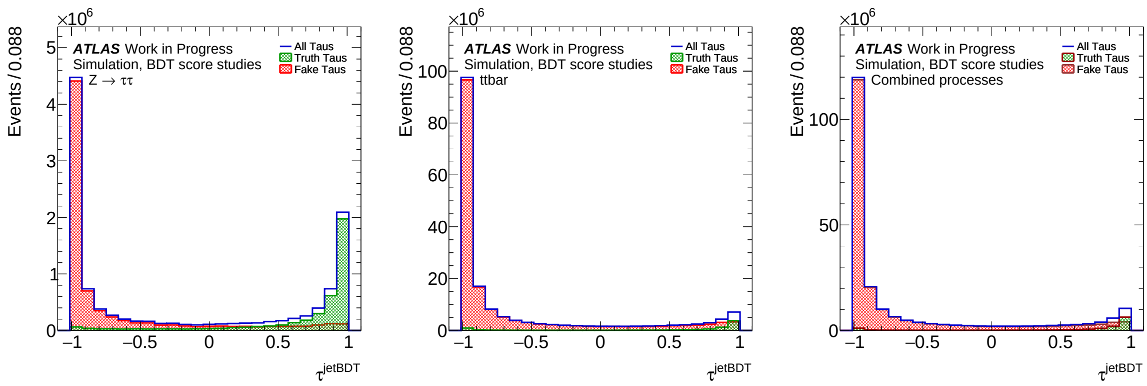
events in the $3h1\ell$ channel without any further cut.

Thus, it is clear that only very delicate cuts can be applied such that as few signal events as possible are lost. A requirement of all taus to fulfil the *Loose* working point would add another efficiency factor of $(0.74 \cdot 0.85 + 0.23 \cdot 0.75)^{4(3)} = 0.41(0.51)$.[5] To actually see some SM di-Higgs events in the 4τ channel, the *Loose* requirement hence can hardly be applied.

Instead, a novel method to distinguish multi-tau events from fake-tau enriched ones is developed. For this purpose, the tau jet BDT values are analysed further as described in section 5.3.1. To gain as high separation power as possible, many additional tau properties are considered in section 5.3.2, before the new method is applied (section 5.3.3) and its gained significances are presented (section 5.3.4).

5.3.1 Truth Tau Likelihood and Probability

The tau candidate requirements as defined in section 4.3 induce many fake-taus which originate from QCD jets. Hence, a separation of truth taus and jets is needed and will be based on the original tau jet BDT scores τ^{jetBDT} and $\tau^{\text{jetBDT}}_{\text{trans}}$ as introduced in section 5.2. In a first step, the distribution of truth and fake taus with respect to the jet BDT values is considered as shown in figure 5.4 for the $hh \to 4\tau$ process. This is not only done for the signal MC sample but for every sample which possibly pollutes the signal region of this work[6] and can look quite different for each sample depending on the samples chance to actually produce four truth taus (cf. figure 5.5(a), 5.5(b) and A.3 in the appendix).



(a) $Z \to \tau\tau$ (b) $t\bar{t}$ production (c) all samples combined

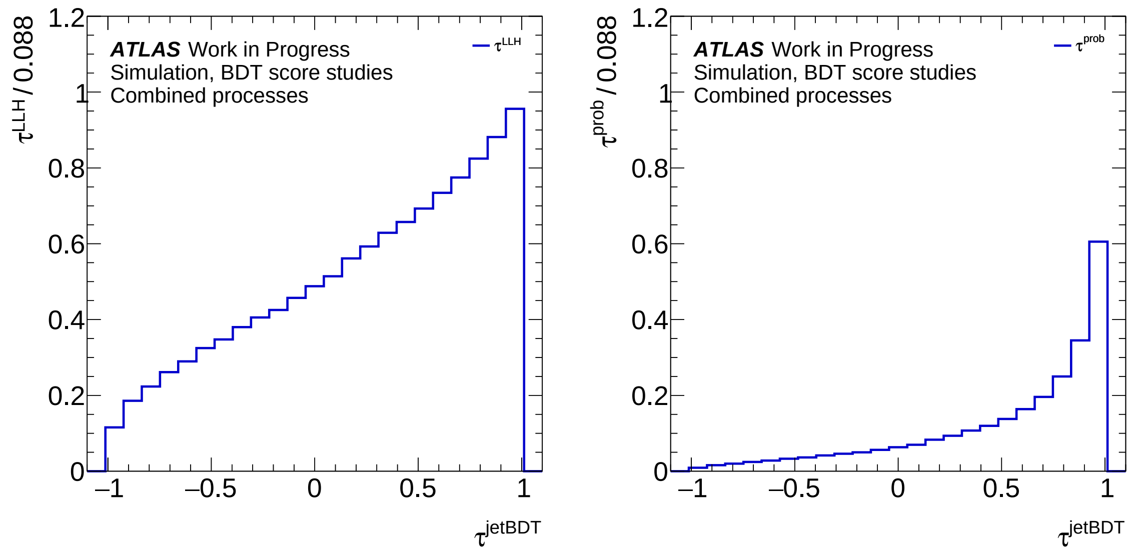
Figure 5.5: The distribution of truth and fake taus for different production modes. While in the $Z \to \tau\tau$ process (a), a considerable amount of truth taus is produced, it can be seen that the vast majority of tau candidates in the $t\bar{t}$ production sample (b) originates from jet fakes. As $t\bar{t}$ production by far dominates the amount of events after the preselection (cf. table 4.4), also a combination of all samples (c) is dominated by fakes.

The overall distribution of truth and fake tau candidates of all samples (cf. figure 5.5(c)) with respect to the jet BDT values is used to assign a new value to each tau candidate. For this purpose, the likelihood τ^{LLH} and the probability τ^{prob} of a tau candidate to be a truth tau is calculated for each BDT value as follows:

$$\tau^{\text{LLH}} = \frac{p^{\text{truth}}\left(\tau^{\text{jetBDT}}\right)}{p^{\text{truth}}\left(\tau^{\text{jetBDT}}\right) + p^{\text{fake}}\left(\tau^{\text{jetBDT}}\right)} \quad \text{and} \quad \tau^{\text{prob}} = \frac{N^{\text{truth}}\left(\tau^{\text{jetBDT}}\right)}{N^{\text{truth}}\left(\tau^{\text{jetBDT}}\right) + N^{\text{fake}}\left(\tau^{\text{jetBDT}}\right)} \tag{5.1}$$

[5] This is only true under the assumption that the probability for tau candidates of the di-Higgs process to fulfil *Loose* requirements is comparable to the one of tau candidates of the $Z \to \tau\tau$ process, which was used to define the working points.
[6] They are introduced in detail in section 6.1. A full list of all used samples can be found in the appendix in section A.1.

Here, $p^{\text{truth/fake}}\left(\tau^{\text{jetBDT}}\right)$ is the probability density function of all truth or fake tau candidates with respect to the jet BDT value and $N^{\text{truth/fake}}\left(\tau^{\text{jetBDT}}\right)$ is the total amount of truth and fake tau candidates with that value. The likelihood and the probability in dependence of τ^{jetBDT} are shown in figure 5.6.



(a) Likelihood

(b) Probability

Figure 5.6: The mapped likelihood value τ^{LLH} (a) and the mapped probability value τ^{prob} (b) with respect to the tau jet BDT value are shown. As the overall amount of fake taus is much higher than the one of truth taus (cf. figure 5.5(c)), the probability value is squeezed to lower values. The corresponding plots for the $\tau^{\text{jetBDT}}_{\text{trans}}$ value are shown in figure A.4 in the appendix.

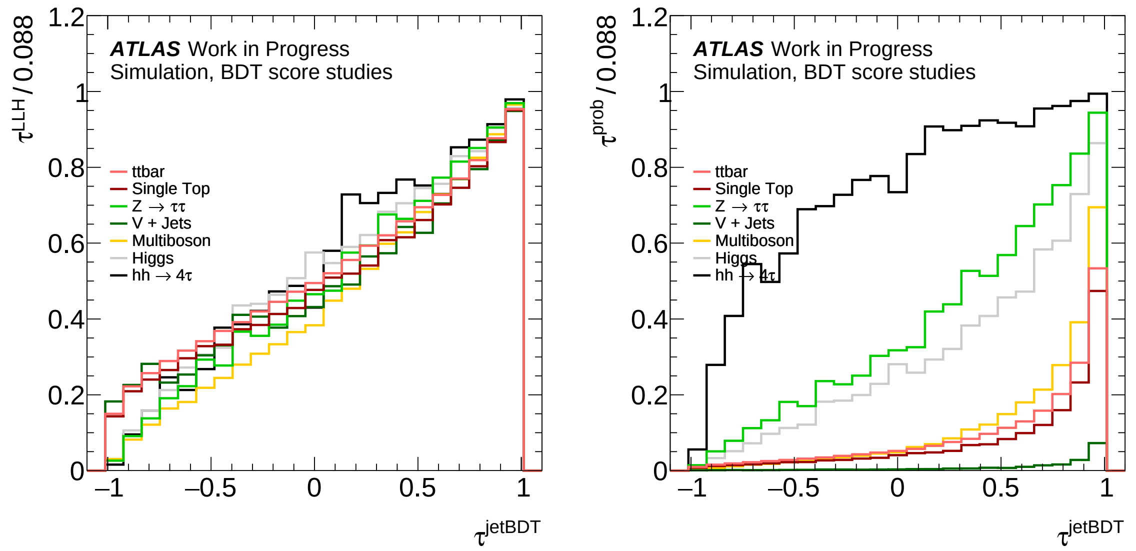
The advantage of τ^{LLH} is that it should be independent of further applied cuts as each distribution (truth or fake) is treated on its own. τ^{prob} is a similar value but it additionally takes into account how many truth taus are produced compared to the fake taus in the considered MC samples. This advantage of information comes with the drawback that the value might fluctuate stronger if the selection of considered samples changes. Thus, in principle it is desirable to use the likelihood τ^{LLH} and to use the information of the overall amount of truth to fake taus additionally. Nevertheless, both alternative approaches are studied in broad detail in the following.

The difference in the values becomes obvious if their behaviour with respect to the underlying process is studied. While the likelihood is roughly similar for all processes, the probability varies significantly due to the different overall ratios of truth and fake taus (cf. figure 5.7). The analogous plots for the $\tau^{\text{jetBDT}}_{\text{trans}}$ value are shown in the appendix in figure A.5.

5.3.2 Sorting Methods to Map the Taus

Although the τ^{jetBDT} score is designed to be relatively flat with respect to several tau properties like their transverse momentum or the pile-up of the event, small dependencies on specific quantities cannot be excluded. Hence, it is worth to study the τ^{jetBDT} behaviour in more detail to maximise the separation power of the defined likelihood or probability of the previous section.

For this purpose, the tau candidates are divided depending on several quantities. First, they are separated depending on their p_{T} in two different ways (`ptsort`, `ptsort2`). It is also studied if their



(a) Likelihood (b) Probability

Figure 5.7: The mapped likelihood value τ^{LLH} (a) and the mapped probability value τ^{prob} (b) with respect to the τ^{jetBDT} value are shown for different processes. While the likelihood is roughly independent of the process (within statistical fluctuations due to the limited statistic in some processes), the probability significantly differs.

performance is changed depending on their eta region (`etasort`). The influence of the number of tracks on likelihood and probability is studied (`prongsort`) and in `bdtsort`, the tau candidates are separated depending on their BDT value. In this sorting scenario, the tau candidate with highest BDT value is always considered separately and a distinct truth and fake distribution is created. The same happens for the tau candidates with second and third highest BDT value. If existing, the remaining tau candidates are collected in a fourth truth and fake distribution. All sorting scenarios are compared to the one, if all taus are taken for one single truth and fake distribution function as explained in the previous section. This "sorting" is referred to as `allinone` in the following.

The different sorting methods are summarised and explained in more detail in table 5.2. As a result, several (mostly four) different τ^{LLH} (cf. figure 5.8 at the end of this section) and τ^{prob} (cf. figure 5.9) distributions are created by the different sorting methods.

As seen in figure 5.8(a) and 5.8(b), the different **likelihood $\boldsymbol{\tau^{\mathrm{LLH}}}$** distributions in dependence of the transverse momentum or pseudorapidity of the tau candidates look very similar. Thus, no real gain of distinguishing these taus is expected compared to the `allinone` method. However, as seen in figure 5.8(e), the tau candidates do show a different behaviour depending on their p_{T} if the τ^{LLH} is calculated with respect to the $\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$ score. In detail, tau candidates with a higher transverse momentum seem to be more likely to be a truth tau than taus with less p_{T} with the same $\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$ score.

By separating the tau candidates with respect to their number of charged tracks (cf. figure 5.8(c) and 5.8(f)), it becomes visible that 3-prong tau candidates with low $\tau^{\mathrm{jetBDT}}_{\mathrm{(trans)}}$ score are more likely to be a truth tau than 1-prong tau candidates. The most diverse distributions can be seen by sorting the tau candidates with respect to their relative τ^{jetBDT} score compared to the other candidates of the event (cf. figure 5.8(d)). Such a strong behaviour can be explained, because this sorting is highly correlated to the underlying process which produces the three or four tau candidates. E.g. a process like $t\bar{t}$-production is very likely to

Cut	ptsort	ptsort2	etasort	bdtsort	prongsort	allinone
1	$p_T < 35$ GeV	$p_T < 25$ GeV	$\|\eta\| < 0.6$	max. τ^{jetBDT}	1-prong	all
2	$p_T \in [35, 60)$ GeV	$p_T \in [25, 40)$ GeV	$\|\eta\| \in [0.6, 1.2)$	second	3-prong	-
3	$p_T \in [60, 90)$ GeV	$p_T \in [40, 70)$ GeV	$\|\eta\| \in [1.2, 1.8)$	third	-	-
4	$p_T \geq 90$ GeV	$p_T \geq 70$ GeV	$\|\eta\| \geq 1.8$	rest	-	-

Table 5.2: The sorting methods to create the τ^{LLH} and τ^{prob} values are shown. E.g. in `ptsort` (and similarly in `ptsort2` and `etasort` with different cuts), all tau candidates with a transverse momentum of less than 35 GeV are handled separately and a truth and fake distribution of them is created. With these distributions, a specific likelihood and probability curve with respect to the $\tau^{\text{jetBDT}}_{\text{(trans)}}$ values is created. The same happens for all tau candidates with $p_T \in [35, 60)$ GeV and so on.
As a result, four different τ^{LLH} and τ^{prob} distributions are created which are shown for different sorting methods in figure 5.8 and 5.9. In the `bdtsort` method, the tau candidates of each event are sorted with respect to their τ^{jetBDT} value. For the candidates with highest τ^{jetBDT} of each event, a separate distribution is calculated. Further, a second and third distribution is created for the tau candidates with second and third highest τ^{jetBDT} values and eventually all remaining tau candidates of the event are collected to form a forth distribution. In the `prongsort` case, only two distributions are created, one for 1-prong and one for 3-prong tau candidates and in `allinone`, only a single distribution is created from all tau candidates independently of their further properties.

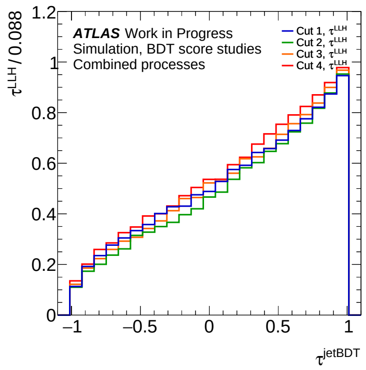
produce many fake taus while only a few truth taus are produced (cf. figure 5.5(b)). In this process, it is much more likely that the highest achieved τ^{jetBDT} score of all tau candidates is still relatively low as all candidates (or at least a significant amount of them) are fakes. This, however, looks very different if four truth taus are produced. The probability that any of these taus looks very tau-like and thus is tagged with a high τ^{jetBDT} score is very high. Hence, events which create two, three or four truth taus are very likely to have at least one tau candidate with a high τ^{jetBDT} score. In contrast to that, mostly events without truth taus are likely to produce the event's highest τ^{jetBDT} score to still be comparably low. Thus, the major amount of tau candidates with a relatively low τ^{jetBDT} score is likely to be a fake tau if this candidate simultaneously has the event's highest τ^{jetBDT} (blue curve).

This discriminating power reverses for the tau candidates which yield the lowest τ^{jetBDT} values (orange and red curve). They are only likely to have relatively high τ^{jetBDT} scores if the event already produces several truth taus such that two or even three tau candidates have higher scores. Hence, if the tau candidate with lowest τ^{jetBDT} score still has a score of e.g. -0.2, it is very likely to be part of an event with many truth taus and correspondingly it is also very likely to be a truth tau itself as well. In summary, the position of a specific τ^{jetBDT} value reveals a lot about the underlying event and thus its absolute value is interpreted very differently depending on the relative position of the value to the other tau candidates.

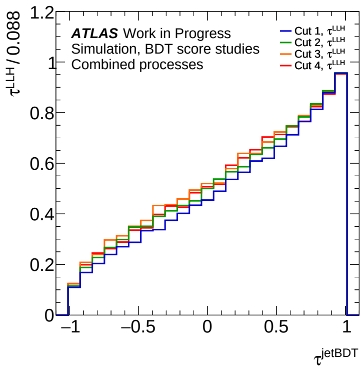
Concerning the **probability** $\boldsymbol{\tau^{\text{prob}}}$ (cf. figure 5.9), a similar qualitative behaviour is observed. Namely, the `ptsort`, `ptsort2` (cf. figure 5.9(a)) and `etasort` method (cf. figure 5.9(b)) reveal that tau candidates only show slight differences in their distribution of truth and fake taus in τ^{jetBDT} with respect to their $|\eta|$ or p_T values.

The sorting method which utilises track information about the tau candidates (`prongsort`, cf. figure 5.9(c)) interestingly shows the opposite behaviour compared to τ^{LLH}. A distinct difference between 1-prong (blue curve) and 3-prong (green curve) tau candidates can be seen if the probabilities are created with respect to the τ^{jetBDT} score. In fact, 3-prong tau candidates are much more often faked by jets and thus their probability to be a truth tau is significantly reduced. This behaviour is expected as hadronic jets tend to produce much more tracks than taus. As a result, the number of hadronic jets with three tracks is much higher than the corresponding amount with only one track.

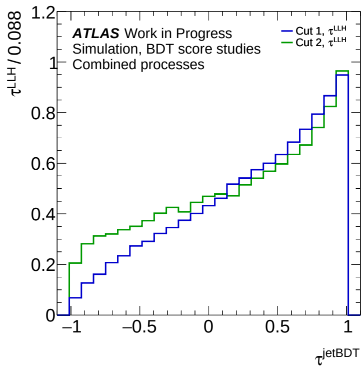
Similarly to the τ^{LLH} distributions, some qualitative differences between candidates with e.g. different transverse momenta can only be observed if the probability is deduced with the $\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$ value (cf. figure 5.9(e)). Moreover, with this score, tau candidates in the outermost $|\eta|$ region tend to be faked more often (cf. figure 5.9(f)).



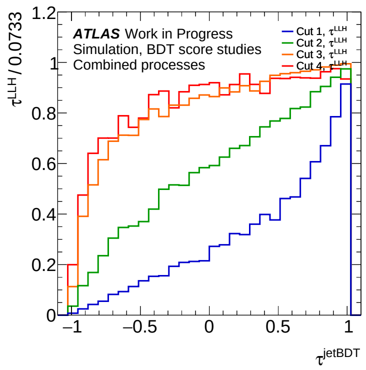
(a) `ptsort`



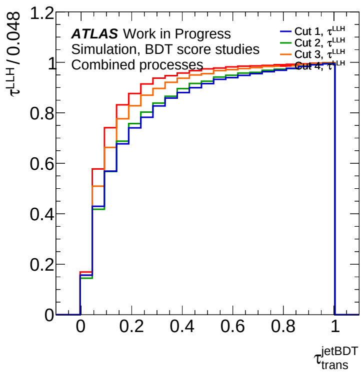
(b) `etasort`



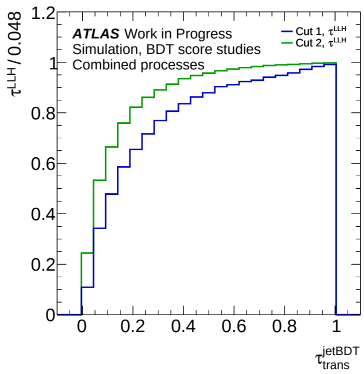
(c) `prongsort`



(d) `bdtsort`

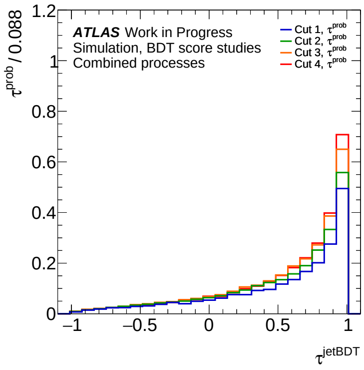


(e) `ptsort`

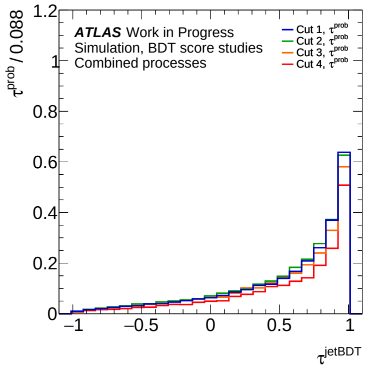


(f) `prongsort`

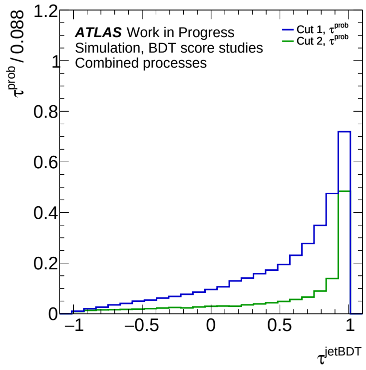
Figure 5.8: The plots show the created τ^{LLH} distributions by the different sorting methods. The different cuts are outlined in table 5.2. In (a) and (b), no significantly different τ^{LLH} shape can be seen. In (c) and especially in (d), however, the likelihood seems to depend significantly on the number of tracks (c) of each tau candidate or which position the candidate has in term of its τ^{jetBDT} score compared to the other candidates in the event (d). (e) and (f) show the τ^{LLH} distribution with respect to the transformed $\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$ score for selected sorting methods. It is worth to note that for this score, a different behaviour in dependence of the p_{T} of the tau candidate can be seen (e) while this was not observable for the τ^{jetBDT} value (a).



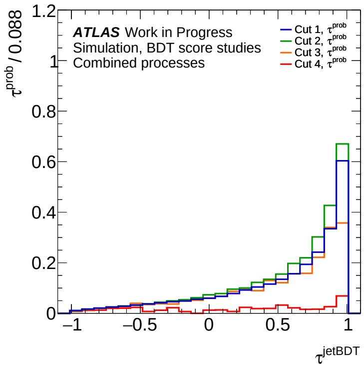
(a) `ptsort2`



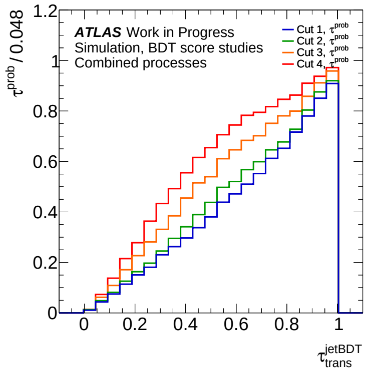
(b) `etasort`



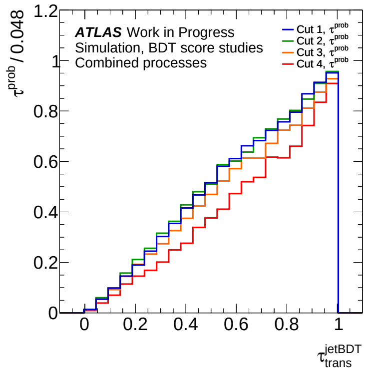
(c) `prongsort`



(d) `bdtsort`



(e) `ptsort2`



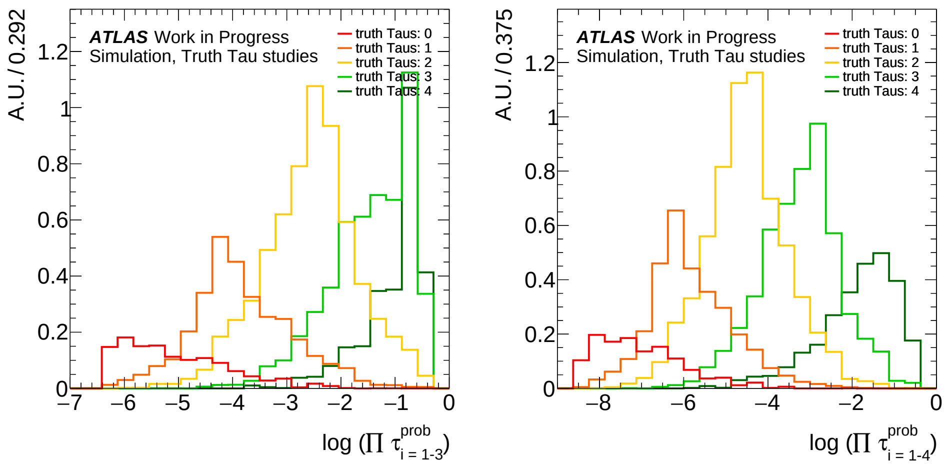
(f) `etasort`

Figure 5.9: The plots show th created τ^{prob} distributions by the different sorting methods. . The different cuts are outlined in table 5.2. Similarly to the τ^{LLH} shapes in figure 5.8, only very slight different τ^{prob} shapes can be observed for both `ptsort` methods or `etasort` (a) and (b). In other sorting methods or the use of $\tau^{\text{jetBDT}}_{\text{trans}}$, different shapes are observable. E.g. in (c) it can be seen that 3-prong tau candidates in green are significantly more probable to be faked than 1-prong tau candidates in blue. To avoid repetitions with no information gain, partly different sorting methods are presented compared to figure 5.8.

5.3.3 Separation of Multi-Truth and Multi-Fake Tau Events

In the novel tau identification approach of this work, the tau candidates are no longer treated separately for their identification. Instead, the shape of an overall multi-tau event is studied and it is tried to separate this from events without many truth taus.

By demanding e.g. four *Loose*-taus in each event, as it might be done by using the classical tau identification method, all events are rejected in which less than four tau candidate's $\tau^{\text{jetBDT}}_{\text{trans}}$ scores exceed the value 0.15. Instead of this demand on four separate $\tau^{\text{jetBDT}}_{\text{trans}}$ scores, the following is done: The $\tau^{\text{LLH/prob}}$ values of each event's tau candidates are multiplied to result in a likelihood- or probability-product $\prod \tau_i^{\text{LLH/prob}}$.[7] While the effective separation power of the single $\tau^{\text{LLH/prob}}$ values to distinguish truth and fake taus compared to the original $\tau^{\text{jetBDT}}_{\text{(trans)}}$ scores might be manageable, their product is expected to separate multi-tau events from background better than single tau cuts (cf. figure 5.10 for the τ^{prob} product and A.6 for the τ^{LLH} product).



(a) τ^{prob} product of three tau candidates (b) τ^{prob} product of four tau candidates

Figure 5.10: The separation power of the logarithm of the τ^{prob} product is shown for the `Sherpa` $\ell\ell\ell\ell$ sample. Events with a high amount of truth taus (light and dark green) peak at higher values while events which are mainly fake dominated peak at the left. The overlap is expected to be less than the truth-fake overlap of the $\tau^{\text{jetBDT}}_{\text{(trans)}}$ scores (cf. figure 5.4).

In detail, the calculated likelihood- and probability-products are used as a threshold value which needs to be fulfilled in every event. Opposite to that, in the classical approach, every tau candidate is considered on its own. Each of them usually is demanded to exceed a certain $\tau^{\text{jetBDT}}_{\text{trans}}$ threshold depending on the chosen working point. Only if each of the tau candidates (in this work mostly three or four) fulfils the $\tau^{\text{jetBDT}}_{\text{trans}}$ requirement, the event is not rejected in the classical approach.

As an example which favours the new method, in a hypothetical four truth tau event, it might be possible that two taus leave a very tau-like trace in the detector and thus gain a high $\tau^{\text{jetBDT}}_{\text{(trans)}}$ score. The two remaining taus, however, are reconstructed worse and both get a medium score[8] which is slightly below the *Loose* threshold. In the classical tau identification approach, only two taus would be identified

[7] In this section the $\tau^{\text{LLH/prob}}$ values for the `allinone` sorting method are used for the sake of simplicity. The differences between the sorting methods are studied in section 5.3.4.

[8] This should not be confused with the *Medium* working point of tau candidates. Working point are highlighted by using italic font.

and the event would be rejected. With the new $\prod \tau_i^{\text{LLH/prob}}$ approach, however, the event might exceed the needed $\prod \tau_i^{\text{LLH/prob}}$ threshold and thus would be "rewarded" for having several $\tau^{\text{LLH/prob}}$ values with at least medium scores. In contrast to that, an event with a very jet-like fake tau would not reach the product threshold, as the whole event is "punished" for one very low $\tau^{\text{LLH/prob}}$ score.

Both methods – first the introduced cut on $\prod \tau_i^{\text{LLH}}$ and second a cut on each single τ^{jetBDT} value – are compared in figure 5.11 and A.7. In this section, the multiplication method is first studied for its general capability to separate events with three or four truth taus from events which mainly consist of fakes. Thus, for now all events in which three (or four) of the tau candidates are truth matched, are labelled as signal events. The corresponding results for the $\prod \tau_i^{\text{prob}}$ are presented in figure A.8 and look very similar. This also accounts for the curves when the transformed $\tau_{\text{trans}}^{\text{jetBDT}}$ score is taken to extract the $\tau^{\text{LLH/prob}}$ values. Hence, all observations which are outlined in the following paragraphs of this section also account for them. Their detailed performance compared to utilising the τ^{LLH} value as well as the significance gain of the different methods is studied in section 5.3.4.

Multiple conclusions can be drawn. First, in the signal process $hh \to 4\tau$ (cf. figure 5.11(a)) as well as in multiboson processes (cf. figure 5.11(b)), hardly any difference between the methods can be seen. In the signal sample, it even seems to worsen, which can probably be explained by fluctuations due to limited statistics. In the multiboson case, a small improvement is visible which seems to be statistically stable. The small changes are probably due to the fact that in these processes, indeed many truth taus are created. Thus, most of the reconstructed tau candidates are featured with a high τ^{LLH} score and maybe one tau candidate falls behind. If it is only one tau candidate that either does not fulfil the τ^{jetBDT} requirement to still be identified as a tau or the very same candidate which holds the $\prod \tau_i^{\text{LLH}}$ value low, the process is equally rejected after the respective thresholds are reached.

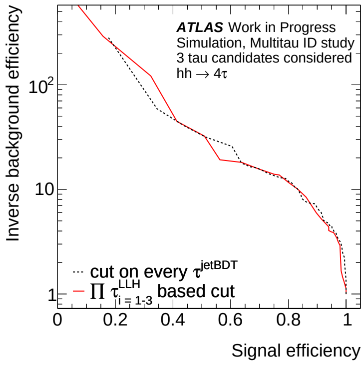
The separation power of the $\prod \tau_i^{\text{LLH}}$ method evolutes if several tau candidates with medium τ^{LLH} scores exist. This can be observed especially in the major background processes of this work, $t\bar{t}$-production (cf. figure 5.11(c)) and $Z \to \tau\tau$ (cf. figure 5.11(d)). Here, the rejection method based on the $\prod \tau_i^{\text{LLH}}$ cut clearly exceeds a cut-based method on single tau candidate properties. By summing over all relevant signal and background processes, this effect dominates as fake-tau processes outnumber events with three or four truth taus.

This is illustrated in figure 5.11(e) in which the $\prod \tau_i^{\text{LLH}}$ based method clearly leads to better results. The promising results duplicate if not only the first three but four tau candidates are considered (cf. figure 5.11(f) for the combination and figure A.7 for separated processes). It can also be observed that the background rejection is much higher when four objects are considered. In these processes, however, the ROC-curves suffer from low statistics as the amount of events with four truth matched taus (which is treated as the signal in the four-tau consideration) is very limited (cf. figure A.7(d) to A.7(f)).

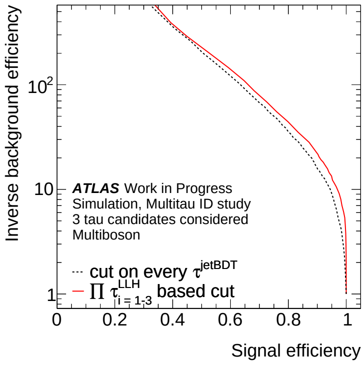
By requiring three or four *Loose* taus (as in the classical identification approach), signal efficiencies of 51 % or 39 % are expected.[9] In these regions, the background rejection is already sufficiently large such that the discrepancy between both methods is not visualised ideally any more. Instead, the signal significances will be studied in more detail in the following section.

It should be emphasised that such a multi-object approach like the $\prod \tau_i^{\text{LLH}}$ method introduced here is interesting in general. The presented figures already illustrate that this method is applicable to separate multi-truth-tau events from fake ones in general. Similar methods might also be interesting to separate e.g. events with many b-jets like in the $hh \to b\bar{b}b\bar{b}$ channel.

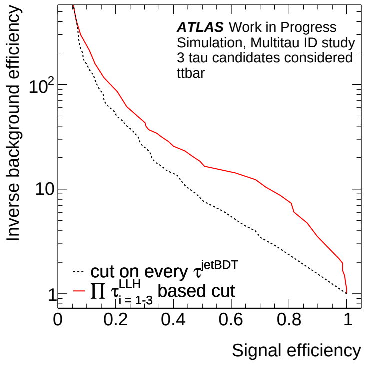
[9] This is simply estimated by the reconstruction efficiencies of *Loose* taus as introduced in section 5.2 via $(0.74 \cdot 0.85 + 0.23 \cdot 0.70)^3 = 0.49$ (with the power of four we get 0.39). The signal efficiency depicted in the visualised plots solely refers to the events which were already successfully reconstructed.



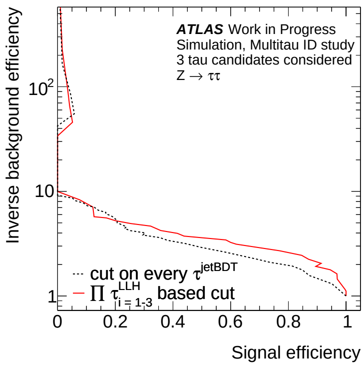
(a) $hh \to 4\tau$, 3 tau



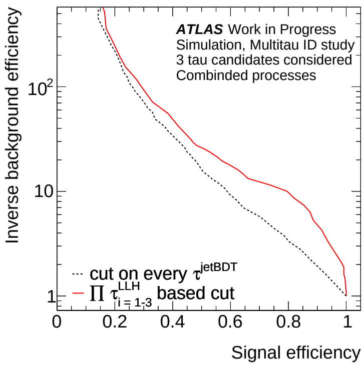
(b) Multiboson, 3 tau



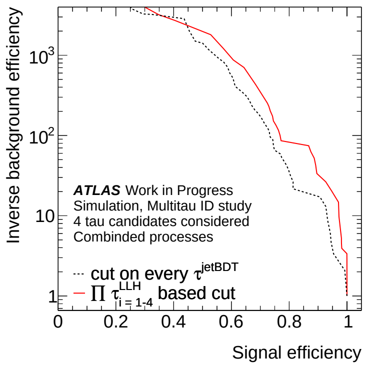
(c) $t\bar{t}$ production, 3 tau



(d) $Z \to \tau\tau$, 3 tau



(e) Sum of all samples, 3 taus



(f) Sum of all samples, 4 taus

Figure 5.11: The background rejection of two separation methods with respect to the signal efficiency is shown for several processes (ROC curves). In the classical method, events are rejected if all three/four tau candidates with highest τ^{jetBDT} value do not fulfil a certain τ^{jetBDT} threshold which runs from -1 to $+1$. In the new method, the τ^{LLH} product of the first three/four tau candidates is demanded to exceed a running threshold between 0 and 1.
Signal events are all events of the specific process in which indeed the considered tau candidates are truth matched taus. Correspondingly, all events are labelled as background if one or more of them are fake. In figure (e), all processes are combined when the first three tau candidates are studied and in figure (f), the combined results are shown if the first four taus are considered. Analogous figures for the τ^{LLH} based cuts for the remaining samples as well as process-related studies with four taus are found in A.7. The results when utilising the probability τ^{prob} are presented in figure A.8.

The strange shape of the ROC curve for $Z \to \tau\tau$ process (d) can be explained by several negatively weighted events with three or more truth matched taus due to higher-order corrections of the MC samples. This effect is not physical but indicates insufficient statistics of the corresponding process and will be further studied in section 6.4.

5.3.4 Significance Gain of the Multiplied Probability Method

In the $hh \to 4\tau$ channel alone (cf. figures 5.11(a), A.7(d)), the multiplied-probability method presented in the previous section might not improve the separation of three or four truth matched taus from fake-enriched events of the same process. However, its main interest is to separate the di-Higgs signal from background processes.

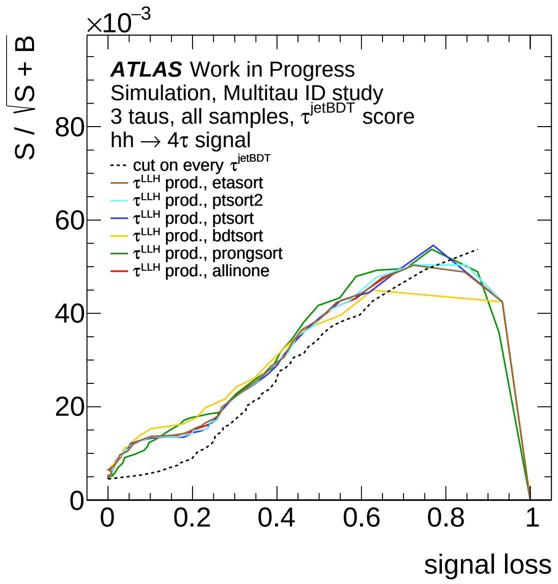
The majority of events to deal with actually originates from background processes like $t\bar{t}$ production or $Z \to \tau\tau$ events. Both processes have in common that they hardly produce four truth taus. Accordingly, most of the background is highly fake-enriched. Additionally, the multiplied probability method is not only capable to reject background fake-processes (cf. e.g. figure 5.11(c)) but the $\tau^{\text{LLH/prob}}$ product also rejects background processes with four truth taus better than single BDT cuts. This is probably due to the fact that e.g. in $t\bar{t}$ processes, truth taus are mainly created in decays of heavy quarks. These taus are not meant to be identified as taus in the ATLAS detector and thus, their $\tau^{\text{jetBDT}}_{\text{(trans)}}$ value is probably not very high.[10] Hence, even a four-truth-tau process in $t\bar{t}$ probably does not achieve such a high $\prod \tau_i^{\text{LLH/prob}}$ as the di-Higgs signal with four isolated taus.

In this analysis, the most interesting result is the possible significance gain to separate di-Higgs signal from background processes. This can be quantified by calculating the significance $S/\sqrt{S+B}$ with S being the amount of signal events surviving a selection and B is the corresponding amount of background processes. Moreover, several methods to map the $\tau^{\text{jetBDT}}_{\text{(trans)}}$ scores to the $\tau^{\text{LLH/prob}}$ score were introduced in section 5.3.2. They are also evaluated with respect to their significance gain. For this purpose, the $\prod \tau_i^{\text{LLH/prob}}$ value is demanded to exceed running values from 0 to 1 and the amount of remaining signal and background events is noted. The resulting plots are presented in figure 5.12 for $\prod \tau_i^{\text{LLH}}$ and figure 5.13 for $\prod \tau_i^{\text{prob}}$. Moreover, the corresponding plots were cross checked with the $ZZ \to 4\tau$ process and are presented in figure A.9 and A.10 in the appendix.

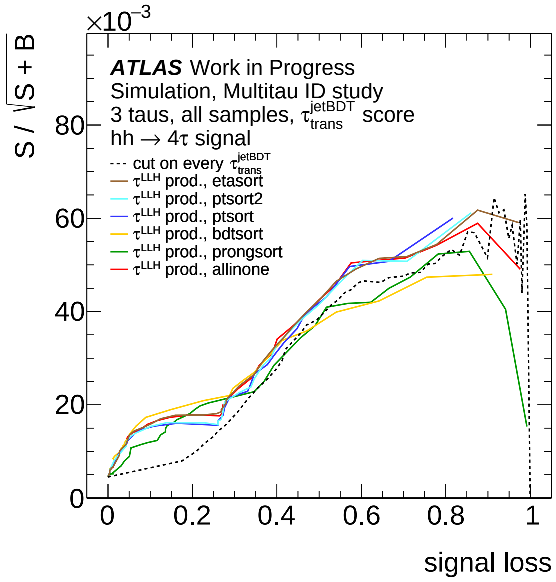
In all plots, a clear enhancement of the signal significance can be observed when cutting on the $\prod \tau_i^{\text{LLH/prob}}$ value instead of using the $\tau^{\text{jetBDT}}_{\text{(trans)}}$ value of each candidate on its own (depicted in the black dashed line). Moreover, the likelihood seems to be very stable with respect to properties of the tau candidates. As a drawback, no significant improvement of the significances can be observed by utilising more of the properties named before.

When using the probability product, however, the `prongsort` method clearly exceeds all other sorting methods. Similar observations are made when using the $ZZ \to 4\tau$ process as the signal (cf. figures A.9 and A.10). The original $\tau^{\text{jetBDT}}_{\text{(trans)}}$ scores already distinguish between 1-prong and 3-prong taus and two different BDT scores were designed [95]. Thus, it might be natural to assume an information gain by creating two different τ^{LLH} scores for 1-prong and 3-prong taus as well. In contrast to that, the BDT scores were designed to be flat with respect to most other tau candidate properties. Hence, the similarities between `etasort`, `ptsort`, `ptsort2` and `allinone` document this flatness and unfortunately no further significance gain can be achieved by these methods.

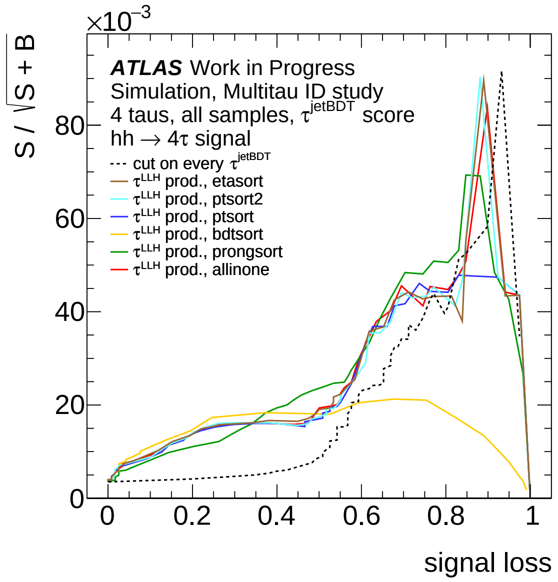
[10] Probably, most of the time it is a medium value such that the event is not rejected by requirements on each tau candidate's $\tau^{\text{jetBDT}}_{\text{(trans)}}$ score.



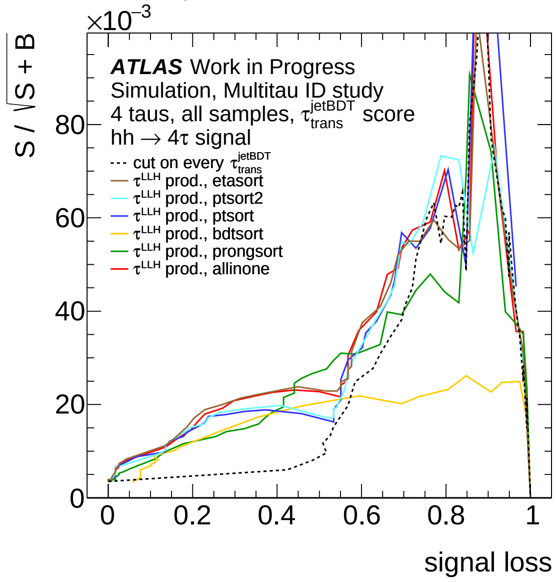
(a) $\prod \tau^{\text{LLH}}_{i=1-3}$ created with the τ^{jetBDT} score



(b) $\prod \tau^{\text{LLH}}_{i=1-3}$ created with the $\tau^{\text{jetBDT}}_{\text{trans}}$ score

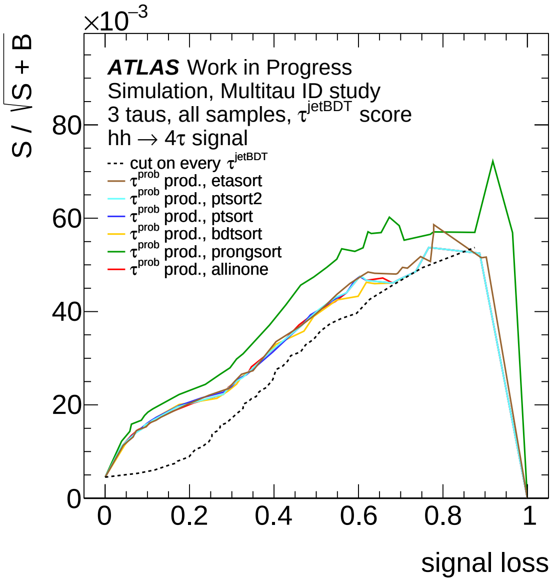


(c) $\prod \tau^{\text{LLH}}_{i=1-4}$ created with the τ^{jetBDT} score

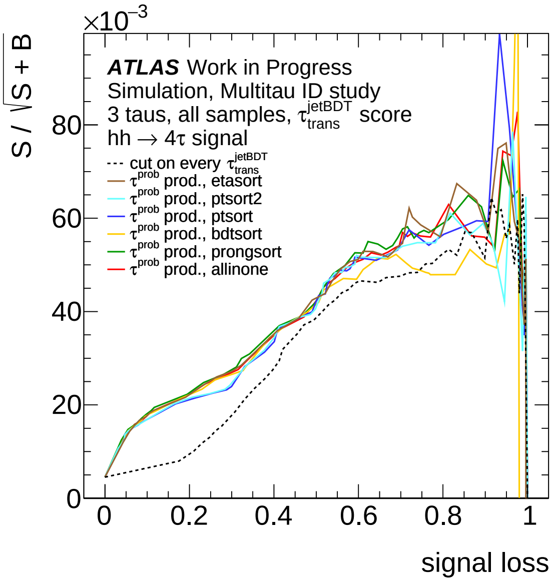


(d) $\prod \tau^{\text{LLH}}_{i=1-4}$ created with the $\tau^{\text{jetBDT}}_{\text{trans}}$ score

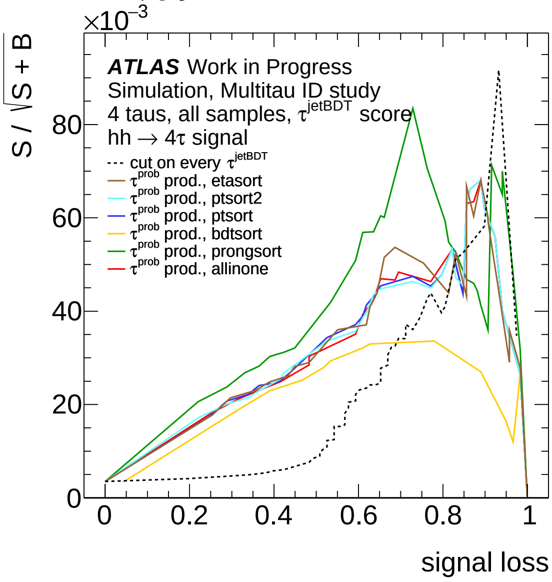
Figure 5.12: The significances while cutting either on the τ^{LLH} product of the first three tau candidates (a) and (b), or on four of them (c), (d) with respect to the signal loss are presented. Here, every product is first evaluated when the τ^{LLH} values were gained with the normal τ^{jetBDT} score and second when the transformed score $\tau^{\text{jetBDT}}_{\text{trans}}$ was used. Different sorting methods were further used to map the BDT scores to the likelihood (cf. section 5.3.2). Compared to cutting on $\tau^{\text{jetBDT}}_{\text{(trans)}}$ (like in the classical tau identification approach, visualised by the black dashed line) a clear significance gain in most sorting methods is observed. However, no sorting method really singles out to be the best one.



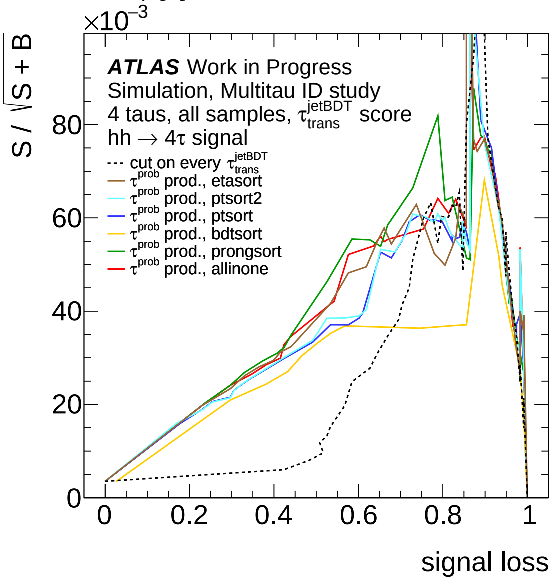
(a) $\prod \tau^{\text{prob}}_{i=1-3}$ created with the τ^{jetBDT} score



(b) $\prod \tau^{\text{prob}}_{i=1-3}$ created with the $\tau^{\text{jetBDT}}_{\text{trans}}$ score



(c) $\prod \tau^{\text{prob}}_{i=1-4}$ created with the τ^{jetBDT} score



(d) $\prod \tau^{\text{prob}}_{i=1-4}$ created with the $\tau^{\text{jetBDT}}_{\text{trans}}$ score

Figure 5.13: The significances while cutting either on the τ^{prob} product of the first three tau candidates (a) and (b), or on four of them (c), (d) with respect to the signal loss are presented. Here, every product is first evaluated when the τ^{prob} values were gained with the normal τ^{jetBDT} score and second when the transformed score $\tau^{\text{jetBDT}}_{\text{trans}}$ was used. Different sorting methods were further used to map the BDT scores to the probabilities (cf. section 5.3.2). Compared to cutting on $\tau^{\text{jetBDT}}_{\text{(trans)}}$ (like in the classical tau identification approach, visualised by the black dashed line) a massive significance gain in nearly all sorting methods is observed. In all cases, the `prongsort` sorting method in green either clearly or slightly outperforms the other sorting methods.

In a next step, the separation power of τ^{LLH} and τ^{prob} is compared directly. Moreover, the usage of τ^{jetBDT} or $\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$ is evaluated.

Additionally, it is tested if the consideration of all samples to map the BDT scores to $\tau^{\mathrm{LLH/prob}}$ leads to the best result or if it is more sensitive to exclusively study the fake and truth distributions of signal-like samples. In this side-study, only samples are used to create the $\tau^{\mathrm{LLH/prob}}$ scores, which are at least close to create many truth taus.[11] In summary, eight different modes are tested against each other. They are summarised in table 5.3 and their performance in the `prongsort` method is visualised in figure A.11 for di-Higgs and di-Z events in the appendix. As the plots are very busy, they are also depicted without the signal-like modes in figure 5.14 for di-Higgs processes.

Mode name	Cut based observable	used BDT score	used processes
τ^{LLH} prod., All	$\prod \tau_i^{\mathrm{LLH}}, i = 1-3/4$	τ^{jetBDT}	all
τ^{LLH} prod., Sig.	$\prod \tau_i^{\mathrm{LLH}}, i = 1-3/4$	τ^{jetBDT}	signal-like
τ^{LLH} prod., All (trans)	$\prod \tau_i^{\mathrm{LLH}}, i = 1-3/4$	$\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$	all
τ^{LLH} prod., Sig. (trans)	$\prod \tau_i^{\mathrm{LLH}}, i = 1-3/4$	$\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$	signal-like
τ^{prob} prod., All	$\prod \tau_i^{\mathrm{prob}}, i = 1-3/4$	τ^{jetBDT}	all
τ^{prob} prod., Sig.	$\prod \tau_i^{\mathrm{prob}}, i = 1-3/4$	τ^{jetBDT}	signal-like
τ^{prob} prod., All (trans)	$\prod \tau_i^{\mathrm{prob}}, i = 1-3/4$	$\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$	all
τ^{prob} prod., Sig. (trans)	$\prod \tau_i^{\mathrm{prob}}, i = 1-3/4$	$\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$	signal-like

Table 5.3: Different modes which were used to map the $\tau^{\mathrm{jetBDT}}_{\mathrm{(trans)}}$ scores to $\tau^{\mathrm{LLH/prob}}$ values

In all figures, it can be observed that the different τ^{prob} product cuts lead to much higher significances than cutting on the τ^{LLH} product. This is mostly also true when the $\tau^{\mathrm{LLH/prob}}$ value is created while using only signal-like samples. However, on average, considering all samples (i.e. also all background processes) leads to slightly better significances such that only their results are presented in figure 5.14.

As the τ^{prob} value utilises the additional information of the ratios of truth and fake taus which are considered, this might have been expected. This knowledge is of increased importance when studying the `prongsort` method. As 3-prong taus are much more likely to be faked (cf. figure 5.9(c)) they gain smaller probabilities to be a truth tau. If an event contains many 3-prong taus, is it thus more likely to have its origin in a background process and this is quantised by the τ^{prob} product. In opposition to that, this information is lost due to normalisation in the τ^{LLH} calculation.

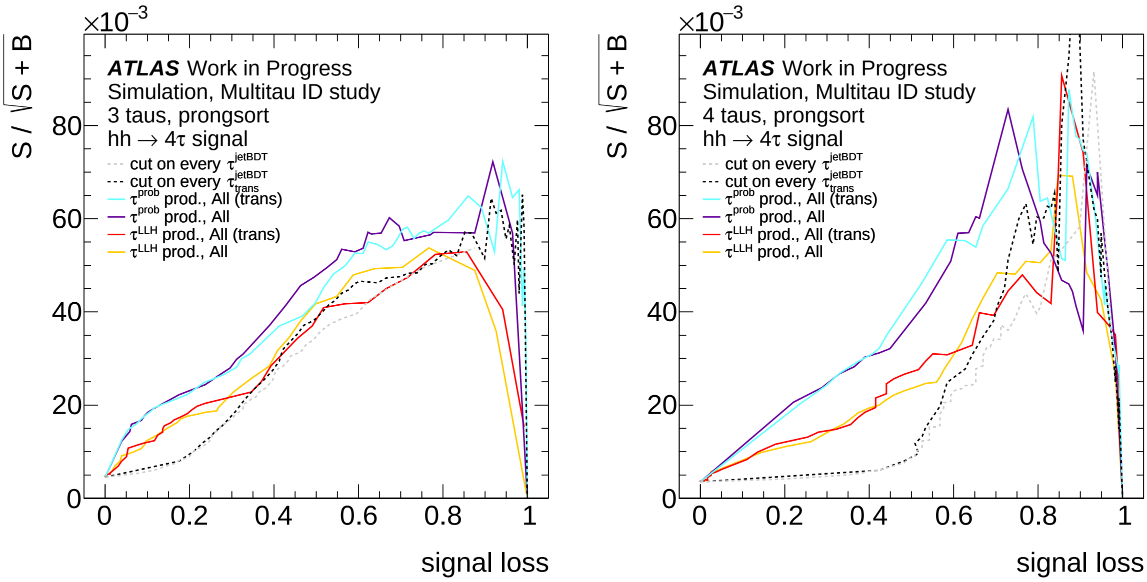
As a result, in the following chapter, the τ^{prob} product will be used to separate signal processes from background. While for the $3h1\ell$ channel, the τ^{prob} value which originated from the τ^{jetBDT} value seems to be best, in the $4h0\ell$ channel, the corresponding value from the $\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$ score is used. However, their differences might simply be caused by statistical fluctuations in the used samples and the regions in which either of them performs best alternate.

In summary, the presented new identification method (based on multiplied probability cuts) outperforms the classical identification method (based on separate cuts on each tau candidates $\tau^{\mathrm{jetBDT}}_{\mathrm{(trans)}}$ score) by far.

For the same signal efficiency, much higher significances are achieved when cutting on $\prod \tau_i^{\mathrm{prob}}$. For example, for a signal loss of 61 %[12] the classical approach achieves signal significances of roughly

[11] This refers to the observation, that the τ^{prob} value strongly depends on the selection of considered samples. For the signal-like samples, apart of the $hh \to 4\tau$ process, the `Sherpa` $\ell\ell\ell\ell$ sample (which e.g. contains $ZZ \to 4\tau$) as well as $\ell\ell\ell\nu$, $\ell\ell\nu\nu$ and $\ell\nu\nu\nu$ were also added to the overall truth and fake distributions.

[12] As expected when demanding four *Loose* taus.



(a) $hh \to 4\tau$ signal, 3 tau candidates

(b) $hh \to 4\tau$ signal, 4 tau candidates

Figure 5.14: The separation power of τ^{LLH} and τ^{prob} is compared. Both values are either produced with the normal τ^{jetBDT} score or the transformed $\tau^{\text{jetBDT}}_{\text{trans}}$ score. It can be observed that the τ^{prob} values in purple or turquoise by far outperform the τ^{LLH} values in red or gold.
These plots were also created when only signal-like processes were used to create the truth and fake distributions for the $\tau^{\text{LLH/prob}}$ scores. In figure A.11 in the appendix, it can be seen that in most cases the significance is reduced if the values are created only with the information of signal-like samples. Thus, they are discarded in the presented figures here for the sake of clarity.

$30 \cdot 10^{-3}$ (cf. figure 5.14(b), the dashed lines). When using the multi-tau approach, this significance is doubled.

Moreover, the new approach enables similar significances of $30 \cdot 10^{-3}$ for a signal loss of only 35 %. Hence, it is possible to achieve the same significances while keeping much more signal events. For the three-tau channel, (cf. figure 5.14(a)), the effect is verified although the full power of this method evolutes best for more objects.

In summary, the new multi-tau identification method represented in this chapter is ideally suited to search for the extremely rare signal events of the $hh \to 4\tau$ channel. It will further be exploited in the following chapter, in which a signal enriched region is designed.

CHAPTER 6

Approaching the Di-Higgs Signal

In section 4.1, it is proven that apart from the $hh \to 4\tau$ process, only a very limited amount of processes in the SM produces four truth τ leptons. Nevertheless in a vast amount of processes such a multi-tau topology can be faked e.g. by tau-like jets. The fake-rate increases drastically if the tau identification threshold is reduced as much is it is done in this thesis (cf. table 4.4).

Hence, background processes with the capability to produce a considerable amount of truth or fake taus have to be estimated and controlled. For this purpose, in section 6.1 heavy-particle processes which are likely to produce four tau-like objects are presented. As they can partly be treated perturbatively, MC simulations are used to estimate their properties.

In contrast to that, pure QCD events suffer from model uncertainties for the essential process of hadronisation (cf. section 4.2). Hence, MC predictions usually do not provide an adequate estimation of the fake-rates in these processes. Instead, in section 6.2, a data-driven method is designed to estimate their production of four fake objects.

In section 6.3, the pairing of the four leptons in each event is studied. To utilise the unique event topology of the $hh \to 4\tau$ process, the leptons have to be divided such that the pairs emerging from the same Higgs boson are identified. An essential task is the alignment of this combination to the data-driven fake estimation.

Eventually, different methods are probed to separate the di-Higgs signal from named background processes. For this purpose, multiple observables are scanned with respect to a possible significance gain (cf. section 6.4). Based on that, phase space regions with maximised signal signatures are designed. In section 6.5, the signal-region selection is improved by specially designed boosted decision trees and in section 6.6 the capability of the most promising phase space regions is interpreted. Moreover, they are used to calculate an upper limit on the di-Higgs cross section.

Several observables are used in the following studies. The used notation is outlined below.

- $\boldsymbol{p_{\mathrm{T}}(\tau_i)}$ The transverse momentum of the *i*th tau. Here, the tau candidates are sorted with respect to their τ^{jetBDT} score. For instance, τ_1 is the tau candidate with the highest score of the event.

- $\boldsymbol{m_{\tau_1\tau_2}/\, m_{\tau_3\tau_4}}$ $m_{\tau_1\tau_2}$ is the invariant mass of the ditau system which contains the positively charged tau candidate with highest τ^{jetBDT} score (further referred to as the *first* ditau system). The tau candidates are paired as described in section 6.3. Analogously, $m_{\tau_3\tau_4}$ is the invariant mass of the ditau system which contains the positively charged tau candidate with lower τ^{jetBDT} score. As both ditau systems are treated analogously, possible differences of these values are not utilised in this work.

- $\boldsymbol{m_{\tau\tau}^{\mathrm{low}}}$/ $\boldsymbol{m_{\tau\tau}^{\mathrm{high}}}$ The lower/higher of both possible invariant masses of the ditau systems. These quantities are beneficial for significance scans and the BDT training as a cut on them directly implies that both ditau system masses are bounded from below/above.

- $\boldsymbol{m_{\tau\tau,\,\mathrm{false}}^{\mathrm{low}}}$ The lower of both possible invariant masses of the ditau systems when they are assigned in opposition to the introduced combination approach of section 6.3. In di-Higgs processes, both Higgs bosons are produced back to back. By combining one tau of each Higgs decay, relatively high invariant masses are achieved. If the tau candidates are produced randomly, this is not necessarily the case. This quantity is used in the BDT training.

- $\boldsymbol{\Delta R_{\tau_1\tau_2}}$/ $\boldsymbol{\Delta R_{\tau_3\tau_4}}$ The distance of both tau candidates of the first ($\Delta R_{\tau_1\tau_2}$) or second ($\Delta R_{\tau_3\tau_4}$) ditau system, respectively.

- $\boldsymbol{\Delta R_{\tau\tau}^{\mathrm{low}}}$/ $\boldsymbol{\Delta R_{\tau\tau}^{\mathrm{high}}}$ The lower/higher tau-distance of both ditau systems. Again, these quantities are useful in significance scans and the BDT training as directly a threshold on both ditaus is implied.

- $\boldsymbol{\sum^{\mathrm{low}} p_{\mathrm{T}}(\tau\tau)}$ The lower of both possible scalar sums $p_{\mathrm{T}}(\tau_i) + p_{\mathrm{T}}(\tau_j)$ of both ditau systems. The scalar sum of a specific ditau system is written out explicitly.

- $\boldsymbol{\sum p_{\mathrm{T}}(\mathrm{lep}_{1\text{-}4})}$ The scalar sum $|p_{\mathrm{T}}(\tau_1)| + |p_{\mathrm{T}}(\tau_2)| + |p_{\mathrm{T}}(\tau_3/\ell_1)| + |p_{\mathrm{T}}(\tau_4/\ell_{1/2})|$ of the four leptons (light leptons and/or tau candidates) which form the two ditau systems. In the $4h0\ell$ channel, also the notation $\sum p_{\mathrm{T}}(\tau_{1\text{-}4})$ is used for this quantity.

- $\boldsymbol{p_{\mathrm{T}}(\tau_1\tau_2\tau_3\tau_4)}$ The vector sum of the four leptons (light leptons and tau candidates) which form the two ditau systems. In the $hh \to 4\tau$ case, this quantity describes the boost of the overall di-Higgs system. This observable should not be confused with $\sum p_{\mathrm{T}}(\tau_{1\text{-}4})$.

- $\boldsymbol{m_{\tau_1\tau_2\tau_3\tau_4}}$ The invariant mass of the four leptons (light leptons and tau candidates) which form the two ditau systems. In the $hh \to 4\tau$ case, this quantity describes the invariant mass the overall di-Higgs system.

- $\boldsymbol{\Delta R_{\tau_1\tau_2\tau_3\tau_4}}$ The distance between both ditau systems. In the $hh \to 4\tau$ case, this quantity describes the distance between both Higgs bosons.

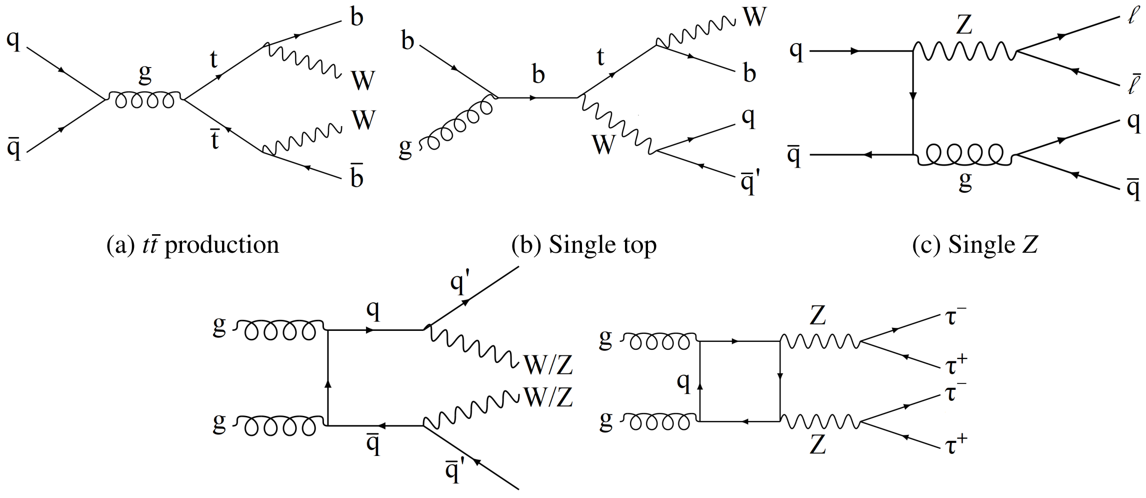
- $\boldsymbol{N_{\mathrm{objects}}}$ The sum of all tau candidates, light leptons and jets. As fake systems tend to produce more jets and/or more lepton candidates, this value tends to be lower for the di-Higgs process.

- $\boldsymbol{N_{\mathrm{jets,\,40\,GeV}}}$ The amount of jets with a transverse momentum of $p_{\mathrm{T}} > 40\,\mathrm{GeV}$

6.1 Tau Fake Production by Heavy Particles

In this thesis, the investigation of four-tau events is investigated for the first time. Hence, the background estimation was built up from scratch. Due to the very low signal cross section, the identification threshold for tau candidates was loosened as much as possible (cf. section 4.3). As a drawback, tau candidates are very likely to be faked by hadronic jets. Hence, several SM processes may be capable to produce a multi-tau-like structure which is caused by fakes. A selection of such processes is visualised in figure 6.1.

As an example, in $t\bar{t}$ production (cf. figure 6.1(a)), both t quarks almost exclusively decay into Wb. The b quark decays further and creates a hadronic jet which might be misidentified as a tau. The W boson might decay directly into a tau or into a light lepton (and then occur in the $3h1\ell$ or $2h2\ell$ channel). Apart from that, a hadronic W decay can also be misinterpreted as a hadronically decaying tau. Analogously,



(a) $t\bar{t}$ production (b) Single top (c) Single Z

(d) Diboson (e) $ZZ \to 4\tau$

Figure 6.1: Possible processes which might fake a four-tau event topology. For most of them, several production and decay modes exist. All visualised processes contain at least four final states. If quarks are produced, they create a hadronic jet which might be misidentified as a τ_{had}. W bosons might either decay hadronically and fake a τ lepton, or they directly decay into a tau. Z bosons might even decay into a $\tau^+\tau^-$ lepton pair which looks nearly like a tau pair of a Higgs boson. Accordingly, the SM process which is expected to mimic the di-Higgs signal best is the $ZZ \to 4\tau$ process (e). Further explanations are given in the text.

four tau candidates possibly emerge from an associated Wt process (cf. figure 6.1(b)) which will further be referred to as “single top”. Fortunately, the probability to misidentify four jets as taus at the same time is expected to be relatively low. Nevertheless, due to the comparably huge cross section of these t quark processes at the LHC, they form one of the dominant backgrounds of this work.

In single Z boson events (cf. figure 6.1(c)) two truth τ leptons are produced directly, such that the signal topology is partly mimicked. For instance, a dominant distribution of the background consists of $Z \to \tau\tau$ events where the remaining two tau candidates are faked by jets. Also Z boson decays into light leptons are considered as they might serve as background processes in the $3h1\ell$ or $2h2\ell$ channel, especially if one of the light leptons is lost or misidentified as a 1-prong tau. The latter ones are included in the single vector boson background distribution and referred to as “V + Jets”. However, they are found to play a comparably insignificant role.

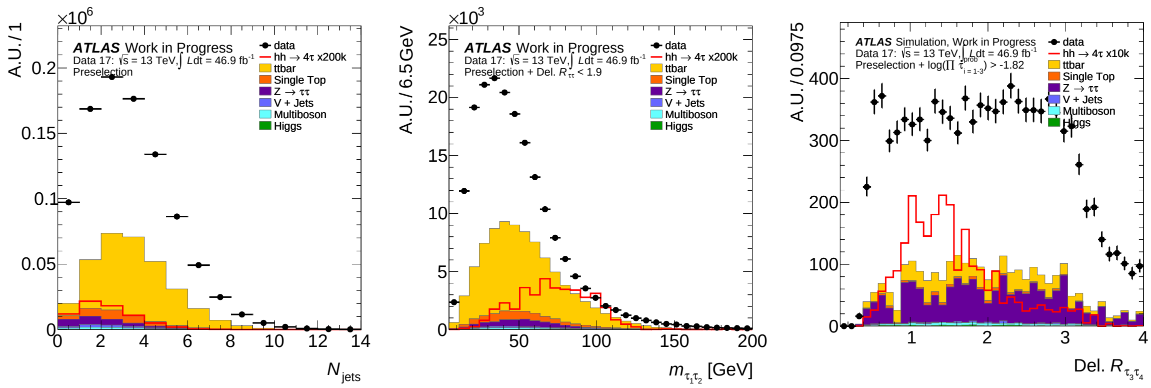
For events in which more than one W or especially Z boson is produced, the number of emerging light leptons, taus or collimated jets increases. Thus, events with two or more vector bosons are studied and referred to as “multiboson” (cf. figure 6.1(d)). They also include the $ZZ \to 4\tau$ process (cf. figure 6.1(e)) which is thought to resemble the di-Higgs process most closely.

Last but not least, events which already contain one Higgs boson are likely to look similar to a di-Higgs process. Especially vector-boson-fusion or Higgs radiation (cf. figure 2.7(b), 2.7(c)) either produces two additional jets which potentially fake the residual taus or an electroweak boson which might decay further to truth light leptons or taus. Although Higgs processes feature a relatively low cross section, they are still several orders of magnitude more frequent than di-Higgs processes. In some of their decay channels, they come close to the signal topology.

A list of used samples as well as details of their MC production can be found in section A.1 in the appendix.

The designed background composition has already been used in the truth study in section 4.1 and the effect of the preselection cuts has been tested in section 4.3. Indeed, single Higgs and multiboson events are most similar to the signal process. Accordingly, they form the exclusive background on truth level (cf. table 4.1). However, due to much higher cross sections in the other processes, the overall event distribution which passes the preselection is dominated by far by t events (cf. table 4.4). Hence, observables to isolate signal events from t quark fakes are an essential part of this work. Nevertheless, due to their similar event topology, multiboson and single Higgs events are expected to become more dominant when a sharp signal region is defined.

To verify that all important processes are selected which might account for the background, they are tested against the data for several phase space regions (cf. figure 6.2). A massive discrepancy is observed. The missing events probably do not have their origin in processes as described in this section but are pure QCD events. They are very hard to model with MC simulations. Therefore, a data-driven estimation of their contribution is designed in the following section.



(a) Preselection (b) $\Delta R_{\tau\tau} < 1.9$ (c) $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) > -1.82$

Figure 6.2: A comparison of events predicted by the considered MC processes and recorded data is presented for different observables and cuts. In (a) the number of jets is shown after the preselection introduced in section 4.3 is applied. In (b) the invariant mass of the first ditau pair is presented after $\Delta R < 1.9$ is demanded and in (c), the distance $\Delta R_{\tau\tau}$ between two taus which are assigned to the same ditau system is shown for all events which fulfil the additional requirement $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) > -1.82$. In all plots, the amount of recorded events significantly exceeds the predicted amount by the considered MC simulations.

6.2 Data Driven Fake Estimation with Four Objects

The overwhelming amount of processes at the LHC are QCD processes. Often multiple jets are produced such that the multi-tau topology can be mimicked. Unfortunately, due to the strong coupling in the hadronisation process of the jets, they cannot be described by perturbation theory. As this is the essential process which determines weather a jet is identified as a tau or not, their modelling is error-prone.

Hence, QCD or multi-jet processes are usually estimated by data-driven methods. Here, the data in orthogonal regions of the phase space is used to estimate the amount of missing events in the signal region. To ensure that no physical information is lost with this method or that a data-to-model agreement is artificially created, the orthogonal regions to extract data information have to be chosen carefully. Ideally, the data-driven fakes solely contain pure QCD events and all other processes are estimated with well controlled MC simulations.

In this thesis, the data driven fakes are estimated by demanding unphysical charge combinations.

In physical processes like Higgs decays into $\tau^+\tau^-$, the tau charges have to equalize. Hence, for four correlated objects, two positively and two negatively charged objects are expected. If the tau candidates, however, mainly consist of randomly misidentified objects, their charges are expected to have no correlation at all and a binomial distribution of 1 : 4 : 6 : 4 : 1 is expected.[1] In table 6.1, the ratios of charge combinations for some selected processes as well as for the recorded data are presented.

Cut	$hh \to 4\tau$	$ZZ \to 4\tau$	Higgs	$Z \to \tau\tau$	Diboson	V + Jets	$t\bar{t}$	Data
4+0−	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
3+1−	23.5	20.7	82.1	16.6	9.9	8.8	8.5	4.4
2+2−	164.5	84.5	175.4	30.3	17.6	15.9	15.3	6.6
1+3−	24.0	21.2	77.5	15.2	8.9	7.4	8.4	4.1
0+4−	1.0	1.0	0.9	0.8	0.8	0.7	1.0	0.9

Table 6.1: The ratios of charge combinations for some selected processes are presented. The corresponding charge requirement of the preselection was abdicated for this study. Each process is normalised with respect to their amount of events in the 4+0− channel. If the charges of all four objects are distributed completely random, a binomial distribution of 1 : 4 : 6 : 4 : 1 is expected. A small asymmetry towards positive charges can be explained by the two-times positively charged initial state of two protons in each collision. Further interpretations are given in the text.

The charge distribution of all MC processes manifests to be far from binomial. Especially di-Higgs or di-Z events strongly peak at 2+2− as expected due to the underlying physical process. In the data, however, a shape close to the binomial distribution is revealed. As physical processes hardly produce four taus in the 3+1− or even 4+0− channel, this is a strong hint that the data indeed mostly contains QCD multi-jet processes.

By assuming that multi-jet events share the same properties independently of their charge distribution, the unphysical charge combinations can be utilised to predict their appearance in the 2+2− channel. Correspondingly, the amount of QCD processes in the unphysical channels is estimated by subtracting all MC predictions from the recorded data. The remaining number of events is rescaled to the 2+2− channel subsequently. Unfortunately, a simple upscaling of all four unphysical channels in the 1 : 4 : 6 : 4 : 1 scheme is not sufficient to dispose of the observed data-MC disagreement as visualised in figure 6.3.

Instead an additional information is utilised, which is already foreshadowed in table 6.1. Namely, in the last four processes ($Z \to \tau\tau$, Diboson, V + Jets and $t\bar{t}$ production) the distributions of the 3+1−, 2+2− and 1+3− channels can roughly be described by 1 : 2 : 1, i.e. the binomial distribution of two objects. Exactly this distribution is expected when only the charges of two objects are created randomly while the other two are produced in correlation to each other. For instance, in the $Z \to \tau\tau$ channel, two truth taus with opposite charges emerge as decay products of the Z. To pass the preselection, two other tau candidates (or the sufficient amount of light leptons) have to be faked. As the fake-tau candidates do not necessarily originate from the same physical process, their charges are not correlated and thus form a 1 : 2 : 1 distribution. Analogously, in $t\bar{t}$ production processes, the decay products certainly contain one W^+ and one W^- boson which decay further. In short, again two charged objects are produced with correlated charges and the other two emerge randomly.

[1] Of course, four positively charged jets in a single event cannot emerge without the creation of negatively charged objects as well. However, if pure QCD processes are expected to fake tau leptons, several jets are produced. The jets which are reconstructed only with one or three tracks such that they are tagged as tau candidates may have a random charge distribution. The equalising negative charges can be absorbed by other jets which are not identified as tau candidates. Moreover, charges can be misidentified in the reconstruction process, especially when trajectories of high-p_{T} objects feature a small curvature.

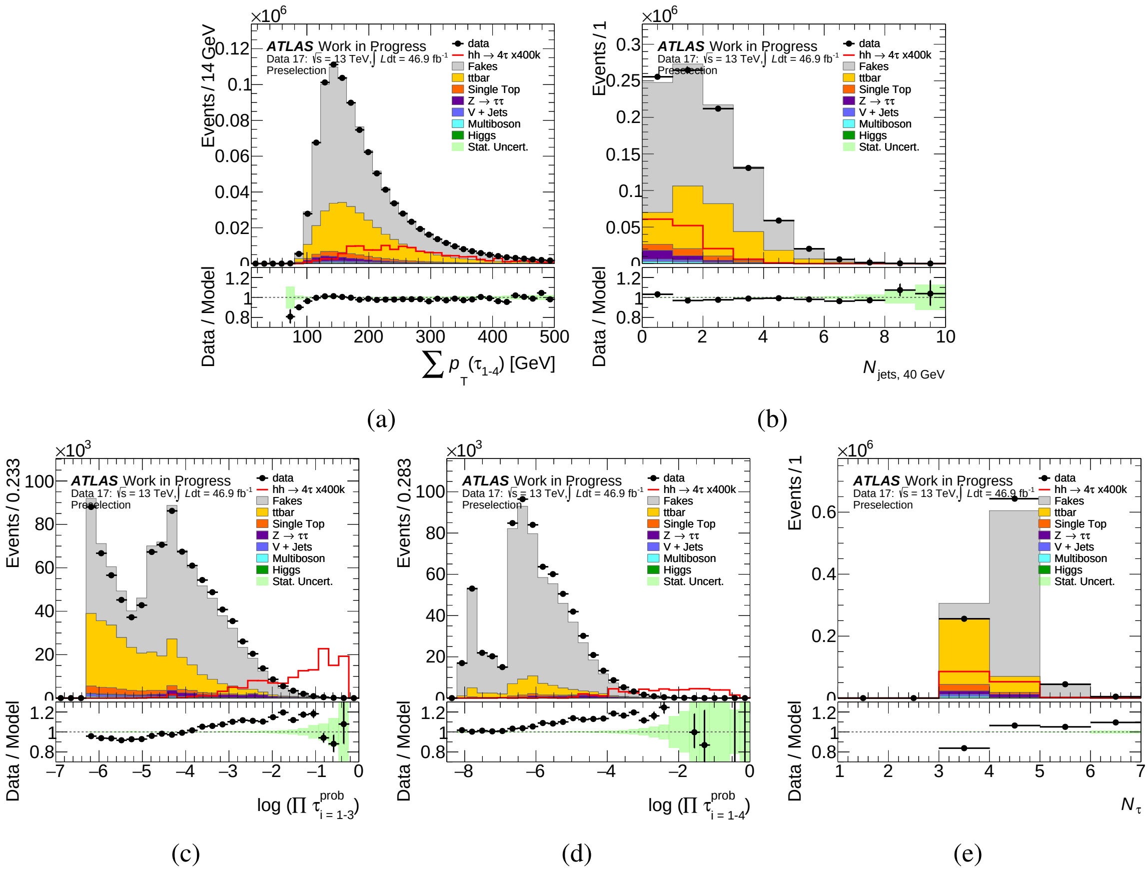
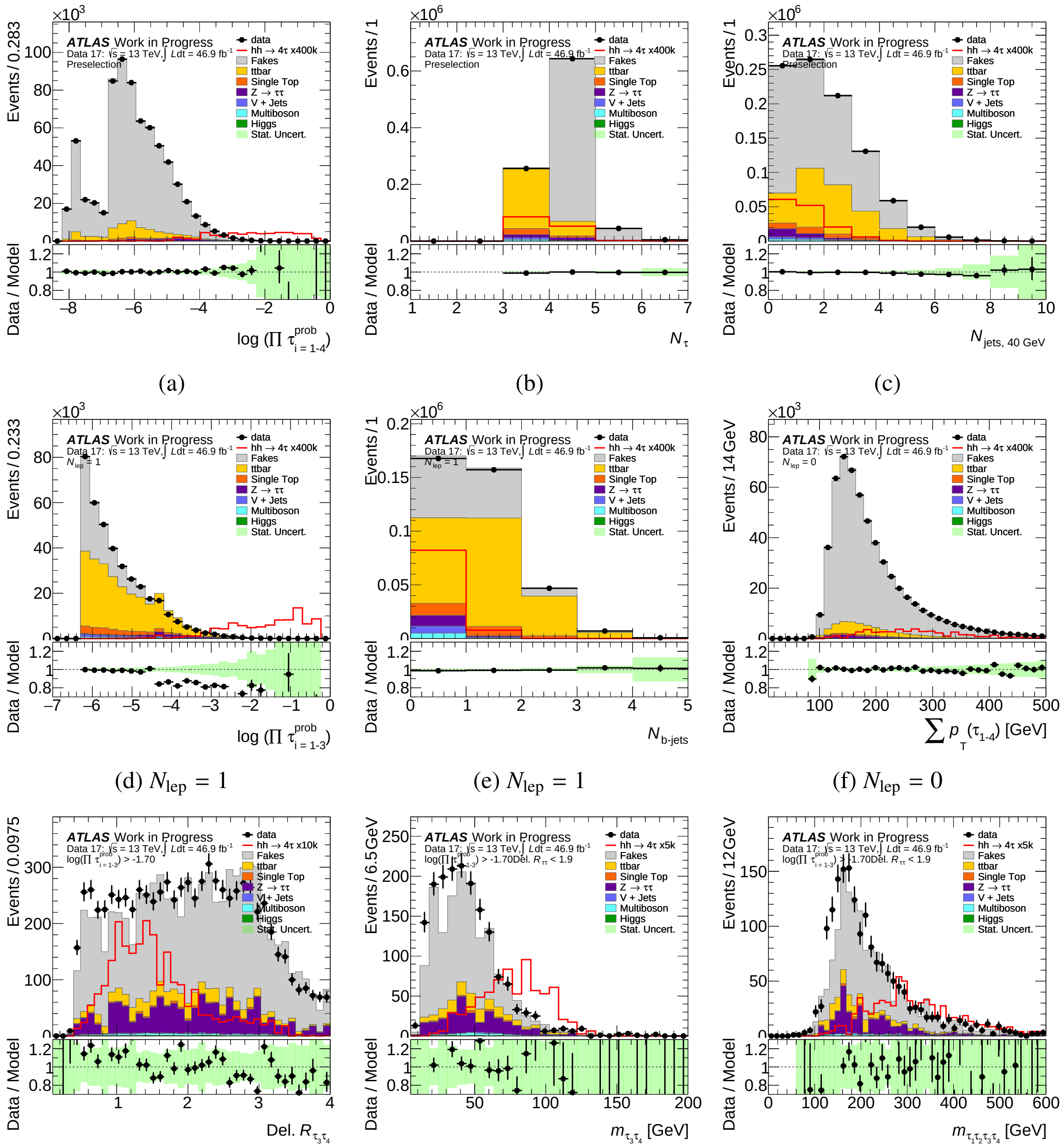


Figure 6.3: The MC predictions together with the data-driven fake estimation are compared to data. In the presented figures, the fakes are predicted by naively upscaling the unphysical 4+0− and 3+1− regions (and the corresponding ones with more negatively charged leptons) by a factor of 6 or 1.5 respectively. If all four objects are produced completely random, a charge distribution of 1 : 4 : 6 : 4 : 1 is expected and the named factors of 6 or 1.5 correspond to the ratio of the 4+0− or 3+1− region and the physical 2+2− region respectively.
While the estimated fakes seem to suit the gap between MC predictions and data quite well for $\sum p_{\mathrm{T}}(\tau_{1\text{-}4})$ (a), or $N_{\text{jets, 40 GeV}}$ (b), this is not the case for the latter three observables. Especially for $\log\left(\prod \tau_i^{\text{prob}}\right)$ ((c) and (d)), a systematic aberration is observable. The predicted amount of produced tau candidates (e) is also not in agreement with the observed data. The latter quantity is of special interest as the different channels which are investigated in this thesis are highly correlated with the amount of produced tau candidates.

Such a fake-production is accounted for as follows. To properly estimate the produced amount of four fake taus, the excess of measured events in the 4+0− and 0+4− channels compared to the MC predictions are averaged and sextupled to serve as a first fake estimation in the 2+2− channel. Additionally, the named distributions are quadruplicated and subtracted from the average of the data-to-MC excess in the 3+1− and 1+3− channel. The remaining events in the latter channels are then upscaled in the 1 : 2 : 1 scheme and complete the fake estimation in the physical channel. An advantage of this method is that possible additional processes which create two correlated taus and two fakes are also accounted for and are added to the overall fake estimation.

The resulting fake estimation in the 2+2− channel is visualised in figure 6.4. Apart of some tail regions, the overall event estimation agrees very well with the data due to the introduced fake-tau estimation. This is also true after applying several cuts. Thus, the data-driven method designed in this section is well



(a) (b) (c)

(d) $N_{\text{lep}} = 1$ (e) $N_{\text{lep}} = 1$ (f) $N_{\text{lep}} = 0$

(g) $\log\left(\prod \tau^{\text{prob}}_{i=1-3}\right) > -1.70$ (h) $\log\left(\prod \tau^{\text{prob}}_{i=1-3}\right) > -1.70$, $\Delta R_{\tau\tau} < 1.9$ (i) $\log\left(\prod \tau^{\text{prob}}_{i=1-3}\right) > -1.70$, $\Delta R_{\tau\tau} < 1.9$

Figure 6.4: The MC predictions in combination with the data-driven fake estimation are compared to data. Here, the more sophisticated scaling method introduced in the text is used to predict the amount of fakes in the 2+2− channel. Apart of some exceptions in the tail regions a strong agreement can be observed. This agreement still holds after several cuts which are summarised in each caption. The cuts are meant to be additional to the preselection.

suited to predict the properties of fake-taus due to pure QCD events.

The robustness of the fake estimation can further be tested by separating fake-like and tau-like production channels from each other in a different way. For this purpose, the tau identification properties are inverted. As the tau identification in this thesis is directly correlated to the $\log\left(\prod \tau_i^{\text{prob}}\right)$ values, a strongly tau-fake enhanced region can be defined by demanding the events to undershoot a specific $\log\left(\prod \tau_i^{\text{prob}}\right)$ threshold. Again the amount of fakes is estimated by the excess of data compared to MC predictions in this phase-space region. As these selected events are assumed to consist almost exclusively of fake-taus, they are expected to share the same properties as the fakes which are estimated by the charge-related method presented in the previous paragraph. The properties of these tau candidates defined by different tau-fake definitions are compared in figure 6.5.

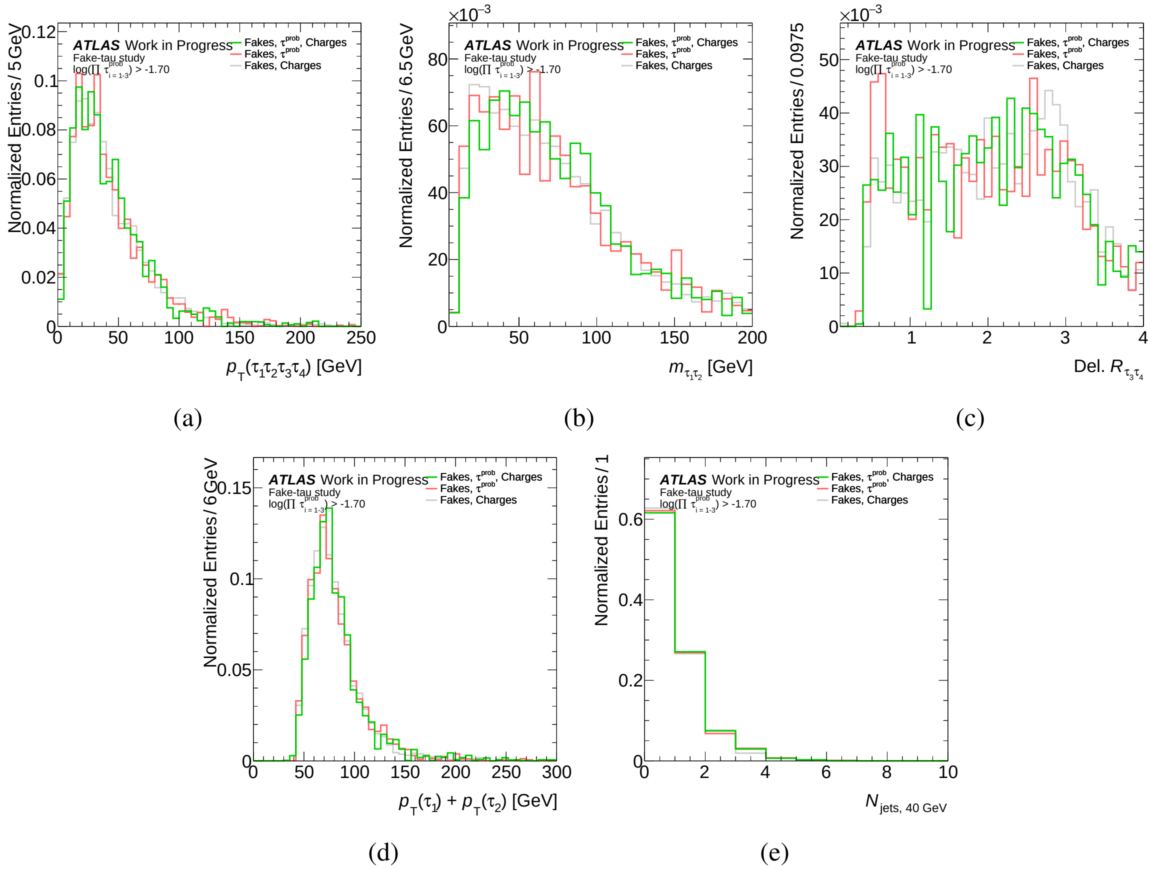


Figure 6.5: The normalised shape of different data-driven fake estimations is compared. The fakes are either defined by the scaling of different channels with unphysical charges (“Fakes, Charges”) or by inverting the requirement of $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) > -1.70$. In detail, these “Fakes, τ^{prob} ” are demanded to fulfil $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) \in (-1.82, -1.74)$ such that their total amount of events is roughly equivalent to the amount which surpasses the $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right)$ threshold. Additionally, unphysical charges are investigated in this inverted $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right)$ region (“Fakes, τ^{prob}, Charges”). All methods are thought to describe highly fake-enriched events of similar origin. This implies that apart of obvious differences like their charges, the differently selected events should resemble each other in most other physical properties. This is confirmed by the figures shown here.

The differently defined fake-taus are confirmed to share such a similar behaviour for the presented observables. This strengthens the reliability of the fake estimations and supports the assumption that they indeed exclusively describe fakes. Moreover, in a next step which exceeds the scope of this thesis, the differences in the presented specific fake-tau definitions can be utilised to estimate the fluctuation strength of the fake-taus. Systematic uncertainties of these events can be estimated and ideally reduced significantly.

A few more remarks can be made. First, hardly any fakes pass the preselection for $N_\tau = 3$ (cf. figure 6.4(b)). This can be explained by the fact that at least one additional light lepton is required if only three tau candidates are produced. In this thesis, at least one of the light leptons is required to fulfil the *Medium* working point. Therefore, it is relatively unlikely that light leptons are faked. Apparently, the existence of such a clean light lepton is extremely rare in pure QCD processes which are thought to dominate the data-driven fakes marked in grey.[2]

Second, in figure 6.4(a) and 6.4(d), the data-MC agreement becomes worse for high values of $\log\left(\prod \tau_i^{\text{prob}}\right)$. Unfortunately, this is exactly the phase space region which will be of highest interest when separating the di-Higgs signal from background (cf. section 6.4 and 6.5). Especially in the $3h1\ell$ channel, the MC predictions exceed the measured data (cf. figure 6.4(d)). In fact, this explains the low appearance of fakes in the $N_\tau = 3$ region. In the data-driven fake estimation, only the excess of data to MC is used. If the MC predictions already overshoot the measured data, naturally no fakes will be predicted. Thus, the observed disagreement does not have its origin in a suboptimal fake estimation but in the MC predictions in general.

This behaviour in particular and the general behaviour of each MC sample should be studied further. Usually, the different production channels are separated in respective control regions which ideally are doubtlessly understood. In these control regions, each process is fitted to the measured data and possible minor mismodelling is diminished by global correction factors for each process.

Moreover, it should be mentioned, that the presented charge-based data-driven fake estimation only describes fakes with exactly four objects which share the same vertex. As in the following, the amount of signal events is tried to keep as high as possible, such a requirement might be suboptimal. A more complete fake-estimation for events with more than four objects can be designed by the introduced inverted $\log\left(\prod \tau_i^{\text{prob}}\right)$ method. Moreover, regions with 4+1− or 5+0− might be included with a sufficient scaling. Maybe this issue can also be solved by an overall global scale factor on the charge-based fakes or a combination of all proposals in a so-called ABCD method can be applied. Possible scaling factors can be calculated by fit programs like `HistFitter` [111] .

For this thesis, the combination of MC predictions (cf. section 6.1) and the presented data-driven fake estimations of this section are thought to model the background sufficiently well such that further investigations are possible. In the following sections, their properties are studied in detail and observables are designed to separate background processes from the signal.

6.3 Ditau Pair Combination

After a sufficient understanding of the dominant background processes has been ensured in the previous sections, the di-Higgs process can be approached. To utilise the unique signal topology, the four leptons of each event have to be combined such that both leptons emerging from the same Higgs boson are paired correctly. Hereby, a specific difficulty is to ensure that the data-driven fakes, which feature different

[2] Data-driven fakes are still observed after demanding $N_{\text{lep}} = 1$ (cf. e.g. figure 6.4(d)). Further investigation revealed that these fakes almost exclusively feature more than three tau candidates. Additionally, the forth tau candidate emerges from another vertex then the other four leptons. Therefore, these events can be reduced significantly by an appropriate cut selection.

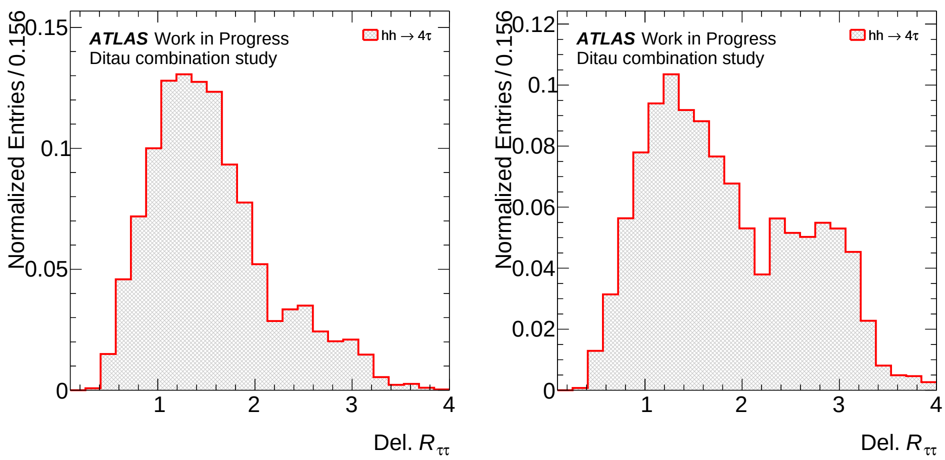
charges, are treated comparably. Otherwise, their ditau properties are not comparable to the events in the 2+2− region.

The property in which a Higgs boson differs most from e.g. a Z boson is its mass. Thus, a very important quantity to distinguish the $h \to \tau\tau$ decay from the corresponding one of a Z is the invariant mass of the tau pair (cf. e.g. [7, 8, 38]). A meaningful calculation of the invariant mass requires the right combination of taus as input objects. While this assignment is straightforward for two objects, the tau combination for four objects is ambiguous.

Each event which passes the required preselection (cf. section 4.3) contains at least four leptons from which at least two are positively and negatively charged respectively. As the identification requirements for light leptons are much higher in this thesis, they are always used to form the ditau pairs, if existing. Nevertheless, in this section they are not always mentioned explicitly for the sake of simplicity. If the amount of leptons of a specific charge exceeds two, only the tau candidates with highest τ^{jetBDT} scores are considered.

In each Higgs decay, the decay products are required to have opposite charges. Therefore, for two leptons of each charge, two ditau pair combinations are possible ($\tau_1^+\tau_1^-$ & $\tau_2^+\tau_2^-$ or $\tau_1^+\tau_2^-$ & $\tau_2^+\tau_1^-$). As already mentioned in the introduction of this section, a selection of tau pairs depending on their invariant masses might suggest itself. Indeed, such a tau pair combination is used in the $hh \to b\bar{b}b\bar{b}$ analysis [52]. Here, the tau pair combination is chosen which is most consistent with the decay of two particles of equal mass.

Unfortunately, in the four-tau case, a significant amount of energy is lost due to undetected neutrinos. As a result, a similar-mass requirement on the ditau pairs is not ideal. Instead, the topology of the Higgs decays in the $hh \to 4\tau$ process is utilised. As shown in figure 6.6(a), the distance $\Delta R_{\tau\tau}$ of both taus emerging from the same Higgs tends to be very low on truth-level. In contrast to that, two distinct peaks in $\Delta R_{\tau\tau}$ are observed on reconstruction level in figure 6.6(b) if the taus are combined with respect to their resulting invariant masses.



(a) Truth ditau ΔR (b) Reco. ditau ΔR, m^{inv} mapping

Figure 6.6: The distance between two leptons which are assigned to form a ditau pair is shown for the $hh \to 4\tau$ process. On truth level (a), a clear peak at low values is observable. When the ditau pairs are chosen with respect to their invariant masses, however, a second peak at roughly $\Delta R = 2.9$ appears (b). In this peak, probably one tau of each Higgs was paired. As the Higgs bosons tend to be boosted back to back, a much higher distance between these taus is measured.

As a result, the distance of possible tau pairs is chosen to select the tau combination in this thesis. In detail, the taus are assigned such that the resulting $\sum \Delta R_{\tau\tau}$ with

$$\sum \Delta R_{\tau\tau} = \Delta R_{\tau_1^+ \tau_i^-} + \Delta R_{\tau_2^+ \tau_j^-} \tag{6.1}$$

is minimised. The gained benefit of this method is visualised in figure 6.7. Only 13 % of the ditau pairs is mismatched by the $\sum \Delta R_{\tau\tau}$ method while for the similar-mass method, this amount rises to 21 %. In both cases, in 18.65 % of all events the ditau pairs cannot be matched correctly because at least one of the decay products is not reconstructed sufficiently.

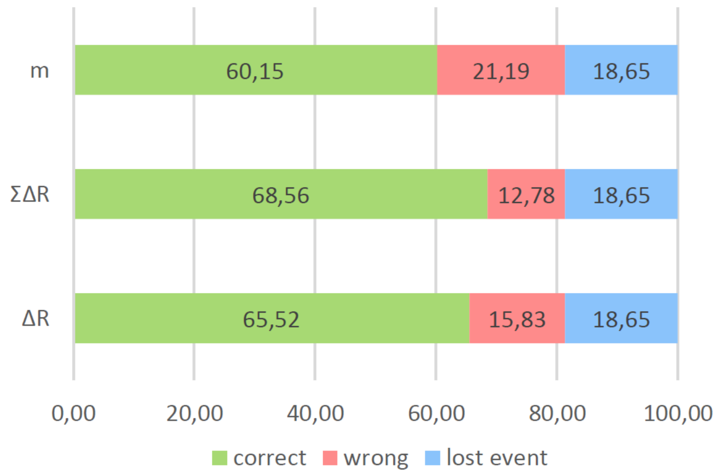


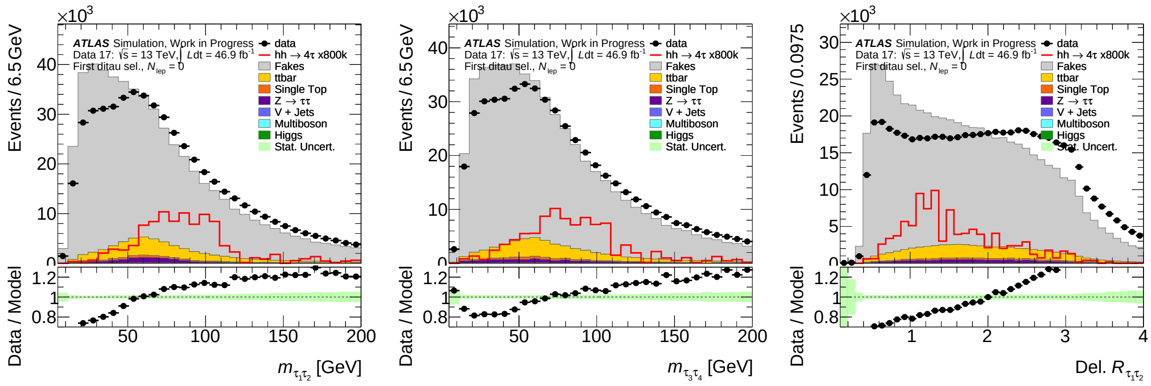
Figure 6.7: The ratio of correctly combined ditau pairs for different selection methods is visualised. The upper bar refers to a ditau combination in which the invariant masses are most similar (m). The requirement to minimise the sum of distances between each tau pair is referred to as $\sum \Delta R_{\tau\tau}$. In $\Delta R_{\tau\tau}$, only the first ditau pair's tau distance is required to be minimal with respect to the alternative mapping.

The pairing of taus becomes considerably more complicated when unphysical charges are allowed. This might not affect the signal region directly, but events with charges like 4+0− or 3+1− are utilised to form the fake-tau distribution as introduced in section 6.2. In these events, one or even both ditau pairs have to be formed with leptons of the same charge, which enables not only two but three possible ditau assignments.[3] The most straightforward solution (referred to as *first selection method*) might be to select the combination out of the three in which the sum of ditau distances is minimal. Unfortunately, this leads to poor fake estimations as visualised in figure 6.8(a), 6.8(b) and 6.8(c).

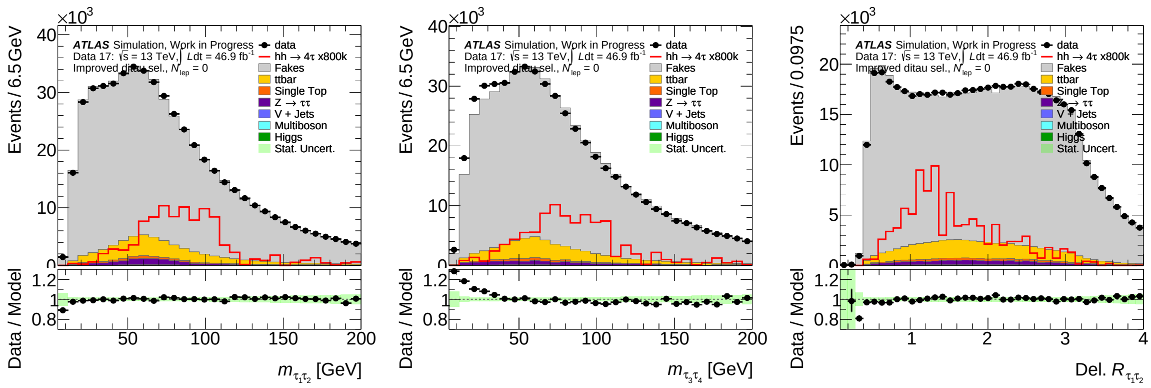
This can be explained by the fact that such a comparison between three possible ditau pairs enables more degrees of freedom than the comparison between only two ditau pairs in the physical 2+2− region. Thus, in a considerable amount of cases, the third option enables a combination in which an even lower sum of $\Delta R_{\tau\tau}$ is found. As a result, in unphysical channels the average $\Delta R_{\tau\tau}$ of the combined taus is smaller (cf. figure 6.8(c)) and correspondingly, their average invariant masses are lower.

As a solution, in the *improved selection method* also in the unphysical region, only two possible ditau combinations are checked against each other with respect to their $\sum \Delta R_{\tau\tau}$. The third option, in which the tau with highest τ^{jetBDT} score (of the majoritarian charge) is mapped to the one with lowest charge

[3] In the 3+1− channel, they read $\tau_1^+\tau_1^-$ & $\tau_2^+\tau_3^+$, $\tau_2^+\tau_1^-$ & $\tau_1^+\tau_3^+$ and $\tau_3^+\tau_1^-$ & $\tau_1^+\tau_2^+$. For the 4+0− case, $\tau_1^+\tau_2^+$ & $\tau_3^+\tau_4^+$, $\tau_1^+\tau_3^+$ & $\tau_2^+\tau_4^+$ and $\tau_1^+\tau_4^+$ & $\tau_2^+\tau_3^+$ are possible. For channels with more negatively charged leptons, the combinations are equivalent but with switched charges.



(a) First selection method (b) First selection method (c) First selection method

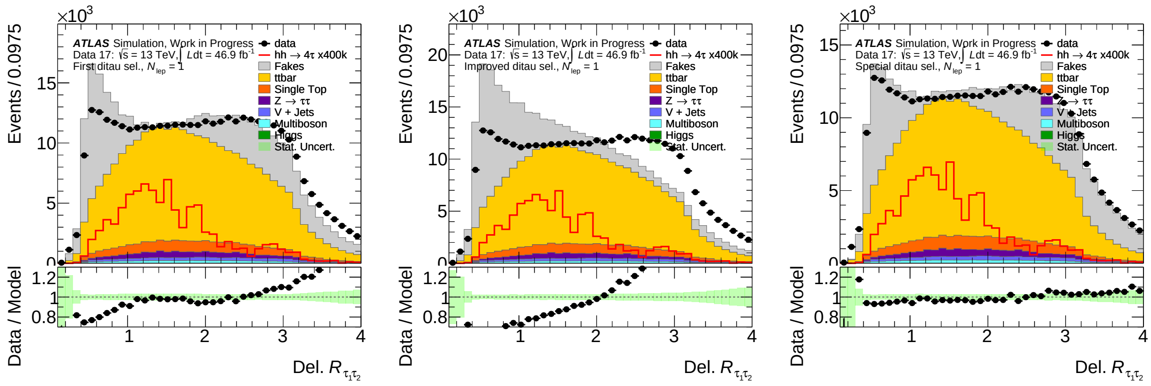


(d) Improved selection method (e) Improved selection method (f) Improved selection method

Figure 6.8: The data-MC agreement of ditau properties for different ditau selection methods is compared. In each figure, light leptons are excluded. The invariant mass $m_{\tau\tau}$ tends to be underestimated in (a) and (b). This has its origin in the fact that in the data-driven fake estimation, an additional degree of freedom is gained by selecting the lowest sum of $\Delta R_{\tau\tau}$ in three and not two possible ditau combinations. If one of the possibilities is rejected, better agreement is achieved as visualised in figure (d) to (f). Further explanations are given in the text.

of the same kind, is discarded. The resulting new fake estimations are presented in figure 6.8(d) to 6.8(f). A good data-Monte-Carlo agreement is achieved with this method. Hence, it is assumed that the unphysical channels undergo a comparable treatment compared to the 2+2− channel and thus, also the ditau combination of the fakes in the latter channel is modelled sufficiently well.

If one of the considered leptons is a light lepton, however, the ditau properties become even worse (cf. figure 6.9(a) and figure 6.9(b)). As the tau candidates are ordered with respect to their τ^{jetBDT} score, the appearance of a light lepton changes the order in which possible $\Delta R_{\tau\tau}$ values are compared. A more comparable handling of unphysically charged fakes with respect to the straight-forward 2+2− region is achieved when only those combinations are considered in which the light lepton is paired with tau candidates of relatively low τ^{jetBDT} score. The resulting $\Delta R_{\tau\tau}$ with this *light lepton selection* is presented in figure 6.9(c). Hence, the ditau combination is slightly altered for events in which light leptons are included.



(a) First selection method (b) Improved selection method (c) Light lepton selection

Figure 6.9: The data-MC agreement of ditau properties for different ditau selection methods is compared. In each figure, exactly one light lepton is demanded. It can be seen that neither by the comparison of the distances of all three possible ditau systems (a) nor by the comparison of only the first two options (b) the unphysically charged regions are handled equivalently to the 2+2− region. As a result, the data in the latter region is not sufficiently described by fakes estimated by the unphysically charged channels. The lepton combination in unphysical channels is aligned to the one of the 2+2− channel, if the light lepton is not combined with the tau candidate with highest possible τ^{jetBDT} score (c). These effects are mainly technical and correspond to the order in which leptons are treated in the N-tuple production.

6.4 Signal Region Definition Based on Significance Scans

The amount of events which passes the preselection requirements is extremely dominated by background (cf. table 4.4). Hence, in order to detect the di-Higgs signal process, the background has to be reduced massively. At the same time, however, hardly any signal should be lost as its amount of totally produced events is very limited.

To solve both tasks at the same time, many observables are scanned in multiple dimensions of the parameter space. The significance $S/\sqrt{S+B}$ is used as a guideline to evaluate possible cuts. Here, S denotes the amount of remaining signal events and B is the corresponding background amount. Moreover, the events are strictly divided into the $3h1\ell$ and $4h0\ell$ channel for this study by demanding either exactly one or no light lepton at all. The $2h2\ell$ channel will also be investigated although the preliminary work of this thesis was focussed on the former two channels. For this channel, exactly two leptons are demanded. To reduce the enormous amount of background events like $Z \to \ell^+\ell^-$, the light leptons are required to be either of the same charge and/or of different generations. The preselection introduced in section 4.3 is altered appropriately for this channel.

In every channel, the main limitation of the significance is the very low prediction of signal events. As already indicated in section 5.1, the general acceptance of signal events might be increased by relaxing the requirements for tau candidates further. To estimate the possible significance gain, the tau-candidate requirements were loosened to a minimal transverse momentum of $p_T > 15\,\text{GeV}$ instead of the previous $p_T > 20\,\text{GeV}$. Additionally, also candidates in the $|\eta| \in [1.37, 1.52]$ were allowed. The requirements of either one or three charged tracks and an overall charge of one were not altered. This new tau candidate selection is referred to as **very loose taus** from now on.[4]

[4] This should not be misidentified with the unofficial *Very Loose* working point which is used by some ATLAS analyses and refers to the tau identification. Here, similar to other working points (cf. section 5.2) a minimal value of each $\tau^{\text{jetBDT}}_{\text{trans}}$ score

Indeed, a very small significance gain of the order of 1 % was observed in some possible signal regions in every channel. However, the introduction of such loose tau candidates comes with several drawbacks as explained in section 5.1. In particular, especially for tau candidates in the constrained $|\eta|$ region, their energy is not recommended to be reconstructed by the calorimeters (as their information in this $|\eta|$ region is incomplete) but only by their track information. In short, they would need to undergo an specially designed new reconstruction process to ensure a maximal performance. Moreover, the uncertainty of very low transverse momenta is strongly limited by pile-up processes. As the high-luminosity LHC upgrade features much more pile-up events, this uncertainty might grow. Thus, low-p_T tau candidates need to be treated very carefully.

It is questionable if such a low significance gain is worth the effort. All of the three signal regions defined in the following subsections require at least $\sum p_T(\tau_{1\text{-}4}) > 200\,\text{GeV}$. It is straight-forward to assume that most low-p_T taus are rejected by this demand anyway. The exploitation of the $|\eta|$-crack region, however, remains interesting. As the reconstruction of affected taus was not altered in this thesis, their erroneous energy information probably lead to the fact that they were rejected by momentum or energy requirements before they could enhance the signal expectation in the signal region of the following channels.

Hence, the the following subsections refer to the previous tau candidate requirements as introduced in table 4.3 with $p_T > 20\,\text{GeV}$ and a veto for the $|\eta|$ crack region. In section 6.5, however, boosted decision trees are trained to separate signal processes. With this approach, adverse correlations (like the appearance of a tau candidate in the $|\eta|$ crack and a correlated strange invariant mass of the ditau system) might be balanced more adequately and the very loose tau selections remains interesting.

6.4.1 $4h0\ell$ Channel

A significance scan over several properties reveals the strong separation power of the $\log\left(\prod \tau^{\text{prob}}_{i=1-4}\right)$ as defined in chapter 5. This is visualised in figure 6.10(a). As the unique signal topology is defined by four truth taus and the named product was developed to separate best between truth and fake events, this is expected. Moreover, the scalar sum of all transverse momenta of the tau candidates $\sum p_T(\tau_{1\text{-}4})$ stands out against other quantities due to its possible high significance gain (cf. figure 6.10(b)). Since Higgs bosons are more massive than most other SM particles, it is natural to assume that their exclusive decay into taus produces leptons with higher momenta than most of the background processes. Especially, if most of the tau candidates are fakes, they tend to have lower momenta. By calculating the scalar sum of all four tau candidates, the separation power is increased.

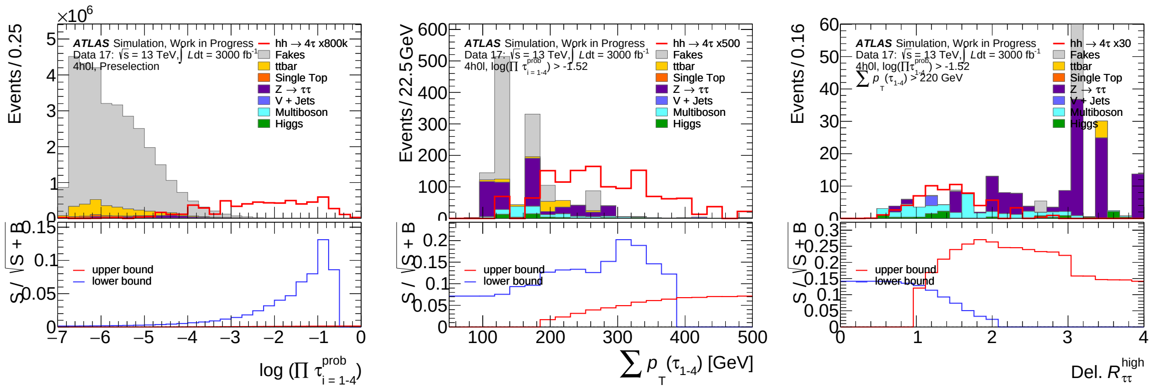
As already discussed in section 6.3, the distances of two taus forming a ditau system are quite low for the $hh \to 4\tau$ signal. This is reflected in the increased significance by demanding an upper cut on $\Delta R_{\tau\tau}$ as visualised in figure 6.10(c). Furthermore, the significance scans reveal that the invariant mass of each ditau system becomes interesting to separate signal from background especially after additional requirements on $\Delta R_{\tau\tau}$ (cf. figure 6.4(h) in section 6.2).

Such correlations between variables occur frequently. Hence, a hard cut on one observable might spoil a possible higher significance by more careful demands on several quantities. To account for this, at first a comparably soft requirement of

$$\log\left(\prod \tau^{\text{prob}}_{i=1-4}\right) > -2 \qquad \text{and} \qquad \sum p_T(\tau_{1\text{-}4}) > 220\,\text{GeV} \tag{6.2}$$

is chosen. With this selection, a two-dimensional significance scan of the invariant mass of each ditau

is demanded. The very loose tau candidates of this work refer to the tau reconstruction process and are much looser than demanded by the *Very Loose* working point.



(a) (b) (c)

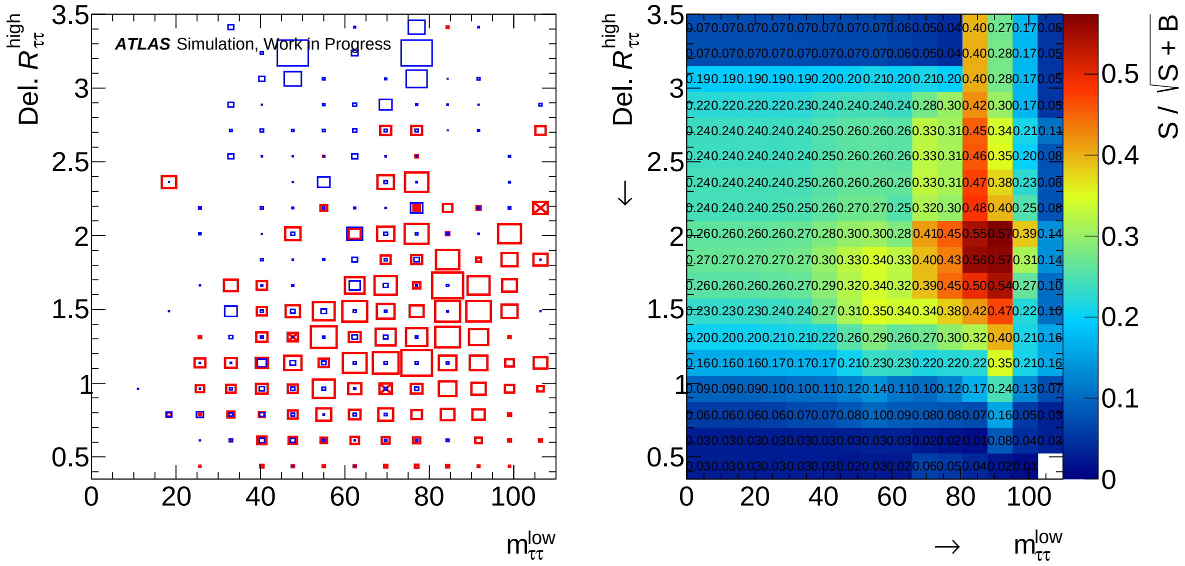
Figure 6.10: The significance scans of the most interesting observables in the $4h0\ell$ channel are shown. In the upper half of each plot, a usual stack plot of all background processes and the artificially enhanced signal process are shown. The lower half visualises the significance alteration depending on changing cuts.
The blue line represents the significance $S/\sqrt{S+B}$ if a cut from below of the specific plot quantity is applied. For instance, by additionally demanding $\sum p_T(\tau_{1\text{-}4}) > 310\,\text{GeV}$ in figure (b), a significance of roughly 0.2 is achieved. The red line visualises the achieved significance, if a cut from above is applied and thus has to be read from right to left. For example in figure (c) an additional requirement of $\Delta R < 1.9$ leads to an overall signal significance of all remaining events of roughly 0.27. A clear rise of either the blue line (from left to right) or the red line (from right to left) indicates a beneficial cut to separate signal from background.
Hence, in (a) a minimal requirement on $\log\left(\prod \tau^{\text{prob}}_{i=1-4}\right)$ is recommendable. Correspondingly, in (b) $\log\left(\prod \tau^{\text{prob}}_{i=1-4}\right) >$ 1.52 is demanded. To further rise the significance, a lower bound of the scalar sum of all tau momenta like $\sum p_T(\tau_{1-4}) > 220\,\text{GeV}$ (or even more) is advisable. In (c), both cuts are applied and it is visualised by the red significance line that an upper bound of $\Delta R_{\tau\tau}$ will increase the significance further.

system and the distance of its taus is performed as visualised in figure 6.11. This scan is repeated for the stricter demand of $\log\left(\prod \tau^{\text{prob}}_{i=1-4}\right) > -1.52$ and/or $\sum p_T(\tau_{1\text{-}4}) > 300\,\text{GeV}$. Moreover, the requirements of $N_{b\text{-jets}} = 0$ or $N_\tau = 4$ (i.e. vetoing b-jets or demanding exactly four taus without allowing more (fake) tau candidates) were found to be interesting in previous studies. Thus, all mentioned two-dimensional significance studies are also repeated with and without these demands.

Cuts	$hh \to 4\tau$	Higgs	Multiboson	Others	Combined	$S/\sqrt{S+B}$
1	0.76	0.47	0.40	0	0.87	0.60
2	0.65	0.36	0.11	0	0.47	0.61
3	1.10	1.26	3.02	0	4.28	0.47

Table 6.2: The predicted events of different processes in possible signal regions of the $4h0\ell$ channel are displayed. The processes V + Jets, $Z \to \tau\tau$, Single Top, $t\bar{t}$ production and fakes are summarised in “others” and are estimated to be completely discarded by these selections. Depending on the amount of remaining signal events, different significances from 0.47 to 0.61 are achieved.
The required cuts in each selection are $\log\left(\prod \tau^{\text{prob}}_{i=1-4}\right) > -1.52$ and $\sum p_T(\tau_{1\text{-}4}) > 220\,\text{GeV}$. Moreover, in cut 1, $\Delta R_{\tau\tau} < 1.925$ and $m_{\tau\tau} > 81\,\text{GeV}$ is required. In cut 2, additionally events with more than exactly four tau leptons are rejected. In cut 3, again more than four reconstructed tau candidates are allowed and the ditau requirements are slightly altered to $\Delta R_{\tau\tau} < 2.08$ and $m_{\tau\tau} > 73\,\text{GeV}$. The slightly different numbers compared to figure 6.11(b) are due to a different binning in this figure which also defines the cuts there.

The highest significance with a signal event expectation close to at least one is achieved for the



(a) (b)

Figure 6.11: The correlation of the lower invariant mass of both ditau systems $m_{\tau\tau}^{\text{low}}$ against the higher of both tau distances $\Delta R_{\tau\tau}^{\text{high}}$ is shown in figure (a). It can be seen that most of the normalised signal events (marked in red) accumulate in the lower right corner, i.e. at high masses and low distances. The background in blue is distributed broader. In (b) a two-dimensional significance scan of these observables is shown. Here, each bin represents a *cut from below* on $m_{\tau\tau}^{\text{low}}$ of the corresponding value and a *cut from above* on $\Delta R_{\tau\tau}^{\text{high}}$. Hence, only events which are located in bins to the right or below each corresponding bin are used to calculate the significance as indicated by the arrows in each axis title. The best significances are achieved by demanding roughly $m_{\tau\tau}^{\text{low}} > 80\,\text{GeV}$ and $\Delta R_{\tau\tau}^{\text{high}} < 2.0$.

following selection:

$$\boxed{\log\left(\prod \tau_{i=1-4}^{\text{prob}}\right) > -1.52, \quad \sum p_{\text{T}}(\tau_{1\text{-}4}) > 220\,\text{GeV}, \quad \Delta R_{\tau\tau} < 1.925 \quad \& \quad m_{\tau\tau} > 81\,\text{GeV}} \quad .$$

Here, the former two cuts were selected since the highest significance values were observed in the two-dimensional significance scan with this selection as shown in figure 6.11(b). The latter two cuts are directly motivated by this figure. As presented in table 6.2, a maximal significance of roughly 0.6 is achieved.

However, the comparably high significance is gained by rejecting almost all background while not even one signal event is expected. It is questionable if such a background rejection is natural or just due to limited statistics of the simulation of corresponding processes. The latter constraint is supported by the fact that especially in the $Z \to \tau\tau$ sample, strong fluctuations are observable (cf. figure 6.10(b) and 6.10(c)). Here, a few events with very high event weights (due to the upscaling to $3\,000\,\text{fb}^{-1}$) pollute phase space regions close to the signal region. The high amount of negatively weighted events in this sample also lead to a strange behaviour of the ROC curve in figure 5.11(d).

Although it was ensured that no highly negatively weighted events of one process pull down the overall background prediction, they can push up the data-driven fake estimation. Here, MC predictions are usually subtracted from the data in unphysical regions. If the overall MC prediction is highly negative, the opposite happens and a highly positive fake estimation results, as observed in the named figures.

These are only effects due to limited statistics and are not physical at all.

Since the $Z \to \tau\tau$ process is one of the main background processes, reliable predictions can only be made if statistical limitations in the simulation are controlled or ruled out. Moreover, no systematic uncertainties were taken into account yet. Their consideration will lower the final significance further.

Nevertheless, the observed background processes in the signal region agree with the ones expected in the truth-level study (cf. table 4.1). The lower event expectation for single Higgs and multiboson processes compared to the truth-level study can be explained with their comparably lower tau momenta and thus lower tau acceptances as outlined in chapter 5. Moreover, several other event selections lead to similar significances which supports the assumption that this is more than statistical fluctuations in a very specific phase space region. Additionally, in some bins, even higher significances were observed (cf. table 6.2). Thus, an assumption of **a possible significance in the $4h0\ell$ channel of roughly 0.56 to 0.61** seems to be reliable as a first guess.

6.4.2 $3h1\ell$ Channel

For the $3h1\ell$ channel, an analogous procedure as in the $4h0\ell$ channel is conducted. Similarly, the $\log\left(\prod \tau_i^{\text{prob}}\right)$ value and the sum of all lepton momenta $\sum p_{\text{T}}(\text{lep}_{1\text{-}4})$ feature a good separation power and a considerable significance gain can be achieved (figure 6.12(a) and 6.12(b)). Therefore, a first requirement of

$$\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) > -1.30 \qquad \text{and} \qquad \sum p_{\text{T}}(\text{lep}_{1\text{-}4}) > 210\,\text{GeV} \tag{6.3}$$

is chosen. To gain additional significance, further requirements on $\Delta R_{\tau\tau}$ as well as the invariant mass of the ditau system seem to be beneficial (cf. figure 6.12(c)). Moreover, a veto on $N_{b-\text{jets}}$ increases the significance appreciably (cf. figure 6.12(d)). As b quarks are produced in any t quark decay and in many other QCD processes, it is expectable that a significant amount of background is reduced by this cut.

Similarly to the previous channel, two-dimensional significance scans are performed with respect to the behaviour of $\Delta R_{\tau\tau}$ versus the invariant ditau mass $m_{\tau\tau}$. A further tightening of the τ^{prob}-product cut to $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) > -0.90$ is tested and the effect of an b-jet veto is compared in all scenarios (cf. figure 6.13). The highest stable significance is achieved in the following event selection (cf. figure 6.13(d)):

$$\boxed{\log\left(\prod \tau_i^{\text{prob}}\right) > -0.90, \quad \sum p_{\text{T}}(\text{lep}_{1\text{-}4}) > 210\,\text{GeV}, \quad N_{b-\text{jets}} = 0, \quad \Delta R_{\tau\tau} < 1.925 \quad \& \quad m_{\tau\tau} > 51\,\text{GeV}}$$

$hh \to 4\tau$	Higgs	Multiboson	$t\bar{t}$ production	Others	Combined	$S/\sqrt{S+B}$
2.06	6.24	19.39	5.30	0	30.93	0.36

Table 6.3: The predicted events of different processes in the signal region of the $3h1\ell$ channel are displayed. The processes V + Jets, $Z \to \tau\tau$, Single Top and fakes are summarised in “others” and are all estimated to be completely discarded by the selection. For all events, $\prod \tau_{i=1-3}^{\text{prob}} > -0.90$, $\sum p_{\text{T}}(\text{lep}_{1\text{-}4}) > 210\,\text{GeV}$, $N_{b-\text{jets}} = 0$, $\Delta R_{\tau\tau} < 1.925$ and $m_{\tau\tau} > 51\,\text{GeV}$ is demanded.

The expected amount of events partitioned in their different processes is presented in table 6.3. Again, single Higgs and multi-tau events serve as the major background. In total **a significance of 0.36 is achieved**. Again, this result is meant to be a careful first estimation of the general capability of the $hh \to 4\tau$ channel rather than a doubtless result. However, negative influences on the significance such as an underestimation of remaining background events due to limited statistics in the simulations might as well be balanced by positive effects. For instance, it is indicated in figure 6.4(d) that the MC simulations

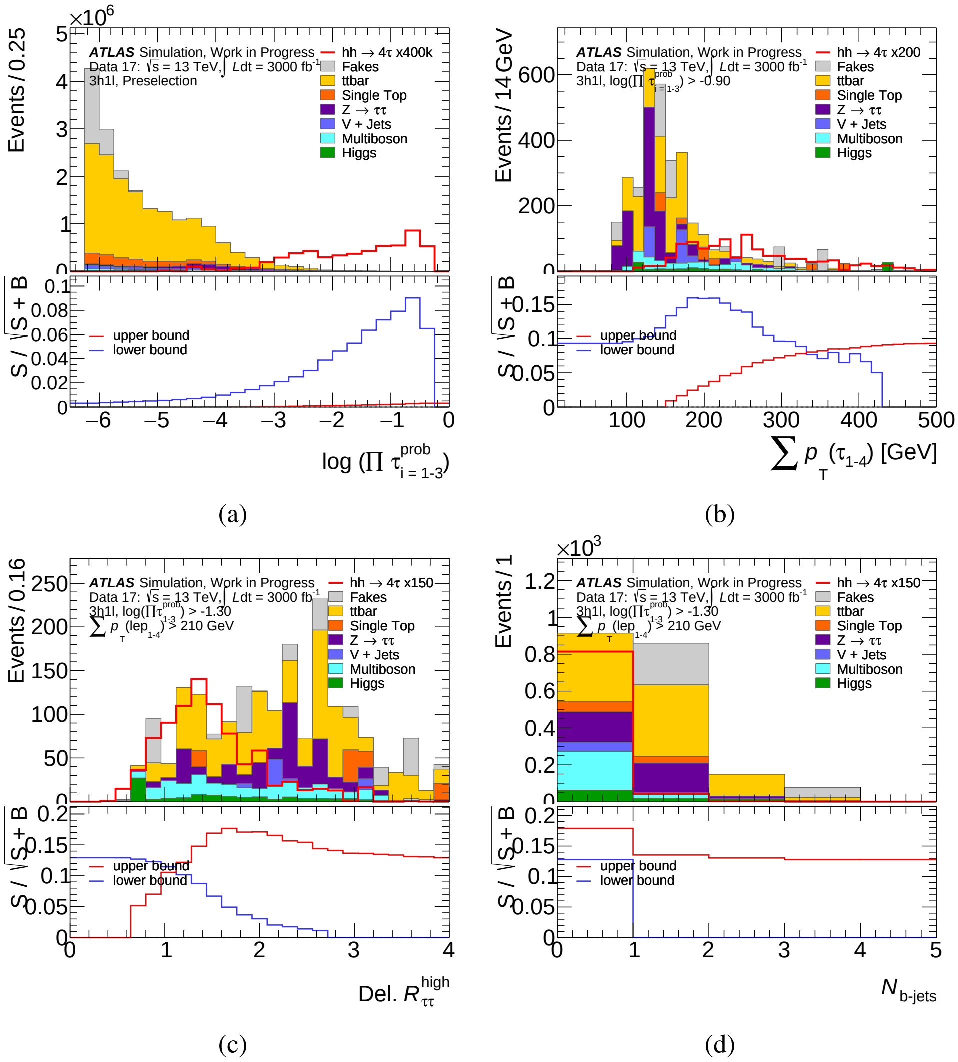
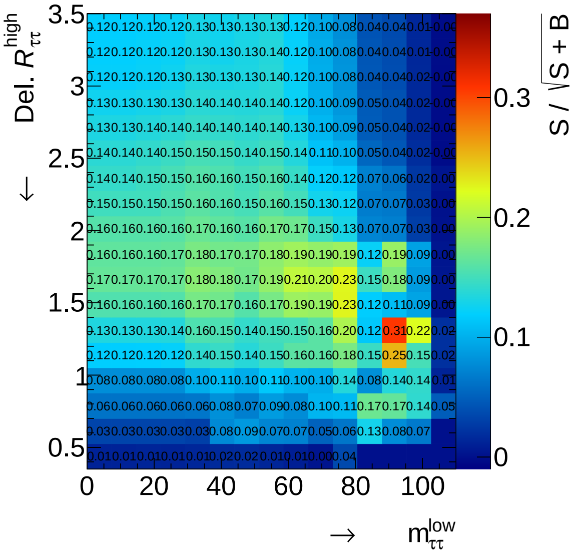
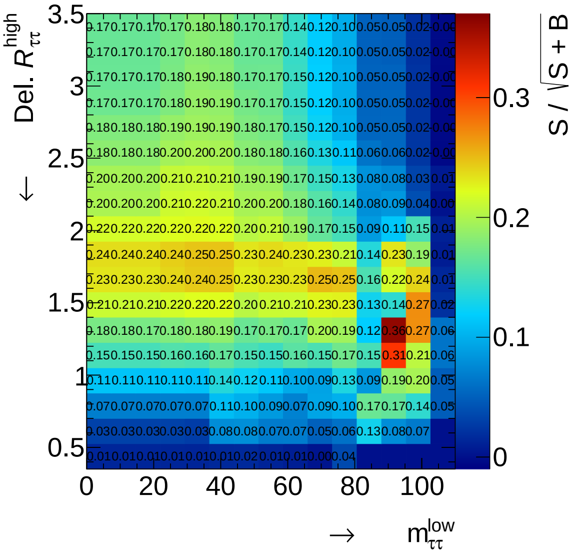


Figure 6.12: The significance scans of the most interesting observables in the $3h1\ell$ channel are shown. In the upper half of each plot, a usual stack plot of all background processes and the artificially enhanced signal process are shown. The lower half visualised the significance alteration depending on changing cuts. They are explained in detail in the caption of figure 6.10.

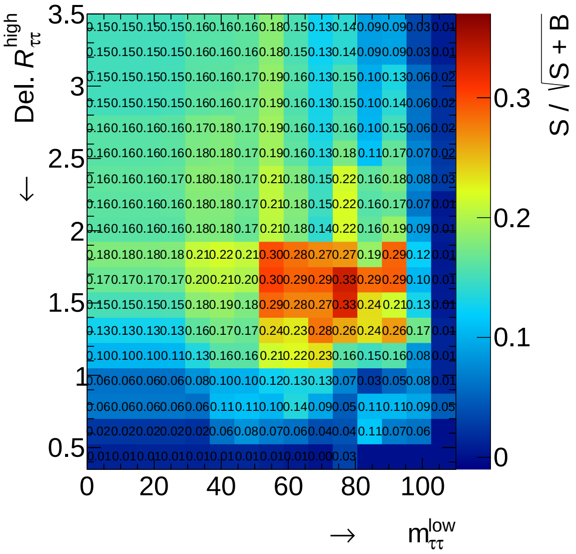
In (a), it is visualised that a minimal requirement on $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right)$ is recommendable. Correspondingly, in (b) $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) > -0.90$ is demanded. To further rise the significance, a lower bound of the scalar sum of all tau and light lepton momenta like $\sum p_{\text{T}}(\text{lep}_{1\text{-}4}) > 210\,\text{GeV}$ is advisable. In (c) and (d), cuts on both previous observables are applied and it is visualised that an upper bound of $\Delta R_{\tau\tau}$ or $N_{b\text{-jets}}$ will increase the significance further.



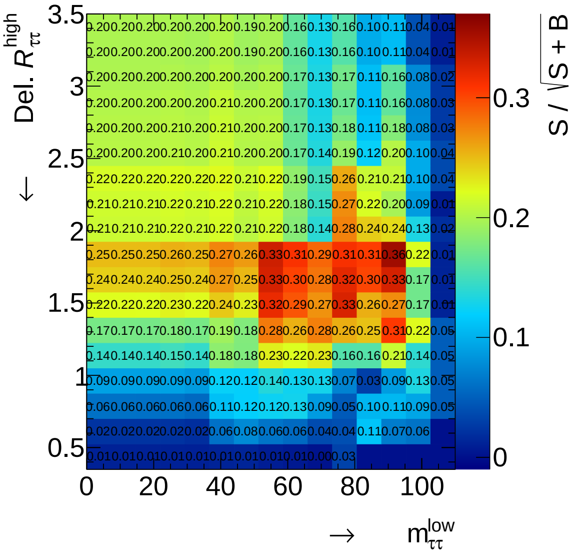
(a) $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) > -1.30$, $\sum p_T(\tau_{1\text{-}4}) > 210\,\text{GeV}$



(b) $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) > -1.30$, $\sum p_T(\tau_{1\text{-}4}) > 210\,\text{GeV}$, $N_{b-\text{jets}} = 0$



(c) $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) > -0.90$, $\sum p_T(\tau_{1\text{-}4}) > 210\,\text{GeV}$



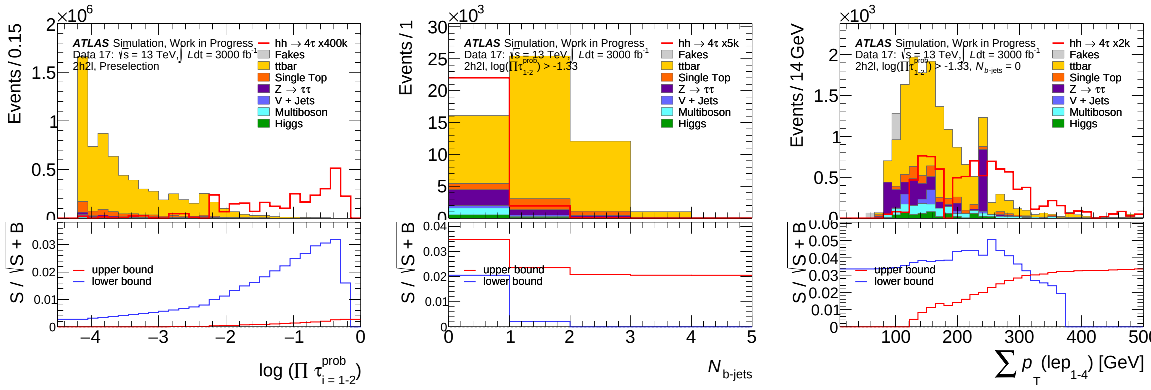
(d) $\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) > -0.90$, $\sum p_T(\tau_{1\text{-}4}) > 210\,\text{GeV}$, $N_{b-\text{jets}} = 0$

Figure 6.13: Two-dimensional significance scans of the lower invariant mass of both ditau systems $m_{\tau\tau}^{\text{low}}$ against the higher of both tau distances $\Delta R_{\tau\tau}^{\text{high}}$ are shown for the $3h1\ell$ channel. Each bin represents a *cut from below* on $m_{\tau\tau}^{\text{low}}$ of the corresponding value and a *cut from above* on $\Delta R_{\tau\tau}^{\text{high}}$. Hence, only events which are located in bins to the right or below each corresponding bin are used to calculate the significance $S/\sqrt{S+B}$. This is indicated by the arrows in each axis title.
For each plot, slightly different selection criteria are applied. The highest significances are observed for the selection in (d) by additionally demanding roughly $m_{\tau\tau}^{\text{low}} > 51\,\text{GeV}$ and $\Delta R_{\tau\tau}^{\text{high}} < 1.9$.

tend to overestimate the actual events for high values of $\log\left(\prod \tau^{\text{prob}}_{i=1-3}\right)$. Hence, in a more complete treatment of the background processes in which they are compared to data in specific control regions, they might scale down slightly which would result in a higher significance.[5]

6.4.3 $2h2\ell$ Channel

Even though the $2h2\ell$ channel is not expected to benefit as much as the previous two channels from the novel tau identification method of chapter 5, a superficial investigation of a possible significance gain due to this channel is investigated. For the sake of simplicity, a very analogous procedure to the previous channels is chosen.



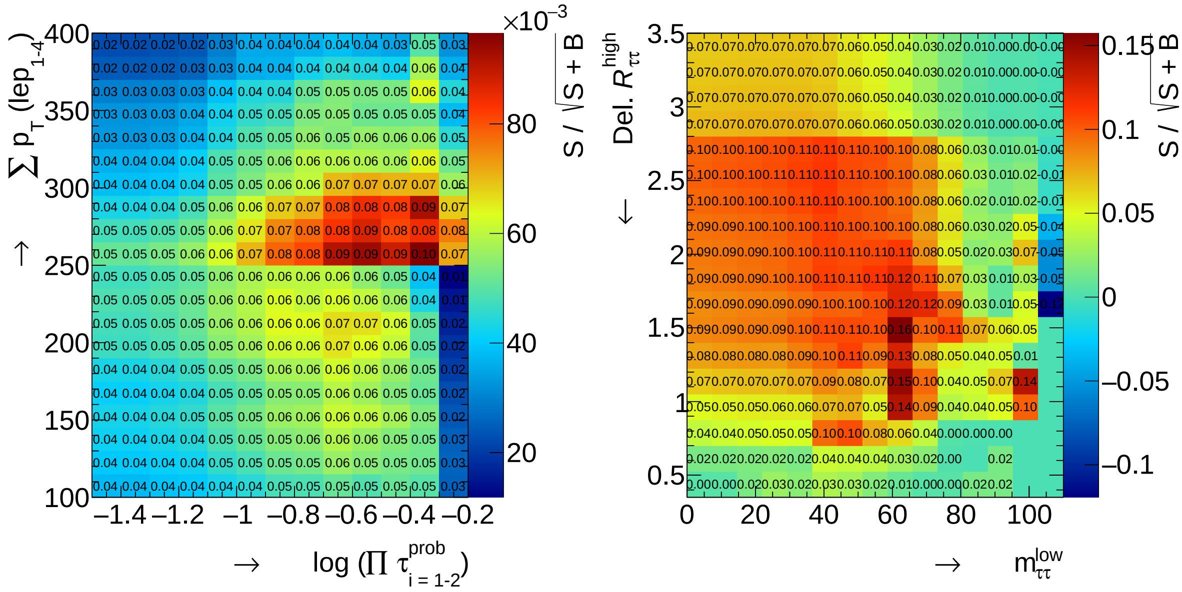
(a) (b) (c)

Figure 6.14: The significance scans of the most interesting observables in the $2h2\ell$ channel are shown. In the upper half of each plot, a usual stack plot of all background processes and the artificially enhanced signal process are shown. The lower half visualised the significance alteration depending on changing cuts. They are explained in detail in the caption of figure 6.10.
Analogously to the cited figure of the $4h0\ell$ channel as well as in the $3h1\ell$ channel (cf. figure 6.12), a lower threshold on $\log\left(\prod \tau^{\text{prob}}_{i}\right)$ (a) and on $\sum p_{\text{T}}(\text{lep}_{1\text{-}4})$ (c) rises the significance. By rejecting all events which contain b-jets, the significance is increased further as visualised in figure (b).

First, again a lower threshold on $\log\left(\prod \tau^{\text{prob}}_{i=1-2}\right)$ is strongly advisable (cf. figure 6.14(a)). Moreover, the significance gain due to the b-jet veto exceeds most other quantities (cf. figure 6.14(b)). The dominance of $t\bar{t}$ production in the $2h2\ell$ background explains the positive properties of such a b-jet veto as it especially rejects t quark events. Therefore, before further investigations $N_{b-\text{jets}} = 0$ is demanded. Analogously to the previous channels, the sum of all lepton transverse momenta singles out to be the observables with highest additional separation power (cf. figure 6.14(c)). Again, two-dimensional significance scans are produced, first to investigate a possible ideal combination of requirements on $\log\left(\prod \tau^{\text{prob}}_{i=1-2}\right)$ and $\sum p_{\text{T}}(\text{lep}_{1\text{-}4})$ (visualised in figure 6.15(a)) and second of $m_{\tau\tau}$ vs. $\Delta R_{\tau\tau}$ after several different cuts of the former two variables. The best significance is achieved for the following requirements (cf. figure 6.15(b)):

$$\log\left(\prod \tau^{\text{prob}}_{i=1-2}\right) > -0.72, \quad \sum p_{\text{T}}(\text{lep}_{1\text{-}4}) > 200\,\text{GeV}, \quad N_{b-\text{jets}} = 0, \quad \Delta R_{\tau\tau} < 1.575 \quad \& \quad m_{\tau\tau} > 58.5\,\text{GeV}$$

[5] Such a simplified conclusion can only be drawn if all background processes scale down globally and if their qualitative behaviour is not altered.



(a) $\log\left(\prod \tau^{\text{prob}}_{i=1-2}\right) > -1.33$, $N_{b-\text{jets}} = 0$

(b) $\log\left(\prod \tau^{\text{prob}}_{i=1-2}\right) > -0.72$, $\sum p_{\text{T}}(\tau_{1\text{-}4}) > 200\,\text{GeV}$, $N_{b-\text{jets}} = 0$

Figure 6.15: Two-dimensional significance scans of either $\log\left(\prod \tau^{\text{prob}}_{i=1-2}\right)$ against $\sum p_{\text{T}}(\text{lep}_{1\text{-}4})$ (a) or the lower invariant mass of both ditau systems $m^{\text{low}}_{\tau\tau}$ against the higher of both tau distances $\Delta R^{\text{high}}_{\tau\tau}$ (b) are shown for the $2h2\ell$ channel. Different selection cuts are applied in both figures and are listed in their captions. In (a), each bin represents a *cut from below* on the corresponding value of the x or y axis (as indicated by the arrows in the axis titles). Thus, only events in the upper right corner of each bin are used to calculate the significance $S/\sqrt{S+B}$. In (b), each bin corresponds to a *cut from below* on $m^{\text{low}}_{\tau\tau}$ of the corresponding value and a *cut from above* on $\Delta R^{\text{high}}_{\tau\tau}$. Thus, events of the lower right corner are used to calculate the significance.
Plots similar to the one in (b) were studied for several different selection cuts. Interestingly, the one displayed here leads to the highest significance of roughly 0.16 by additionally demanding $\Delta R_{\tau\tau} < 1.575$ and $m_{\tau\tau} > 58.5\,\text{GeV}$ although figure (a) indicates that a stricter cut combination of roughly $\log\left(\prod \tau^{\text{prob}}_{i=1-2}\right) > -0.7$ and $\sum p_{\text{T}}(\text{lep}_{1\text{-}4}) > 250\,\text{GeV}$ would be advisable. This motivates the multidimensional significance analysis.

$hh \to 4\tau$	Single Higgs	Multiboson	V + Jets	Other	Combined	$S/\sqrt{S+B}$
0.62	2.52	7.50	3.79	0	13.81	0.16

Table 6.4: The predicted events of different processes in the signal region of the $2h2\ell$ channel are displayed. The processes $Z \to \tau\tau$, single top $t\bar{t}$ production and fakes are included in "other" and are estimated to be rejected by the selection. The events displayed here fulfil $\log\left(\prod \tau^{\text{prob}}_{i=1-2}\right) > -0.72$, $\sum p_{\text{T}}(\tau_{1\text{-}4}) > 200\,\text{GeV}$, $N_{b-\text{jets}} = 0$, $\Delta R < 1.575$ and $m_{\tau\tau} > 58.5\,\text{GeV}$.

As summarised in table 6.4, the main background in the signal region is formed by single Higgs and multiboson events. Not even one signal event is predicted in this region. As real events can only be whole numbers, the $2h2\ell$ channel might be only suitable in combination with other channels. Furthermore, in comparison to the other two channels, the significance is considerably lower. This can partly be explained by lower probabilities for leptonically decaying taus to reach required momentum thresholds for the reconstruction of light leptons. The identification requirements for light leptons in general are not loosened in this thesis as it is done for tau candidates. The very successful new tau identification method hardly develops its full potential in the $2h2\ell$ channel but it is conceivable that such a ID-score product could be extended to light leptons.

By comparing the event predictions to data, it is revealed that the actual amount of events is strongly overestimated by the MC simulations in the considered phase space region (cf. figure 6.16). Such a mismodelling must have its origin in a massive overestimation of one or several of the background processes and even technical problems suggest itself.[6] Thus, a sophisticated handling of background processes and their rescaling in control regions is essential in this channel. In an ideal case, the mismodelling can be solved by a significant downscaling of the background without further sensible modifications of the overall background shape. If so, the qualitative conclusions drawn in this section still remain true. Moreover, the expectable downscaling of background processes might enable much higher significances in the $2h2\ell$ channel and the **achieved significance of 0.16** only serves as a lower limit.

Nevertheless, it cannot be ruled out in the scope of this thesis that such a mismodelling is caused by some more physical processes in this channel which are not yet understood. In this case, statements about this channel are difficult.

In any case, results of the $2h2\ell$ channel have to be used carefully and more investigations are needed.

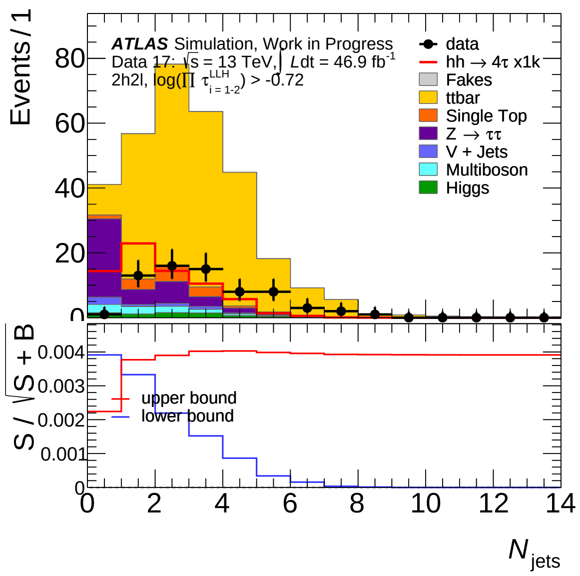


Figure 6.16: The amount of predicted events by MC simulations are compared to data in the $2h2\ell$ channel. A massive discrepancy is observed. The predictions overestimate the actual amount of measured data nearly by a factor of four. The overall shape of the number of jets, however, seems to be comparable.

6.5 Multivariate Analysis with Boosted Decision Trees

As already introduced in section 3.4.2, boosted decision trees are a powerful tool to further separate two classes (like signal and background) from each other. As a first step, observables with possible separation power are defined and serve as the input of the BDT. Next, hundreds of decision trees are created which differ in their splitting of subsets. Ideally, herewith all interesting cut combinations with respect to the input variables are covered. In the boosting process, weights are assigned to each decision tree depending

[6] For example, a cut which is only applied on the data samples would explain such a discrepancy. As data and MC were processed with exactly the same program, such a straight-forward explanation is ruled out.

on its separation power. In summary, especially when multiple input variables are involved, BDTs are expected to enable a significant separation improvement and are widely used in data analysis.[7]

The capability of this tool is studied in this section. For this purpose, multiple BDTs are trained to separate the $hh \to 4\tau$ signal from background processes. The TMVA package [106] is used for the training and several BDT architectures are tested. For instance, apart from other configurations, the number of decision trees N is altered from 400 to 1200. Moreover, the maximal tree-depth D is varied and the number of grid points P in the input variable range is altered to find optimal cuts in the node splitting. These unique features of the different BDTs architectures used in this thesis are referred to as $\mathrm{BDT}^{N}_{D,P}$.

In principle, the minimal amount of degrees of freedom which suits the problem is desirable. More degrees of freedom (like a very high amount of trees or very deep trees) might cause overfitting and suffer from worse separation power. This becomes noticeable when the performance of the BDT is much better on the training sample than on the test sample.

Usually, hundreds or thousands of different architectures are systematically trained and compared. With this trial-and-error procedure, another few percent of efficiency can be gained. This chapter, however, is more meant to be a proof of principle and a very detailed study of different BDT architectures is relinquished.

6.5.1 One-dimensional BDT Approach

Analogously to section 6.4, the signal region is split into the $4h0\ell$, $3h1\ell$ and $2h2\ell$ channel. To utilise the full available statistics, each sample is divided into ten subsamples depending on the event number. Nine out of ten samples are then trained and the resulting BDT is used to assign BDT scores to the tenth sample. Hereby, the full amount of events can be used for training and the latter analysis while full orthogonality is ensured between training and testing samples.

Moreover, the tau reconstruction properties are reduced to the very low settings as introduced in section 6.4 as a trial. In the one-dimensional BDT training of this subsection, the “normal” tau candidate selection with $p_\mathrm{T} > 20\,\mathrm{GeV}$ remains beneficial. Nevertheless, the additional statistics of the very loose taus is found to be beneficial such that the BDTs are trained with very loose taus and after the training the tightened tau momentum cut is applied. In a two-dimensional approach as introduced in section 6.5.2, the loosened requirement of $p_\mathrm{T} > 15\,\mathrm{GeV}$ gives the best results. In both cases, however, the differences are small such that they are not distinguished in broad detail.

Additionally to the observables used in the previous section, several other quantities are tested as input variables. All observables which indicate a possible significance gain are utilised. In figure 6.17 a selection of additional input variables to the BDTs is shown. The complete list of used input variables in each channel is specified in table 6.5.

A small selection of results is displayed in figure 6.18. On average, small BDT scores are assigned to the dominant fake-tau-enriched background processes like $t\bar{t}$ production, $Z \to \tau\tau$ or data-driven fakes. Events with a higher probability of four truth leptons like single Higgs or multiboson processes as well as signal processes peak at high BDT scores, which enables a very pure signal region defined by a lower threshold of this score. With such a threshold, for the $4h0\ell$ channel the significance compared to the previous signal-region selection is improved considerably and values of more than 0.7 are achieved (cf. figure 6.18(a) to 6.18(c)).

The order of magnitude of the previously achieved significances (cf. section 6.4) in the $4h0\ell$ channel is validated. While in the previous event selection, many highly weighted positive or negative $Z \to \tau\tau$ events pollute the signal region of the $4h0\ell$ channel, the region with very high BDT scores is dominated

[7] E.g. in the $hh \to b\bar{b}\tau\tau$ channel, BDTs are used as the final discriminant of signal and background [51].

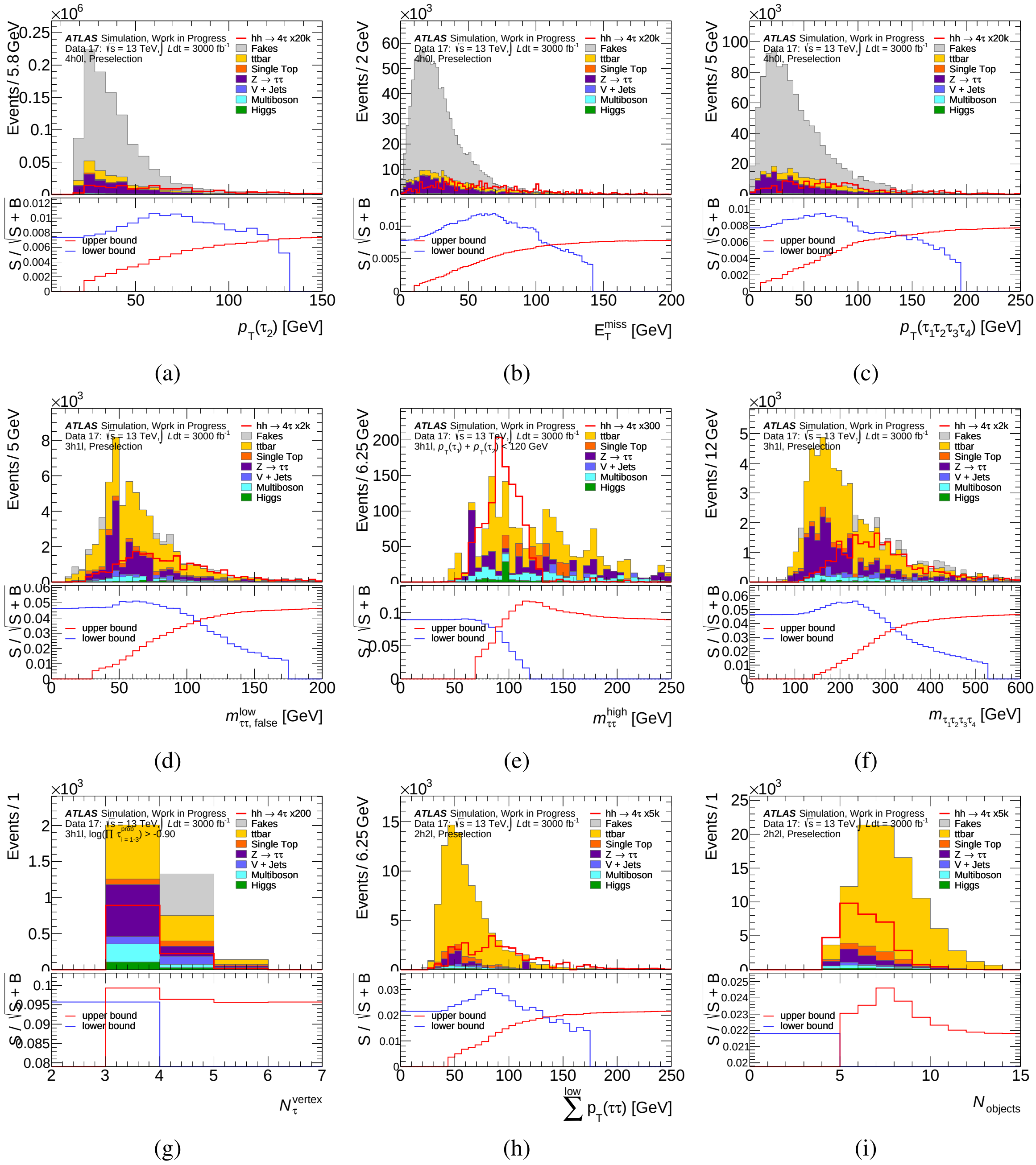


Figure 6.17: Several observables and their possible significance gain are displayed. All of them either recommend a cut from below or from above as indicated by the rising blue or red significance lines in the lower plots. In combination with other cuts, they were outperformed by other observables and thus, they were not used in section 6.4. A BDT, however, enables a multivariate analysis in which their possible additional separation power is utilised.

	Observables	$2h2\ell$	$3h1\ell$	$4h0\ell$
Single tau related	$p_T(\tau_1)$	✓	✓	✓
	$p_T(\tau_2)$	✓	✓	✓
	$p_T(\tau_3)$		✓	
Ditau related	$\Delta R_{\tau\tau}^{\text{high}}$	✓	✓	✓
	$m_{\tau\tau}^{\text{low}}$	✓	✓	✓
	$m_{\tau\tau}^{\text{high}}$	✓	✓	✓
	$m_{\tau\tau,\text{ false}}^{\text{low}}$	✓	✓	✓
	$\sum^{\text{low}} p_T(\tau)p_T(\tau)$	✓	✓	✓
Tetratau related	$m_{\tau\tau\tau\tau}$	✓	✓	✓
	$\sum p_T(\text{lep}_{1-4})$	✓	✓	✓
	ΔR_{hh}		✓	✓
	$p_T(\tau\tau\tau\tau)$			✓
other	$\log\left(\prod \tau_i^{\text{prob}}\right)$	✓	✓	✓
	$N_{b-\text{jets}}$	✓	✓	
	N_{τ}^{Vertex}		✓	
	N_{objects}	✓		
	E_T^{miss}			✓

Table 6.5: BDT input variables depending on the different channels

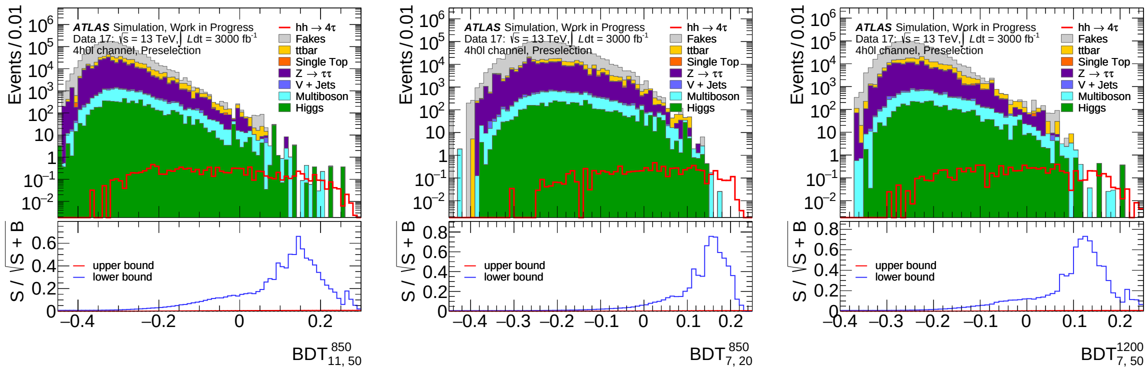
by single Higgs or multiboson events with much more statistics. Moreover, the BDT training cannot distinguish between positive and negative event weights for technical reasons. Therefore, negatively weighted events cannot simulate a cleaner signal region by equalising positive event weights of the same process but are rejected by the BDT equally. In summary, the reliability of the achieved significance is supported by the BDT results.

Cut	$hh \rightarrow 4\tau$	Single Higgs	Multiboson	Other	Combined	$S/\sqrt{S+B}$
$\text{BDT}_{11,50}^{850} > 0.14$	1.17	1.12	0	0	1.12	0.77
$\text{BDT}_{7,20}^{850} > 0.145$	0.78	0.04	0	0	0.04	0.70
$\text{BDT}_{7,50}^{1200} > 0.095$	1.32	1.63	1.84	0	3.47	0.60
$\text{BDT}_{7,50}^{1200} > 0.115$	1.05	0.83	0	0	0.83	0.77

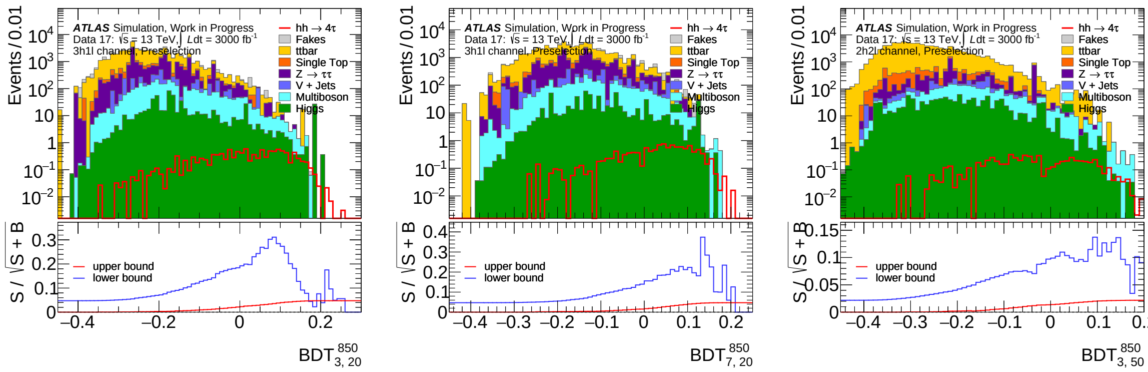
Table 6.6: The predicted events of different processes in the signal region formed by different trained BDTs for the $4h0\ell$ channel are displayed. The processes V + Jets, $Z \rightarrow \tau\tau$, Single Top $t\bar{t}$ production and Fakes are summarised in "others" and are estimated to be completely discarded by the selections. In comparison to the previous signal selection summarised in table 6.2, considerably higher significances are gained by several different BDTs.

In table 6.6, the amount of events fulfilling specific BDT thresholds is displayed for the $4h0\ell$ channel. Again, single Higgs and partly multiboson events serve as the only remaining background processes. In comparison to the previous selection "by hand" in table 6.2, the significance gain is mainly due to a higher amount of expected signal events in the final selection.

The BDTs with best separation properties for the $3h1\ell$ and $2h2\ell$ channel are shown in figure 6.18(d) to 6.18(f). Corresponding signal region cuts are summarised in table 6.7. In both channels, comparable significances as in section 6.4 are achieved, which again strengthens the assumption that their order of magnitude is assumed realistically. In contrast to the $4h0\ell$ channel, however, the previous results are hardly improved or even slightly worse.



(a) $4h0\ell$ channel (b) $4h0\ell$ channel (c) $4h0\ell$ channel



(d) $3h1\ell$ channel (e) $3h1\ell$ channel (f) $2h2\ell$ channel

Figure 6.18: Selected BDT scores are shown which feature the best signal-to-background separation for the $4h0\ell$ (a), (b) and (c), the $3h1\ell$ (d) and (e) and the $2h2\ell$ (f) channel. All plots illustrate that fake-tau-dominated events like $t\bar{t}$ production, $Z \to \tau\tau$ or data-driven fakes peak at low BDT scores while signal events or events with a higher probability of four truth leptons like single Higgs or multiboson processes peak at high BDT scores. Moreover, the achivable significance by cutting on each BDT score is visualised in the blue line of the lower plots.

Cut	$hh \to 4\tau$	Single Higgs	Multiboson	$Z \to \tau\tau$	$t\bar{t}$	Comb.	$S/\sqrt{S+B}$
			$3h1\ell$ channel				
BDT > 0.072	3.76	39.76	32.03	6.92	48.66	127.38	0.33
BDT9 > 0.13	0.83	1.30	2.81	0	0	4.11	0.37
			$2h2\ell$ channel				
BDT2 > 0.105	0.77	3.33	9.52	1.42	15.93	30.20	0.14

Table 6.7: The predicted events of different processes in the signal region formed by different trained BDTs for the $3h1\ell$ (upper two rows) and $2h2\ell$ channel (third row) are displayed. The processes V + Jets, Single Top and data-driven fakes are not included as they are estimated to be completely discarded by the selections. In comparison to the previous signal selections summarised in table 6.3 and 6.4, comparable significances are gained.

6.5.2 Two-dimensional Approach with Kinematic BDTs

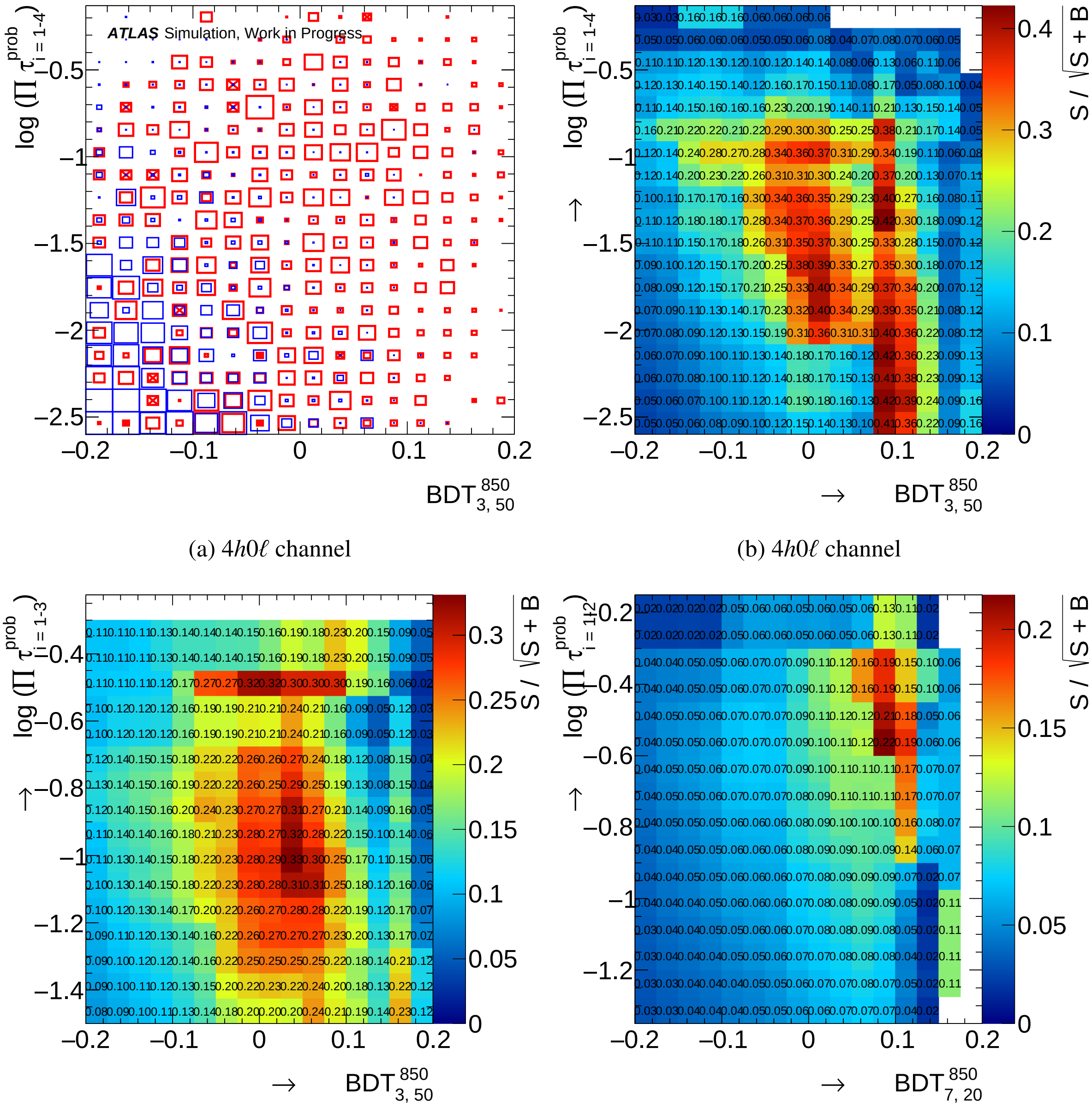
In all three channels, $\log\left(\prod \tau_i^{\text{prob}}\right)$ is the observable with strongest separation power. This is natural, as this quantity directly corresponds to tau identification properties and is designed to reject fake-tau events. As already tested in section 5.3, this product is only slightly correlated to other properties of the tau candidates such as their momenta or direction.

Therefore, it might be beneficial to train a BDT only on kinematic quantities such that the algorithm is "enforced" to separate signal from background with respect to their ditau masses, tau momenta or tau distances. The $\log\left(\prod \tau_i^{\text{prob}}\right)$ information can then be exploited independently and together, higher significances might be achieved. Correspondingly, BDTs are trained in all three channels with the same input variables as summarised in table 6.5 except for the $\log\left(\prod \tau_i^{\text{prob}}\right)$. Their separation power combined with the independent capability of $\log\left(\prod \tau_i^{\text{prob}}\right)$ is visualised in two-dimensional significance plots in figure 6.19 and the event predictions in possible signal regions are summarised in table 6.8.

Cut	$hh \to 4\tau$	Higgs	Multiboson	V + Jets	$Z \to \tau\tau$	$t\bar{t}$	Fakes	Comb.	$S/\sqrt{S+B}$
				$4h0\ell$ channel					
1	1.74	3.70	8.10	0	0	0	2.00	13.80	0.44
2	0.89	1.84	0.14	0	0	0	0	1.978	0.53
				$3h1\ell$ channel					
3	3.12	37.01	32.50	1.82	4.88	5.31	0	81.52	0.34
4	1.40	4.03	7.82	0	0	0	0	11.85	0.38
				$2h2\ell$ channel					
5	0.74	1.03	7.55	1.82	0	0	0	10.40	0.22

Table 6.8: The predicted events of different processes in the signal region formed by a two-dimensional scan of kinematic BDTs and $\prod \tau_i^{\text{prob}}$ for all channels are displayed. Not explicitly included are the event predictions of single Top events as they are estimated to be discarded by the selections.
The cuts applied for each selection are $\left[\text{BDT}_{3,50}^{850} > 0.0 \;\&\; \log\left(\tau_{i=1-4}^{\text{prob}}\right) > -1.95\right]$ for cut 1, $\left[\text{BDT}_{5,20}^{850} > 0.1 \;\&\; \log\left(\tau_{i=1-4}^{\text{prob}}\right) > -2.15\right]$ for cut 2, $\left[\text{BDT}_{3,50}^{850} > 0.024 \;\&\; \log\left(\tau_{i=1-3}^{\text{prob}}\right) > -1.05\right]$ for cut 3, $\left[\text{BDT}_{5,20}^{850} > 0.1 \;\&\; \log\left(\tau_{i=1-3}^{\text{prob}}\right) > -1.275\right]$ for cut 4 and $\left[\text{BDT}_{7,20}^{850} > 0.077 \;\&\; \log\left(\tau_{i=1-2}^{\text{prob}}\right) > -0.6\right]$ for cut 5.
The corresponding figures for the $\text{BDT}_{5,20}^{850}$ score are presented in figure A.13. In comparison to the BDTs which were trained on the full set of input variables (cf. tables 6.6, 6.7), the two-dimensional method leads to diverse results. This is analysed in the text.

The gain of this two-dimensional approach is varying. In the $2h2\ell$ channel, a considerable rise of the maximal significance from previously 0.16 to 0.22 is observed. Unfortunately, this is not replicated



(a) $4h0\ell$ channel

(b) $4h0\ell$ channel

(c) $3h1\ell$ channel

(d) $2h2\ell$ channel

Figure 6.19: The two-dimensional distribution (a) and significance scans of the trained BDTs against the $\log\left(\prod \tau_i^{\text{prob}}\right)$ for the $4h0\ell$ ((a) & (b)), $3h1\ell$ (c) and $2h2\ell$ channel (d). In (a), the normalised signal in red accumulates at the upper right corner while background events in blue show the opposite behaviour. The corresponding plots for the leptonic channels are presented in figure A.12.
Each bin of the significance plots represents a *cut from below* on the corresponding value of the x or y axis. Thus, only events in the upper right corner of each bin are used to calculate the significance $S/\sqrt{S+B}$. Detailed interpretations of the significance scans are given in the text.

in the other channels. While the maximal significances for the 2D approach in the $3h1\ell$ channel are comparable to the 1D case (cf. table 6.7), they cannot compete with previous results in the $4h0\ell$ channel (cf. table 6.6). In the latter channel, only significances of roughly 0.44 to 0.53 are achieved, while in the first approach, significances up to 0.77 are observed. This indicates a higher correlation of the $\log\left(\prod \tau^{\text{prob}}_{i=1-4}\right)$ value to other kinematic variables in the $4h0\ell$ channel. As the dominant background in this channel consists of fakes (before cuts), it is imaginable that they feature significantly other averaged kinematic observables. Thus, a cut on the tau identification property $\log\left(\prod \tau^{\text{prob}}_{i=1-4}\right)$ might be correlated to the demand of a high scalar sum of transverse tau momenta. As a result, these observables should not be treated completely independently as it is done in the 2D approach. Instead, their separation power and their correlations are ideally taken into account by training one BDT with all information.

This correlation decreases for less taus in each event, which explains the comparably better results in the $3h1\ell$ channel and especially the $2h2\ell$ channel.

6.6 Statistical Interpretation

In section 6.4 and 6.5, an amount of possible final signal selections is presented. In each channel, different significances are observed. They depend on the decision of how many signal events should be kept. Especially by selecting a very strict signal region, the question arises, how robust the observed significances are. In this analysis, it is tried to not select a signal region with an artificially high significance due to limited statistics of background events. Discontinuous significance distributions are an indicator for this and event selections based on such significances are avoided.

As an example, the maximal significance for the BDT score of the $3h1\ell$ channel in figure 6.18(d) seems to be much more reliable than in figure 6.18(e) in which a steep rise of nearly $\Delta(S/\sqrt{S+B}) = 0.1$ is observed. One the one hand, such a behaviour strongly indicates that some highly weighted background events pollute this phase space region and limited statistics causes such a discontinuity. On the other hand, the BDT is exactly trained not to assign high scores to the background and thus it might be expectable for a majority of background events not to exceed a certain threshold.

Under the assumption that all three channels are completely orthogonal and all highest possible significances are valid, an overall significance of the $hh \to 4\tau$ channel of 0.89 is achieved. In a more conservative approach, at least a combined significance of 0.66 seems to be realistic.[8] In any case, these values are deduced without taking any systematic uncertainties into account. In the prospect study for the HL-LHC [49], the expected significance of the $hh \to b\bar{b}\tau\tau$ channel (which is assumed to be most comparable to the signal channel of this work) drops by 16 % when systematic uncertainties are included. Thus, a similar reduction might be expected for the $hh \to 4\tau$ channel, as well. In this case, the corresponding reduced significance values of 0.55 to 0.75 for this channel and a dataset of $3\,000\,\text{fb}^{-1}$ are comparable to the expected significance of the $hh \to b\bar{b}b\bar{b}$ channel in the named study. Here, a significance of 0.61 is estimated when systematic uncertainties are taken into account. The significances are further discussed in the next chapter.

Additionally, the production of fake-taus in conjunction with one or two light leptons seems to be overestimated by MC predictions. It should be abstained from speculating but it does not seem to be too far-fetched to assume that a proper handling of the background in control regions lead to an appreciable downscaling of major background distributions, especially in the $2h2\ell$ channel. Higher significances might be possible.

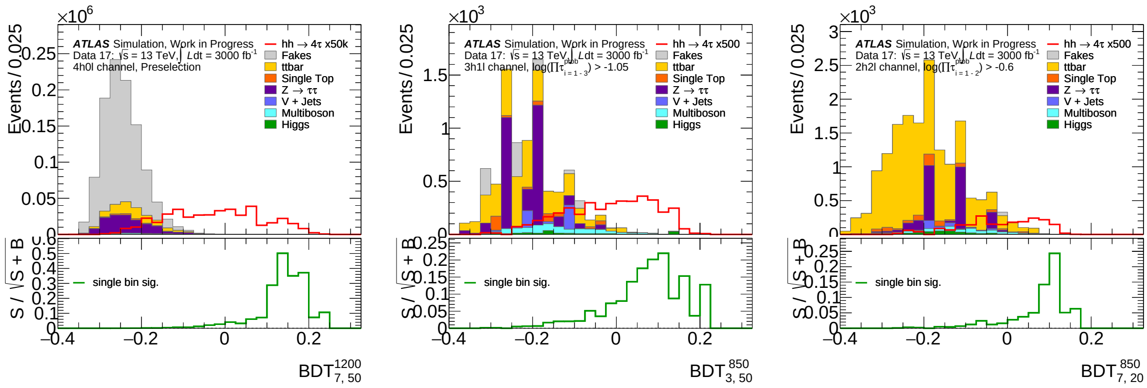
[8] For the optimistic scenario, a significance of 0.77 for the $4h0\ell$ channel is assumed. In the $3h1\ell$ and $2h2\ell$ channel, 0.38 and 0.22 are taken respectively. In the conservative scenario, the highest observed values without any spikes in the significance curve are taken. These read roughly 0.55 (up to 0.60) in the $4h0\ell$ channel, 0.34 ($3h1\ell$) and 0.16 ($2h2\ell$).

All results presented so far are simple counting experiments in each signal region. In contrast to that, the contained information of the observed signal and background distribution in each observable can be utilised and a shape fit of each distribution is performed. This is done by the `HistFitter` framework [111]. In here, the most promising signal regions with a relatively continuous significance curve are taken as signal regions. For the $4h0\ell$ channel, the 1D BDT score distribution as visualised in figure 6.18(c) is used. For the $3h1\ell$ and $2h2\ell$ channel, the two-dimensional distributions as displayed in figure 6.19(c) and 6.19(d) are considered. In the latter two channels, the most promising $\log\left(\prod \tau_i^{\text{prob}}\right)$ cut reads

$$\log\left(\prod \tau_{i=1-3}^{\text{prob}}\right) > -1.05 \quad \text{or} \quad \log\left(\prod \tau_{i=1-2}^{\text{prob}}\right) > -0.60 \tag{6.4}$$

respectively.

In all cases, the very low BDT score values are thought to serve as a control region for the background processes. In here, they dominate significantly and hardly any signal events are expected. As the full range of the BDT is utilised, the fitting program receives enough background information to exploit their expected shapes. The bins with high BDT scores contain a comparably high signal to background ratio and serve as signal-region bins. The resulting distributions are presented in figure 6.20.



(a) $4h0\ell$ channel (b) $3h1\ell$ channel (c) $2h2\ell$ channel

Figure 6.20: Input regions for the `HistFitter` shape fit for all channels. As the significance in each bin is essential for the upper limit calculation, in the lower plot, the significances achieved by each single bin are presented. In all three channels, bins with comparably high significances exist.

Eventually, an upper limit scan is performed. The signal cross section is increased systematically which leads to higher signal event predictions. By assuming that the MC samples describe the measured data perfectly ("Asimov data"), the signal strength is singled out which predicts an amount of signal events that can be excluded at 95 % confidence level.

The 2017 data considered in this thesis can be extrapolated to estimate the exclusion capability of arbitrary amounts of data. For example, for the full Run II dataset with an integrated luminosity of $147\,\text{fb}^{-1}$, an upper limit on the signal cross section in comparison to the expected SM cross section of $\sigma_{\text{ggF}}/\sigma_{\text{ggF}}^{\text{SM}} = 17.1^{+9.6}_{-5.8}$ is calculated. This result as well as similar upper limit expectations for other integrated luminosities are summarised in table 6.9.

The expected shapes of signal and background can further be utilised by considering more than the three one-dimensional signal regions. As the considered phase space regions of the $3h1\ell$ and $2h2\ell$ channel are already extracted from a two-dimensional phase-space scan (cf. figure 6.19(c) and 6.19(d)), it is straight-forward to exploit the whole phase space information. For this purpose, instead of selecting

$\int \mathcal{L}\,\mathrm{d}t$	$\sigma_{\mathrm{ggF}}/\sigma^{SM}_{\mathrm{ggF}},^{+1\sigma}_{-1\sigma},^{+2\sigma}_{-2\sigma}$, 1 bin	$\sigma_{\mathrm{ggF}}/\sigma^{SM}_{\mathrm{ggF}},^{+1\sigma}_{-1\sigma},^{+2\sigma}_{-2\sigma}$, 8 bins
36.1 fb^{-1}	$45.2^{+27.0\ +65.4}_{-16.7\ -25.5}$	$36.9^{+22.5\ +54.8}_{-13.5\ -21.1}$
49.1 fb^{-1}	$37.5^{+22.2\ +53.8}_{-13.5\ -21.1}$	$30.7^{+18.5\ +44.9}_{-11.2\ -17.4}$
147 fb^{-1}	$17.1^{+9.6\ +23.1}_{-5.8\ -9.4}$	$13.9^{+7.9\ +19.0}_{-4.7\ -7.6}$
3 000 fb^{-1}	$2.89^{+1.19\ +2.78}_{-0.76\ -1.22}$	$2.40^{+0.95\ +2.22}_{-0.61\ -0.97}$

Table 6.9: Upper limits at 95 % confidence-level on the relative di-Higgs cross section $\sigma_{\mathrm{ggF}}/\sigma^{SM}_{\mathrm{ggF}}$ are displayed for different integrated luminosities. They are either calculated while using a defined cut on $\log\left(\prod \tau_i^{\mathrm{prob}}\right)$ for the leptonic channels ("1 bin"), or while utilising the full phase space in these channels ("8 bin"). In the latter case, eight BDT distributions are used which differ in the $\log\left(\prod \tau_i^{\mathrm{prob}}\right)$ interval. For the $4h0\ell$ channel, the 1 bin method is used both times. A significant improvement of the limits can be observed when utilising the full phase space.

one single $\log\left(\prod \tau_i^{\mathrm{prob}}\right)$ threshold, the $3h1\ell$ and $2h2\ell$ channel are divided into eight distributions each with different $\log\left(\prod \tau_i^{\mathrm{prob}}\right)$ intervals. In figure A.14 in the appendix, it is visualised that several bins with considerable significance contribute to the overall fit with this method. The shape of each region is displayed in figure A.15 and A.16.

A considerable improvement of the upper limit with this "8 bin" method is visible (cf. table 6.9). The limit extracted from an assumed amount of data with an integrated luminosity of 147 fb^{-1} improves further to $\sigma_{\mathrm{ggF}}/\sigma^{SM}_{\mathrm{ggF}} = 13.9^{+7.9}_{-4.7}$. For this limit, the corresponding p value of the null-hypothesis in dependence of the signal enhancement is visualised in figure 6.21.

The observed limit for 147 fb^{-1} is significantly better than several other ongoing ATLAS analyses like $hh \to b\bar{b}WW^*$, $hh \to WW^*\gamma\gamma$ or $hh \to WW^*WW^*$. To get a comparable result to published studies [51–56], the limit is also estimated for an integrated luminosity of 36.1 fb^{-1}, which corresponds to the equivalent amount of data which was used in these publications. In the signal channel of this work an upper limit at 95 % confidence-level of $\sigma_{\mathrm{ggF}}/\sigma^{SM}_{\mathrm{ggF}} = 36.9^{+22.5}_{-13.5}$ is estimated. While the most interesting channels slightly exceed this value (cf. figure 6.22), the channels which contain W bosons deduce upper limits of 120 to 300.

Thus, the $hh \to 4\tau$ channel analysed in this work outperforms the W decay channels by far and might have the capability to contribute considerably in the combined di-Higgs analyses at the ATLAS detector. This contribution is quantified in the outlook in chapter 7.

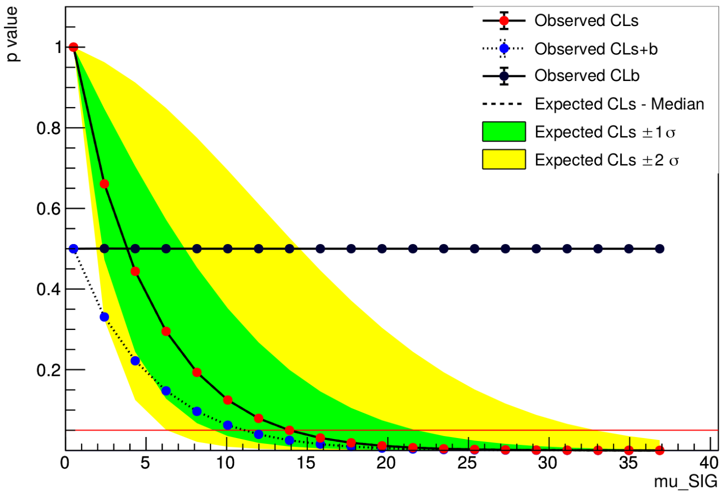


Figure 6.21: The p value of the null-hypothesis as a function of the relative signal strength for an integrated luminosity of 147 fb^{-1}. The full phase space is utilised by dividing the $\log\left(\prod \tau_i^{\text{prob}}\right)$ into eight intervals for the leptonic channels. As the data was assumed to coincide with the expectation (Asimov data), the observed limit coincides with the expected limit. An upper limit of the signal cross section in comparison to the expected SM cross section of $\mu_{\text{SIG}} = \sigma_{\text{ggF}}/\sigma_{\text{ggF}}^{\text{SM}} = 13.9^{+7.9}_{-4.7}$ is observed for this fit configuration. The corresponding plots for 36.1 fb^{-1} and 3 000 fb^{-1} are shown in figures A.17 and A.18 in the appendix.

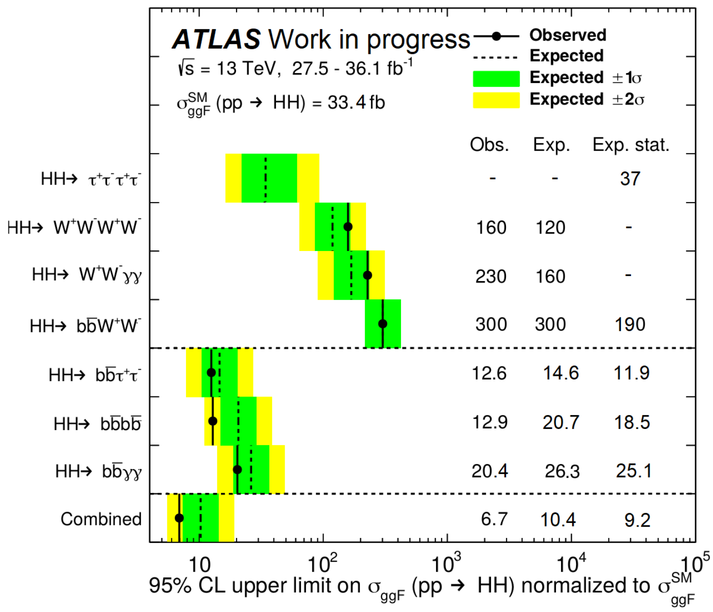


Figure 6.22: The upper limits at 95 % confidence-level on the relative di-Higgs cross section for an integrated luminosity of up to 36.1 fb^{-1} for several channels are summarised. For the combined limit, only the three latter channels divided by the dashed line are considered. The limit extracted by the $hh \to 4\tau$ channel is by far better than comparable limits of channels which contain W bosons.
This figure is based on the combination plot of the latter three channels from [116]. The values of the W channels are taken from [54–56].

CHAPTER 7

Conclusion and Outlook

In this work, the $hh \to 4\tau$ process and its detection potential is studied in detail. Its main background processes are estimated and several methods are introduced to separate them from the di-Higgs process. Moreover, its capability to contribute to the ongoing Higgs pair production studies at the ATLAS detector is investigated and an upper limit on the respective cross section is deduced with this channel.

The main limitation of the named channel is the very small amount of produced events due to the low cross section of di-Higgs events and the low branching ratio for the four-tau decay. In chapter 5, the restrictions of a successful reconstruction and identification of τ leptons in general are investigated. Together with chapter 4, it is motivated that an alternative tau identification approach is necessary to retain a sufficient amount of $hh \to 4\tau$ events for further studies. For this purpose, in section 5.3, a novel tau identification method is introduced which enables a drastic loosening of the usual tau identification requirements.

In this method, it is abstained from identifying every tau on its own. Instead, the tau-identification properties of multiple tau candidates of each event are combined and their information is used to separate multi-tau events from fake-enriched ones. A division of tau candidates into 1-prong and 3-prong ones further improves this method and very promising results are obtained. Compared to the classical tau identification approach at the ATLAS detector, a considerably higher amount of background processes is rejected for the same signal acceptance.

Further improvements of this method are imaginable. For example, BDTs could utilise the correlation of the used τ^{jetBDT} scores to other tau properties in a more comprehensive way. Alternatively, neural nets might serve as ideal methods to be trained on the specific topology of four truth tau events. Comparable methods might be applied for analyses like $hh \to b\bar{b}b\bar{b}$ to improve the multi-b-jet tagging.

In chapter 6, possible heavy-particle processes which mimic the di-Higgs topology are summarised. To predict the influence of pure QCD events, a data-driven fake estimation for four objects is developed. This estimation is based on the object's charges and together with the heavy-particle processes, their prediction show a satisfactory agreement with recorded data. The challenge to align this charge-based fake-estimation with the used tau pair assignment is solved such that a reliable prediction of background processes is assumed.

Based on the new tau identification method of chapter 5, signal regions with highly enriched $hh \to 4\tau$ events compared to background are designed. For this purpose, three decay channels are differentiated by the expected tau decays. Namely, in the $4h0\ell$, $3h1\ell$ and $2h2\ell$ channel events with either four, three or two hadronically decaying taus are considered, which corresponds to 87 % of all $hh \to 4\tau$ decays.

The leptonically decaying taus are reconstructed as light leptons in the ATLAS detector and serve as separation criteria of the named channels. In a first approach, the event selection is designed by investigations of significance scans. Possible correlations between observables are taken into account as several two-dimensional significance scans are performed for different preselections. In a second approach, the achieved separation of signal and background is improved by BDTs which are trained on several quantities with expected separation power.

A combined significance of 0.89 in an optimistic scenario or 0.66 in a very conservative scenario is deduced. The highest contribution comes from the $4h0\ell$ channel in which a significance nearly two times as high as for the $3h1\ell$ channel is expected. In the $2h2\ell$ channel, the achieved significance is roughly divided by another factor of two. In all cases, only statistical uncertainties were considered. However, especially in the $4h0\ell$ channel, the amount of expected background events in the signal region is very low. Possible higher uncertainties on these values should not lead to much lower significances.

The significance values have to be compared to an overall expected significance of the di-Higgs detection at the ATLAS detector of 3.5 without systematic uncertainties or 3.0 with systematic uncertainties [49]. These values are combined predictions of the $hh \to b\bar{b}\tau\tau$, $hh \to b\bar{b}\gamma\gamma$ and $hh \to b\bar{b}b\bar{b}$ channels which are estimated to achieve single significances of 2.5, 2.1 and 1.4 respectively. If systematic uncertainties are included, the values drop to 2.1, 2.0 and 0.6.

Due to the low event expectations, the $hh \to 4\tau$ channel is expected to be dominated by statistical uncertainties at the end of the HL-LHC program and in an optimistic scenario, significance values comparable to the 0.6 of $hh \to b\bar{b}b\bar{b}$ after the consideration of systematic uncertainties might be realistic. The more optimistic value of 0.89 for the signal channel of this work pushes the overall significance up to either 3.61 or 3.13 which corresponds to an improvement of roughly 3 % to 4 %. As the Higgs self-coupling might be altered considerably by physics beyond the SM, this contribution should not be underestimated. However, the four-tau channel is not expected to contribute as much as the most sensible channels.

In section 6.6, an upper limit at 95 % confidence level for the di-Higgs cross section is deduced. For an integrated luminosity of up to 36.1 fb^{-1}, several studies for different channels are published [51–56] such that the respective upper limits can be compared. For the $hh \to 4\tau$ channel, an upper limit of $36.9^{+22.5}_{-13.5}$ is calculated. Again, this value cannot compete with the limits of the previously named three channels in which expected limits of 12 to 25 are predicted for statistical errors only. In decay channels which include W bosons, however, upper limits of only 77 to 240 are expected. Hence, they are outperformed by far by the $hh \to 4\tau$ process.

The promising results are built on several assumptions which have to be tested in future. First of all, a suitable trigger has to be selected which does not discard too many signal events. While for the leptonic channels, a light lepton trigger might enable high efficiencies, a trigger for the $4h0\ell$ channel remains challenging. This problem intensifies if only tau candidates with low transverse momenta populate an event. In section 6.4, the sum of all transverse lepton momenta is singled out as a very promising quantity to separate signal from background and in all three channels, at least $\sum p_T(\text{lep}_{1\text{-}4}) > 200\,\text{GeV}$ is demanded for the signal region. Hence, at least one or two taus in this region might exceed a hypothetical p_T threshold of future tau or ditau triggers, anyway.

Moreover, in this work the data recorded in 2017 was extrapolated to 3 000 fb^{-1} without any further assumptions. At the HL-LHC, however, much more pile-up is expected [76]. Hence, on the one hand, the p_T thresholds for tau candidates have to be reconsidered with rigorous effects on the tau acceptance as outlined in section 5.1. Especially the studied use of very loose taus as defined in this thesis (i.a. $p_T > 15\,\text{GeV}$) is questionable in this context and needs to be verified once first data with comparable pile-up is available. Also the effect of such a pile-up increase on pure QCD fake-taus can only reliably be studied with appropriate data. Nevertheless, the data-driven fake estimation designed in this thesis is expected to be robust, especially when alternative approaches and improvements as outlined in section 6.2 are applied.

On the other hand, the ATLAS detector will cover a solid angle up to $|\eta| = 4.0$ for the HL-LHC upgrade. Moreover, more sophisticated methods might improve the use of tau candidates in the $|\eta|$ crack region. These two effects might equalise possible acceptance loss due to different p_{T} thresholds. Additionally, improvements on b tagging or other identification methods were not considered at all. The signal acceptance as well as possible more exclusive background rejection might benefit from such future developments.

In any case, the presented work provides several possibilities to be further improved and completed. First of all, the dominant background processes should be divided into particular control regions in which they are probed against data. Especially the indicated overestimation of the MC simulations in the $3h1\ell$ channel might result in a slight downscaling of major background processes and higher significances in this channel could be entailed. The enormous discrepancy between data and MC predictions in the $2h2\ell$ channel has to be understood and almost certainly the prediction of this channel will be altered hereby. As the contribution of this channel to the estimated significance of the di-Higgs process is very small at the moment, the potential improvement is very high compared to the setback.

Last but not least, systematic uncertainties have to be included which will decrease the estimated significance and upper limit. Especially due to the novel tau identification approach, the systematic uncertainties due to τ leptons will be more complex.

In summary, it is shown that the $hh \rightarrow 4\tau$ process indeed provides the capability to considerably participate in the ongoing Higgs pair production measurement. As the process is strongly limited statistically, its contribution to the combined upper limit is expected to rise with more recorded data. Thus, further investigations of this channel are recommendable.

APPENDIX A

Additional Tables and Figures

A.1 Details of the Used MC Samples

Sample ID	Generator	Name
342284	Pythia8EvtGen	WH125 inc
342285	Pythia8EvtGen	ZH125 inc
344235	PowhegPythia8EvtGen	VBFH125 ZZ4lep notau
345060	PowhegPythia8EvtGen	ggH125 ZZ4l
345066	PowhegPythia8EvtGen	ggZH125 ZZ4lepZinc
345123	PowhegPythia8EvtGen	ggH125 tautau
345211	PowhegPythia8EvtGen	WmH125J Winc tautau
345212	PowhegPythia8EvtGen	WpH125J Winc tautau
345217	PowhegPythia8EvtGen	ZH125J Zinc tautau
345875	PowhegPythia8EvtGen	ttH125 dilep
345948	PowhegPythia8EvtGen	VBFH125 WWlvlv
346193	PowhegPythia8EvtGen	VBFH125 tautauh
345911	aMcAtNloHerwig7EvtGen	hh 4tau

Table A.1: The used samples which are summarised as single Higgs processes (upper block) and the $hhh \to 4\tau$ sample (lower block).

Sample ID	Generator	Name
410470	PhPy8EG	ttbar nonallhad
410471	PhPy8EG	ttbar allhad
410472	PhPy8EG	ttbar dil

Table A.2: The used samples which are summarised as $t\bar{t}$ production.

Sample ID	Generator	Name
410644	PowhegPythia8EvtGen	singletop schan lept top
410645	PowhegPythia8EvtGen	singletop schan lept antitop
410646	PowhegPythia8EvtGen	Wt DR inclusive top
410647	PowhegPythia8EvtGen	Wt DR inclusive antitop
410654	PowhegPythia8EvtGen	Wt DS inclusive top
410655	PowhegPythia8EvtGen	Wt DS inclusive antitop

Table A.3: The used samples which are summarised as single top processes.

Sample ID	Generator	Name
364128	Sherpa 221	Ztautau p_T range: 0 70 CVetoBVeto
364129	Sherpa 221	Ztautau p_T range: 0 70 CFilterBVeto
364130	Sherpa 221	Ztautau p_T range: 0 70 BFilter
364131	Sherpa 221	Ztautau p_T range: 70 140 CVetoBVeto
364132	Sherpa 221	Ztautau p_T range: 70 140 CFilterBVeto
364133	Sherpa 221	Ztautau p_T range: 70 140 BFilter
364134	Sherpa 221	Ztautau p_T range: 140 280 CVetoBVeto
364135	Sherpa 221	Ztautau p_T range: 140 280 CFilterBVeto
364136	Sherpa 221	Ztautau p_T range: 140 280 BFilter
364137	Sherpa 221	Ztautau p_T range: 280 500 CVetoBVeto
364138	Sherpa 221	Ztautau p_T range: 280 500 CFilterBVeto
364139	Sherpa 221	Ztautau p_T range: 280 500 BFilter
364140	Sherpa 221	Ztautau p_T range: 500 1000
364141	Sherpa 221	Ztautau p_T range: 1000 E CMS

Table A.4: The used samples which are summarised in $Z \to \tau\tau$ processes. They are sliced with respect to the transversal momentum p_T. Moreover, several flavour filters are applied.

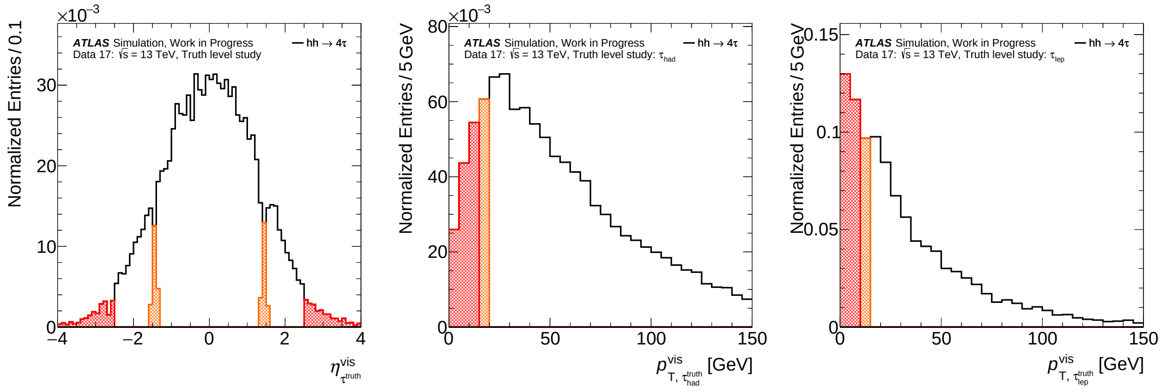
Sample ID	Generator	Name
308092	Sherpa 221	Zee2jets TChannel
308093	Sherpa 221	Zmm2jets TChannel
308094	Sherpa 221	Ztautau2jets TChannel
308095	Sherpa 221	Znunu2jets TChannel
308096	Sherpa 221	Wenu2jets TChannel
308097	Sherpa 221	Wmunu2jets TChannel
308098	Sherpa 221	Wtaunu2jets TChannel
364100	Sherpa 221	Zmumu p_T range: 0 70 CVetoBVeto
364101	Sherpa 221	Zmumu p_T range: 0 70 CFilterBVeto
364102	Sherpa 221	Zmumu p_T range: 0 70 BFilter
364103	Sherpa 221	Zmumu p_T range: 70 140 CVetoBVeto
364104	Sherpa 221	Zmumu p_T range: 70 140 CFilterBVeto
364105	Sherpa 221	Zmumu p_T range: 70 140 BFilter
364106	Sherpa 221	Zmumu p_T range: 140 280 CVetoBVeto
364107	Sherpa 221	Zmumu p_T range: 140 280 CFilterBVeto
364108	Sherpa 221	Zmumu p_T range: 140 280 BFilter
364109	Sherpa 221	Zmumu p_T range: 280 500 CVetoBVeto
364110	Sherpa 221	Zmumu p_T range: 280 500 CFilterBVeto
364111	Sherpa 221	Zmumu p_T range: 280 500 BFilter
364112	Sherpa 221	Zmumu p_T range: 500 1000
364113	Sherpa 221	Zmumu p_T range: 1000 E CMS
364114	Sherpa 221	Zee p_T range: 0 70 CVetoBVeto
364115	Sherpa 221	Zee p_T range: 0 70 CFilterBVeto
364116	Sherpa 221	Zee p_T range: 0 70 BFilter
364117	Sherpa 221	Zee p_T range: 70 140 CVetoBVeto
364118	Sherpa 221	Zee p_T range: 70 140 CFilterBVeto
364119	Sherpa 221	Zee p_T range: 70 140 BFilter
364120	Sherpa 221	Zee p_T range: 140 280 CVetoBVeto
364121	Sherpa 221	Zee p_T range: 140 280 CFilterBVeto
364122	Sherpa 221	Zee p_T range: 140 280 BFilter
364123	Sherpa 221	Zee p_T range: 280 500 CVetoBVeto
364124	Sherpa 221	Zee p_T range: 280 500 CFilterBVeto
364125	Sherpa 221	Zee p_T range: 280 500 BFilter
364126	Sherpa 221	Zee p_T range: 500 1000
364127	Sherpa 221	Zee p_T range: 1000 E CMS

Table A.5: The used samples which are summarised as V + jets processes. In the upper block, *t*-channel processes are summarised. The latter two blocks include all $Z \rightarrow \mu\mu$ and $Z \rightarrow ee$ processes. They are sliced with respect to the transversal momentum p_T. Moreover, several flavour filters are applied.

Sample ID	Generator	Name
363355	Sherpa 221	ZqqZvv
363356	Sherpa 221	ZqqZll
363357	Sherpa 221	WqqZvv
363358	Sherpa 221	WqqZll
363359	Sherpa 221	WpqqWmlv
363360	Sherpa 221	WplvWmqq
363489	Sherpa 221	WlvZqq
364250	Sherpa 222	llll
364253	Sherpa 222	lllv
364254	Sherpa 222	llvv
364255	Sherpa 222	lvvv
364283	Sherpa 222	lllljj
364284	Sherpa 222	lllvjj
364285	Sherpa 222	llvvjj
364288	Sherpa 222	llll lowMllPtComplement
364289	Sherpa 222	lllv lowMllPtComplement
364290	Sherpa 222	llvv lowMllPtComplement
407311	Sherpa 221	6l0v
407312	Sherpa 221	5l1v
407313	Sherpa 221	4l2v
407314	Sherpa 221	3l3v
407315	Sherpa 221	2l4v

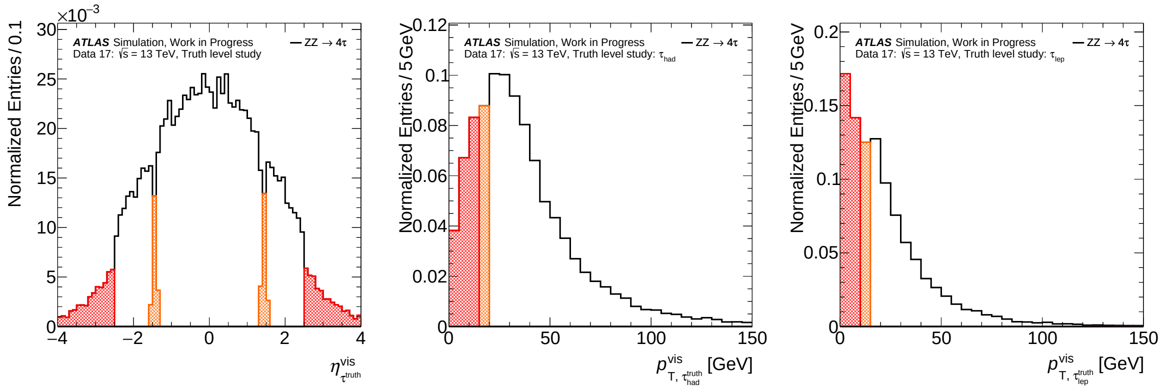
Table A.6: The used samples which are summarised in multiboson processes. They consist of diboson (upper two block) and triboson (lower block) processes. The llll sample contains all SM processes which have exactly four leptons as a final state. As an example, also the process $ZZ \to 4\tau$ is covered by this process. Due to its limited statistic, no $hh \to 4\tau$ processes are contained. Analogously, lllv contains all SM processes which result in exactly three leptons and one neutrino etc.

A.2 Tau Acceptance Plots of the DAOD Files



(a) (b) (c)

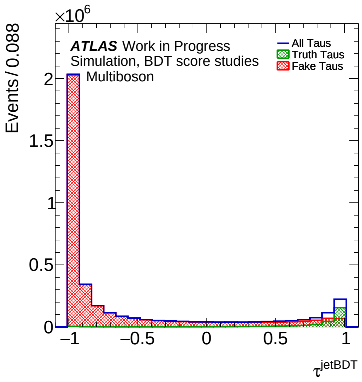
Figure A.1: Fraction of reconstructable truth taus in the $hh \to 4\tau$ channel with respect to their pseudorapidity (a) or their transversal momenta. The latter one is divided into hadronically (b) and leptonically (c) decaying taus. Taus in the red region are not reconstructable or impossible to identify. Taus in the orange region are very hard to successfully identify. However, they might be accessible in general. Further explanations are given in the text. In contrast to figure 5.1, the DAOD sample of the process was taken. This sample already suffers from some preselection cuts on the tau and lepton candidates and thus does not depict the full reality. Due to that, the "forbidden" $|\eta|$ and p_T regions are slightly downscaled.



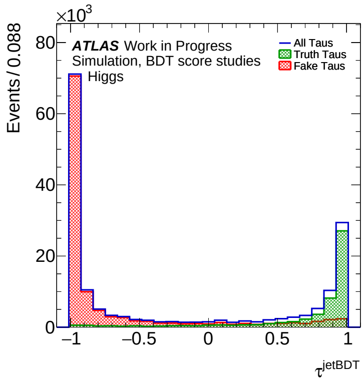
(a) (b) (c)

Figure A.2: Fraction of reconstructable truth taus in the $ZZ \to 4\tau$ channel with respect to their pseudorapidity (a) or their transversal momenta. The latter one is divided into hadronically (b) and leptonically (c) decaying taus. Taus in the red region are not reconstructable or impossible to identify. Taus in the orange region are very hard to successfully identify. However, they might be accessible in general. Further explanations are given in the text. Analogously to figure A.1, the DAOD sample of the process was taken. This sample already suffers from some preselection cuts on the tau and lepton candidates and thus does not depict the full reality. Due to that, the "forbidden" $|\eta|$ and p_T regions are slightly downscaled.

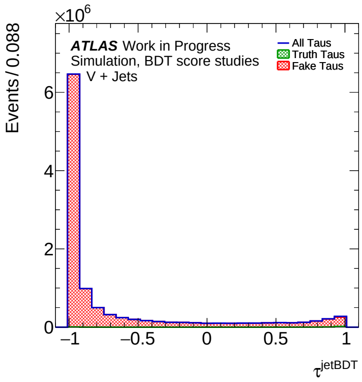
A.3 Additional Plots Concerning the Introduction of Tau Likelihood and Probability



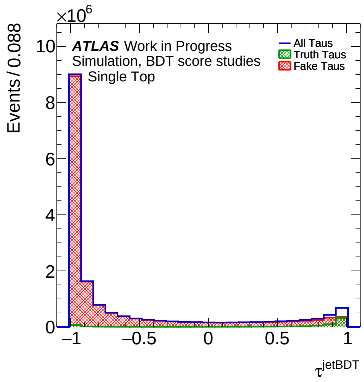
(a) Multiboson



(b) Higgs

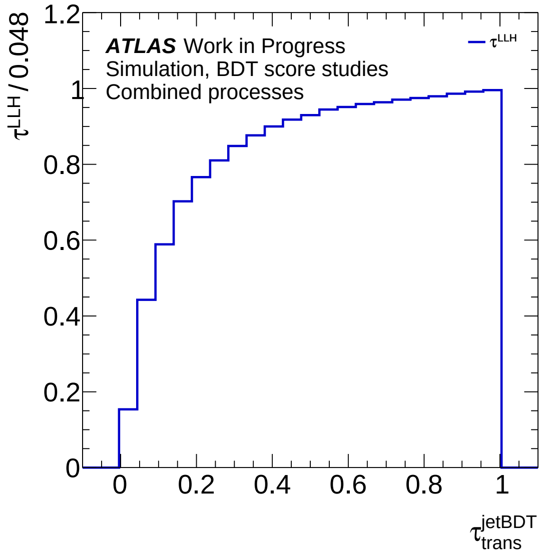


(c) V + Jets

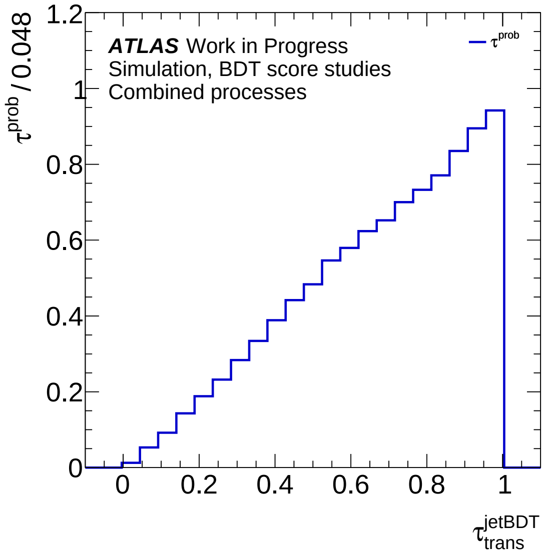


(d) Single Top

Figure A.3: The distribution of truth and fake taus for the background production modes which are not presented in figure 5.5.

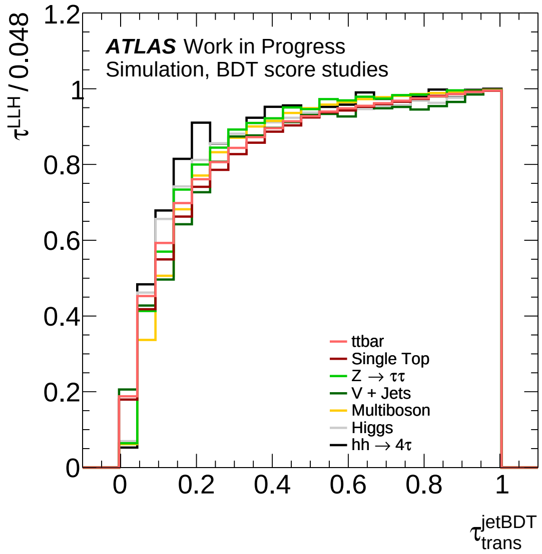


(a) Likelihood

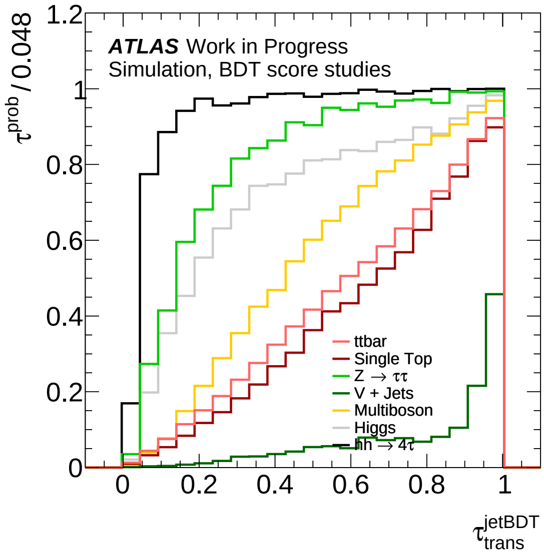


(b) Probability

Figure A.4: The mapped likelihood value τ^{LLH} (a) and the mapped probability value τ^{prob} (b) with respect to the transformed tau jet BDT value are shown. As the overall amount of fake taus is much higher than the one of truth taus (cf. figure 5.5(c)), the probability value is squeezed to lower values.



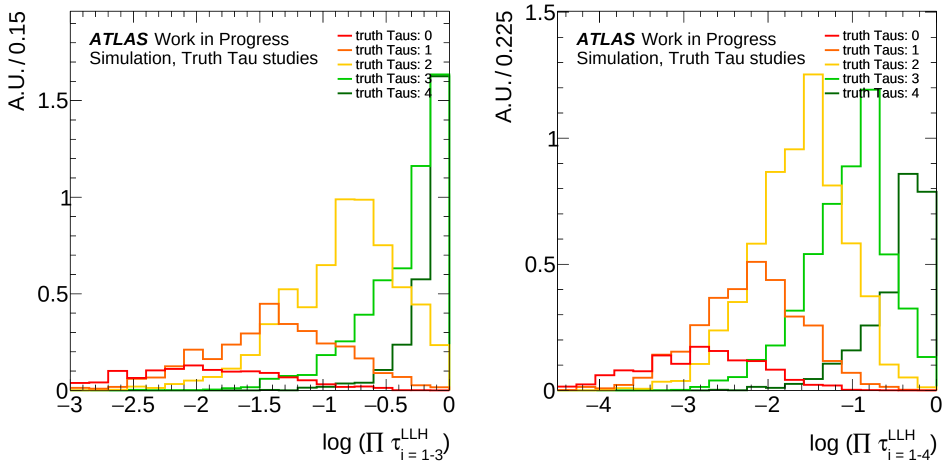
(a) Likelihood



(b) Probability

Figure A.5: The mapped likelihood value τ^{LLH} (a) and the mapped probability value τ^{prob} (b) with respect to the transformed $\tau_{\mathrm{trans}}^{\mathrm{jetBDT}}$ score are shown for different processes. Similar to figure 5.7, the likelihood is roughly independent of the process (within statistical fluctuations due to the limited statistic in some processes), while the probability significantly differs.

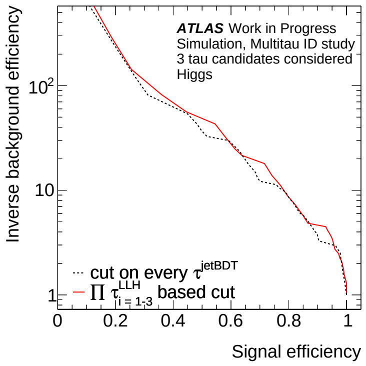
A.4 Additional Plots to Evaluate the Performance of the Multiplied Likelihood/Probability Method



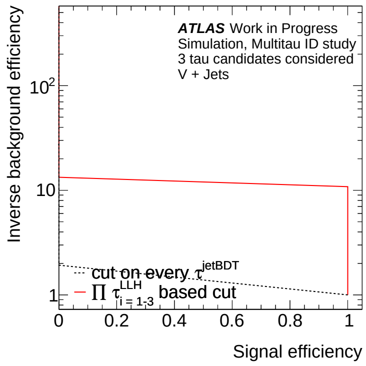
(a) τ^{LLH} product of three tau candidates

(b) τ^{LLH} product of four tau candidates

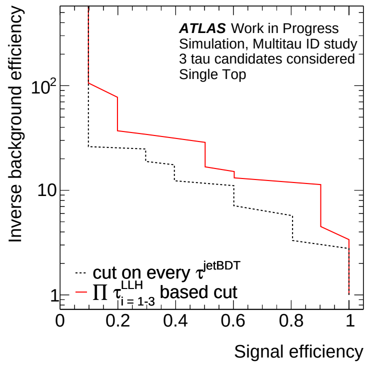
Figure A.6: The separation power of the logarithm of the τ^{LLH} product is shown for the Sherpa $\ell\ell\ell\ell$ sample. Clearly, events with a high amount of truth taus (light and dark green) peak at higher values while events which are mainly fake dominated peak at the left. Compared to the corresponding distributions with τ^{prob} (figure 5.10), the distributions are squeezed to low values.



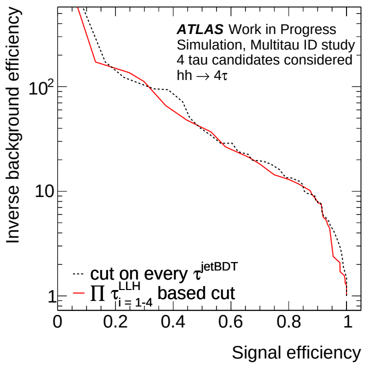
(a) Single Higgs, 3 tau



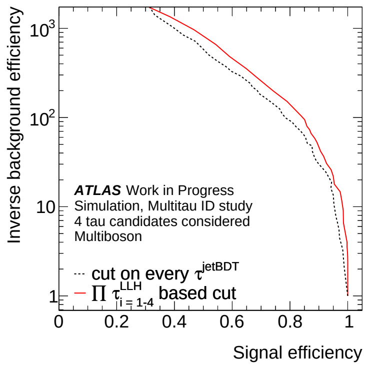
(b) V + Jets, 3 tau



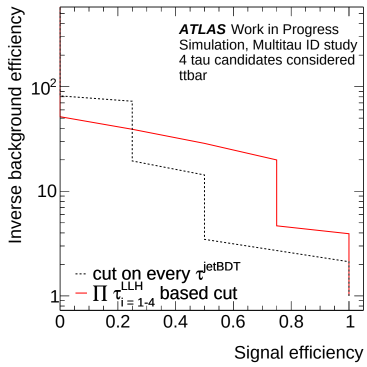
(c) Single Top, 3 tau



(d) $hh \to 4\tau$, 4 tau

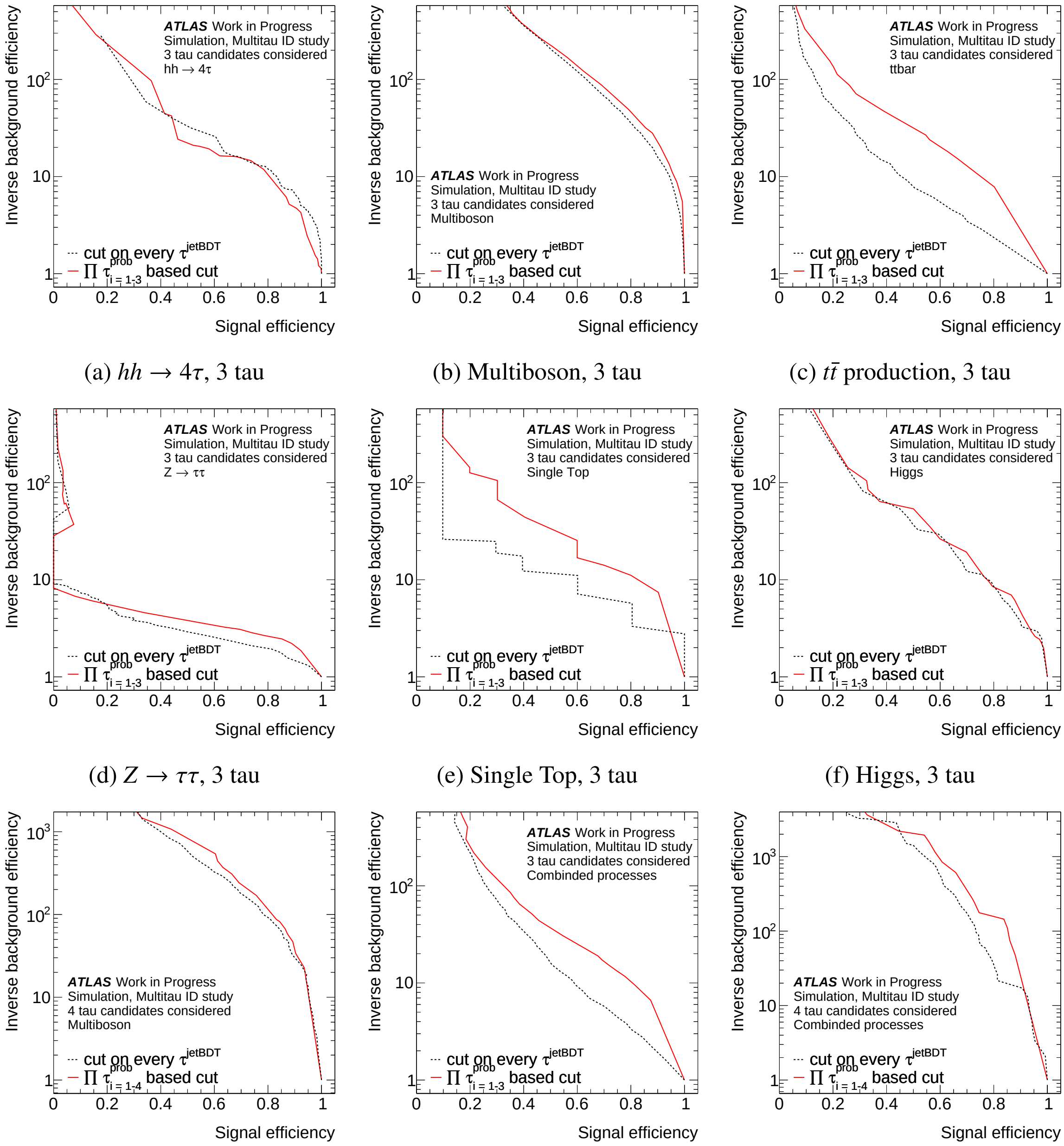


(e) Multiboson, 4 tau



(f) $t\bar{t}$ production, 4 tau

Figure A.7: The background rejection of two separation methods with respect to the signal efficiency is shown for several processes (ROC curves). In the classical method, events are rejected if all three/four tau candidates with highest τ^{jetBDT} value do not fulfil a certain τ^{jetBDT} threshold which runs from -1 to $+1$. In the new method, the τ^{LLH} product of the first three/four tau candidates is demanded to exceed a running threshold between 0 and 1. Signal events are all events of the specific process in which indeed the considered tau candidates are truth matched taus. Correspondingly, all events are labelled as background if one or more of them are fake. In several processes (e.g. the V + Jets process (b) or single top production (c)) the amount of three truth matched taus is very limited. Hence, the ROC curve suffers from many spikes. When considering four taus, the limited amount of such a high number of truth matched taus is even observable in the $t\bar{t}$ process (f). More meaningful is the consideration of the combination of all processes in the four-tau case as presented in figure 5.11(f).

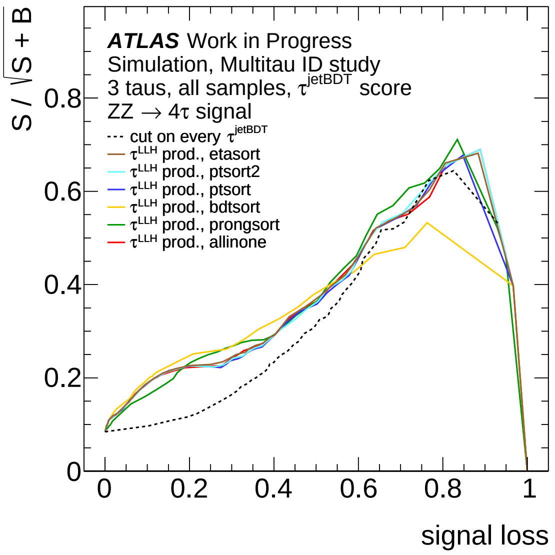


(a) $hh \rightarrow 4\tau$, 3 tau (b) Multiboson, 3 tau (c) $t\bar{t}$ production, 3 tau

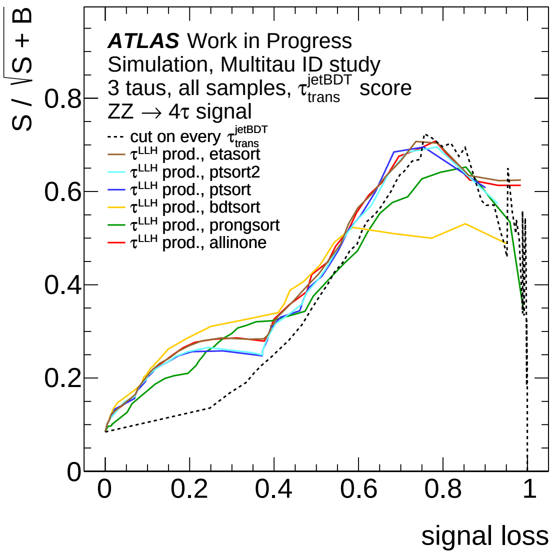
(d) $Z \rightarrow \tau\tau$, 3 tau (e) Single Top, 3 tau (f) Higgs, 3 tau

(g) Multiboson, 4 tau (h) Sum of all samples, 3 taus (i) Sum of all samples, 4 taus

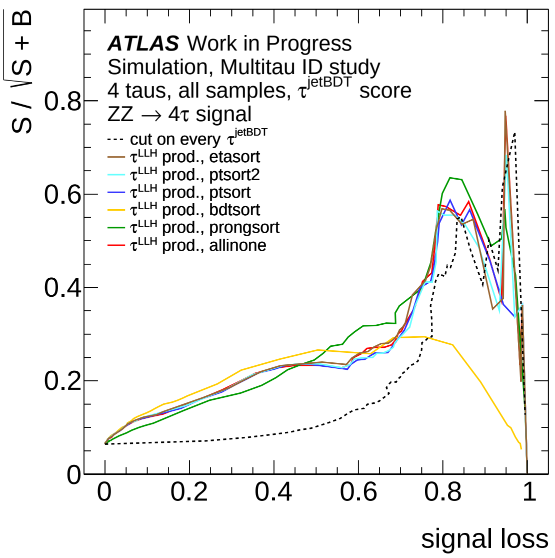
Figure A.8: ROC curves of the multiplied-likelihood method with τ^{prob} compared to single BDT based cuts for different processes. Analogous figures for τ^{LLH} with more explanations are visualised in figure 5.11 and A.7.



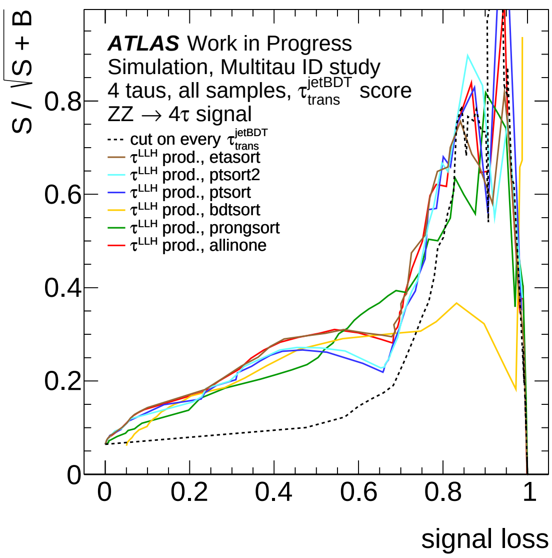
(a) $\prod \tau^{\mathrm{LLH}}_{i=1-3}$ created with the τ^{jetBDT} score



(b) $\prod \tau^{\mathrm{LLH}}_{i=1-3}$ created with the $\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$ score

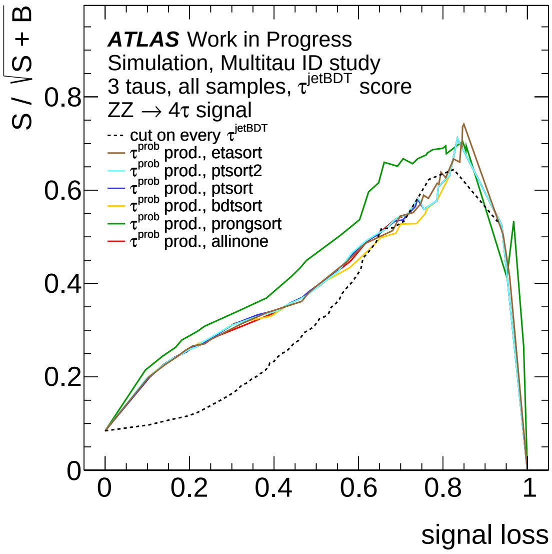


(c) $\prod \tau^{\mathrm{LLH}}_{i=1-4}$ created with the τ^{jetBDT} score

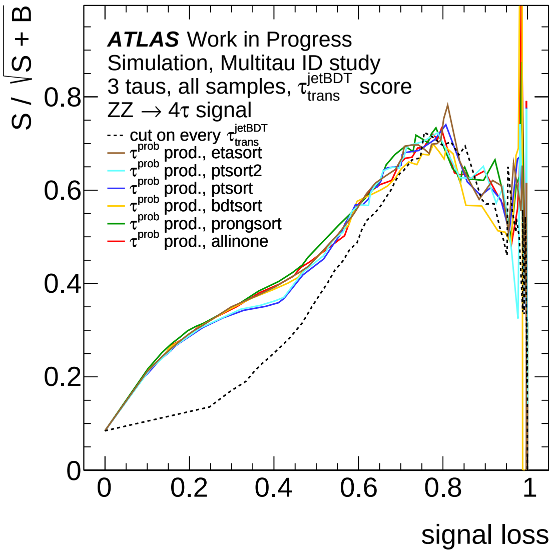


(d) $\prod \tau^{\mathrm{LLH}}_{i=1-4}$ created with the $\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$ score

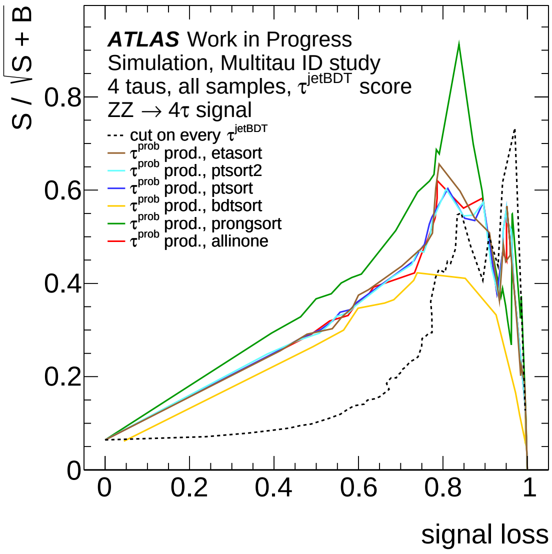
Figure A.9: The significances for the $ZZ \rightarrow 4\tau$ signal while cutting either on the τ^{LLH} product of the first three tau candidates (a) and (b), or on four of them (c), (d) with respect to the signal loss are presented. Here, every product is first evaluated when the τ^{LLH} values were gained with the normal τ^{jetBDT} score and second when the transformed score $\tau^{\mathrm{jetBDT}}_{\mathrm{trans}}$ was used. Different sorting methods were further used to map the BDT scores to the likelihood (cf. section 5.3.2). Compared to cutting on $\tau^{\mathrm{jetBDT}}_{\mathrm{(trans)}}$ (like in the classical tau identification approach, visualised by the black dashed line) a clear significance gain in most sorting methods is observed. However, no sorting method really singles out to be the best one.



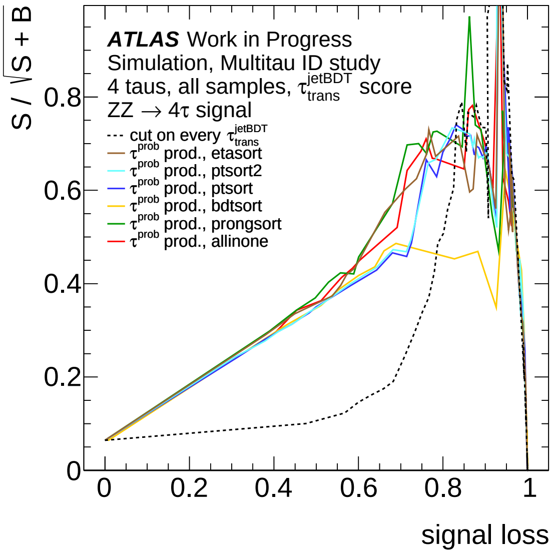
(a) $\prod \tau^{\text{prob}}_{i=1-3}$ created with the τ^{jetBDT} score



(b) $\prod \tau^{\text{prob}}_{i=1-3}$ created with the $\tau^{\text{jetBDT}}_{\text{trans}}$ score

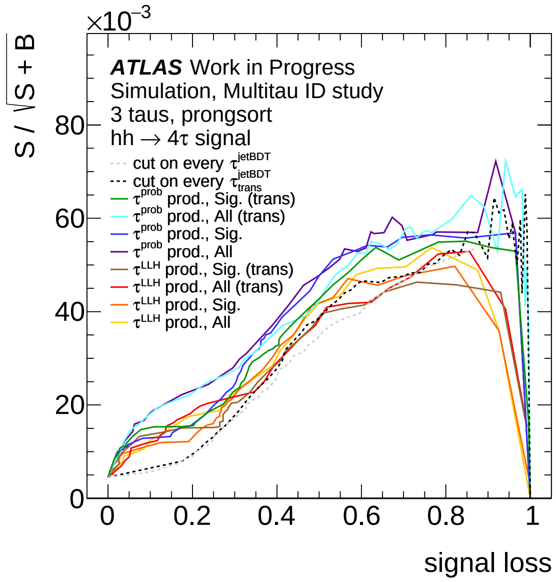


(c) $\prod \tau^{\text{prob}}_{i=1-4}$ created with the τ^{jetBDT} score

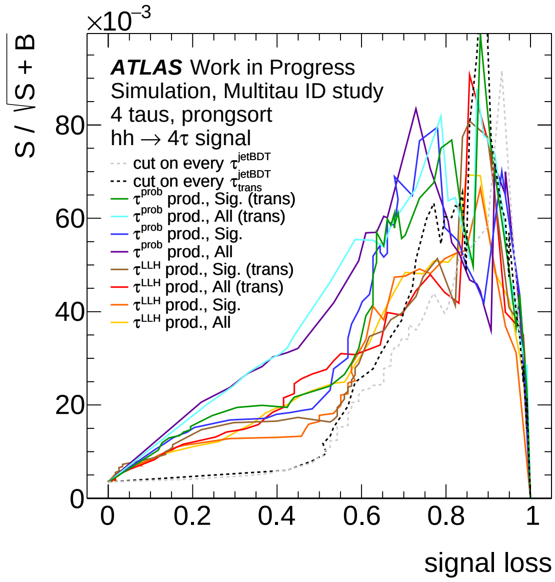


(d) $\prod \tau^{\text{prob}}_{i=1-4}$ created with the $\tau^{\text{jetBDT}}_{\text{trans}}$ score

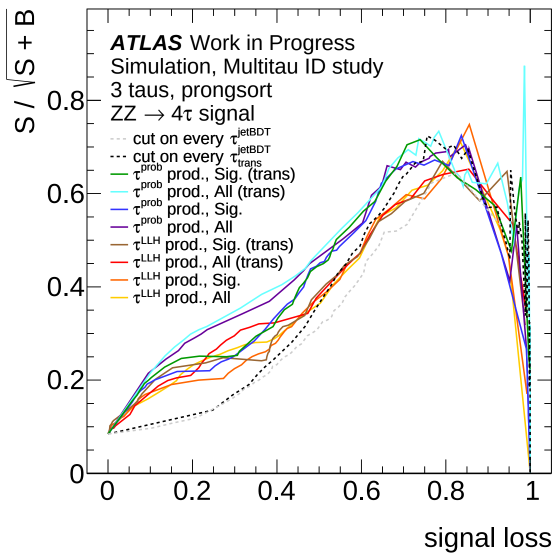
Figure A.10: The significances for the $ZZ \to 4\tau$ signal while cutting either on the τ^{prob} product of the first three tau candidates (a) and (b), or on four of them (c), (d) with respect to the signal loss are presented. Hereby, every product is first evaluated when the τ^{prob} values were gained with the normal τ^{jetBDT} score and second when the transformed score $\tau^{\text{jetBDT}}_{\text{trans}}$ was used. Different sorting methods were further used to map the BDT scores to the probabilities (cf. section 5.3.2). Compared to cutting on $\tau^{\text{jetBDT}}_{\text{(trans)}}$ (like in the classical tau identification approach, visualised by the black dashed line) a massive significance gain in nearly all sorting methods is observed. In all cases, the `prongsort` sorting method in green either clearly or slightly outperforms the other sorting methods.



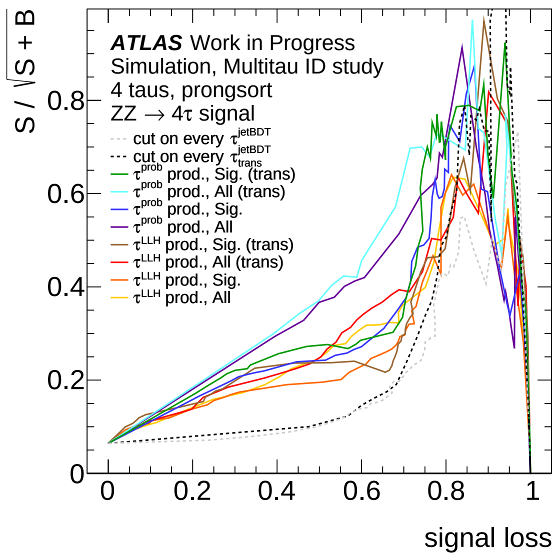
(a) $hh \to 4\tau$ signal, 3 tau candidates



(b) $hh \to 4\tau$ signal, 4 tau candidates



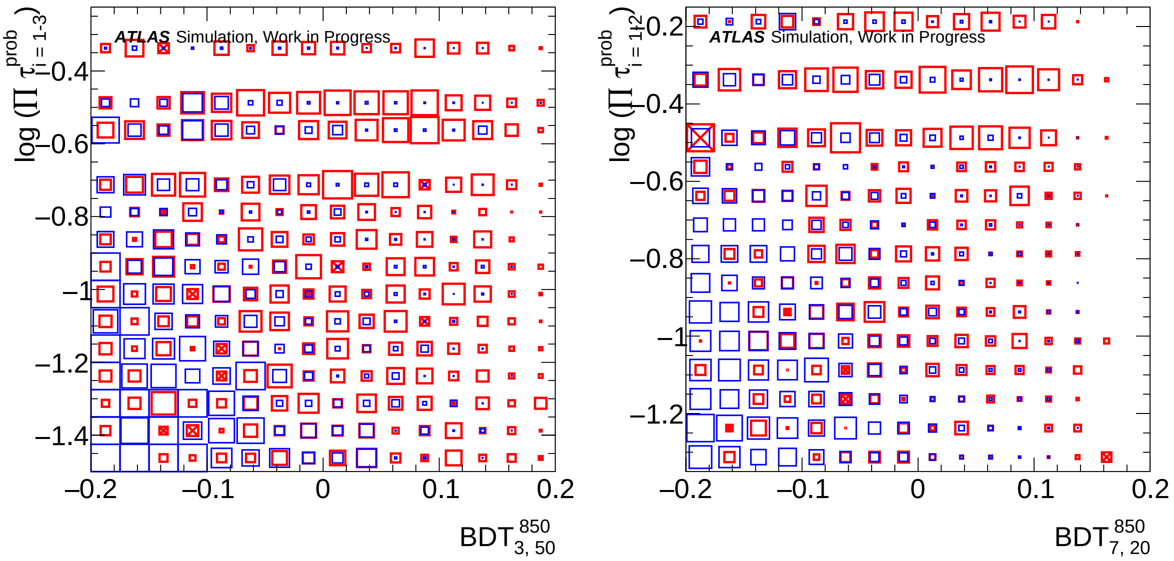
(c) $ZZ \to 4\tau$ signal, 3 tau candidates



(d) $ZZ \to 4\tau$ signal, 4 tau candidates

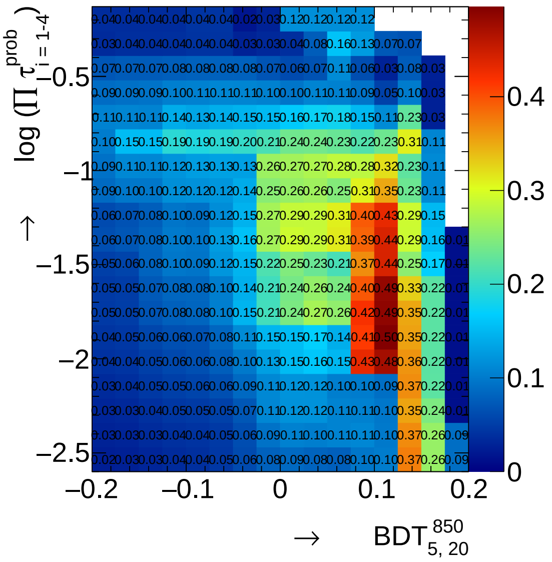
Figure A.11: The separation power of τ^{LLH} and τ^{prob} is compared. Both values are either produced with the normal τ^{jetBDT} score or the transformed $\tau^{\text{jetBDT}}_{\text{trans}}$ score. Moreover, either all processes considered in this work are used to build the truth and fake distributions or only signal-like processes are considered. In the $hh \to 4\tau$ sample as well as in the $ZZ \to 4\tau$ process, the τ^{prob} value by far outperforms the τ^{LLH} value. Moreover, in most cases the significance is reduced if the values are created only with the information of signal-like samples. Thus, they are discarded in figure 5.14 and are not used further.

A.5 Additional Plots of the Significance Scans to Define a Signal Region

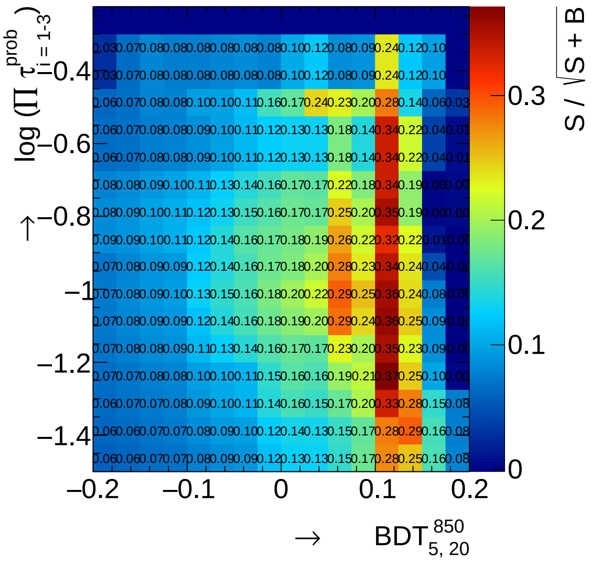


(a) $3h1\ell$ channel (b) $2h2\ell$ channel

Figure A.12: The two-dimensional distribution of the trained BDTs against the logarithm of $\prod \tau_i^{\text{prob}}$ is presented for leptonic channels. The normalised signal in red accumulates at the upper right corner while background events in blue show the opposite behaviour.



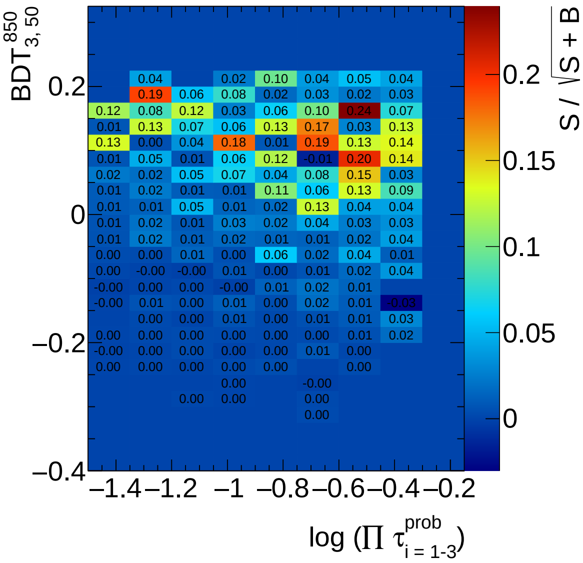
(a) $4h0\ell$ channel



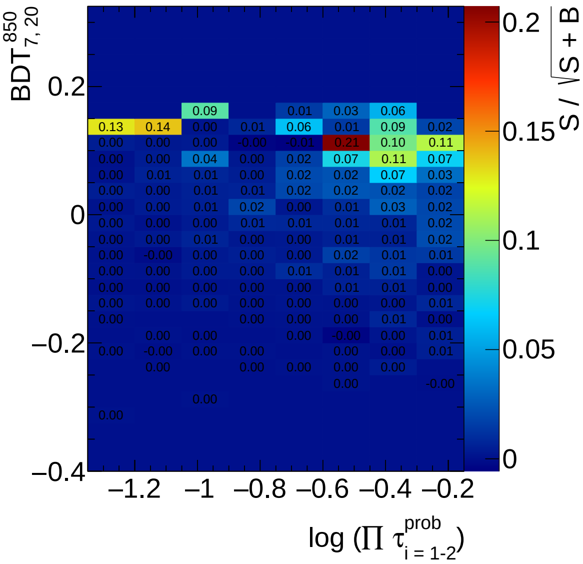
(b) $3h1\ell$ channel

Figure A.13: Alternative two-dimensional significance scans of the trained BDTs against the probability-product for the $4h0\ell$ and $3h1\ell$ channel are shown. Each bin of the significance plots represents a *cut from below* on the corresponding value of the x or y axis. Thus, only events in the upper right corner of each bin are used to calculate the significance $S/\sqrt{S+B}$. Compared to figure 6.19, higher significances are achieved. However, the distribution is very discontinuous and abruptly rises after a certain BDT threshold. As this might be caused by statistical limitation of the MC simulations close to the signal region, these plots are not used for the shape fit analysis in section 6.6.

A.6 Additional Plots of the Shape Fit Analysis

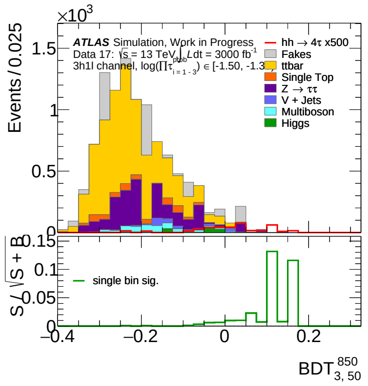


(a) $3h1\ell$ channel

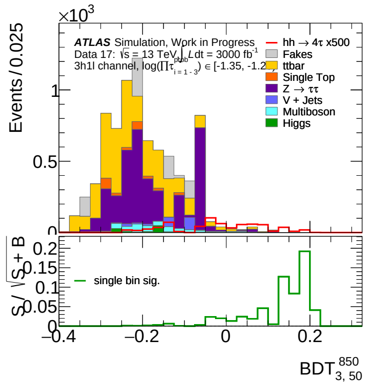


(b) $2h2\ell$ channel

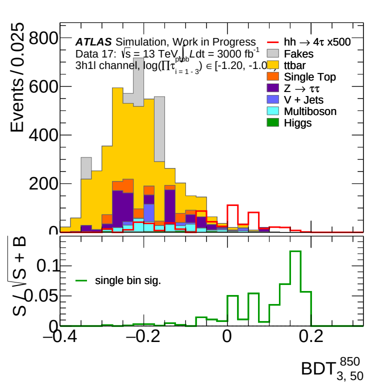
Figure A.14: Two-dimensional significance scans of the trained BDTs against the probability-product for the leptonic channel are shown. In each bin, only the amount of signal and background of this bin is considered. For the $3h1\ell$ channel (a), several $\log\left(\prod \tau_i^{\text{prob}}\right)$ intervals show a continuous shape of significances and several bins with high $S/\sqrt{S+B}$ are observed. For the $2h2\ell$ channel (b), however, this is only true for two or maybe three $\log\left(\prod \tau_i^{\text{prob}}\right)$ intervals. A broader binning might be advisable.



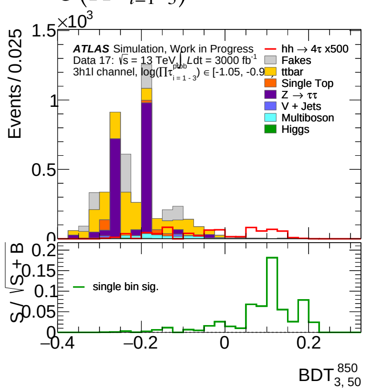
(a) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-1.50, -1.35)$



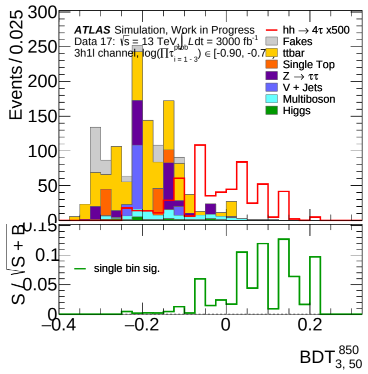
(b) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-1.35, -1.20)$



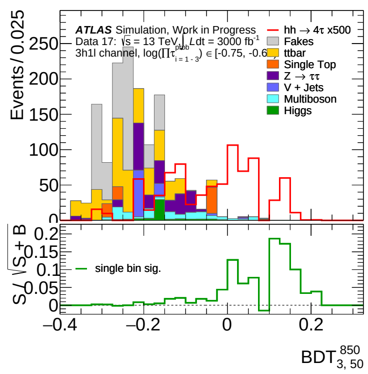
(c) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-1.20, -1.05)$



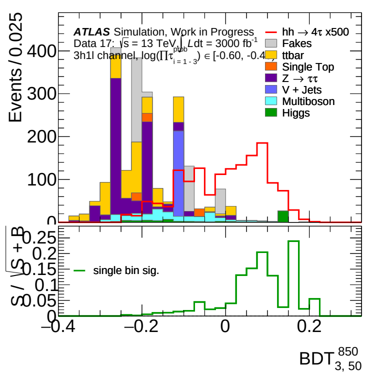
(d) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-1.05, -0.90)$



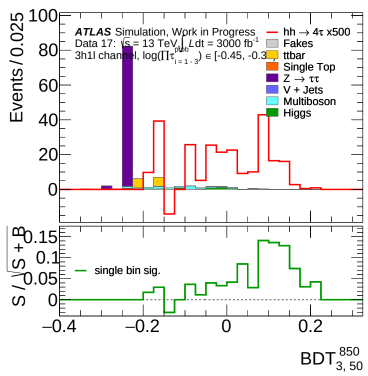
(e) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-0.90, -0.75)$



(f) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-0.75, -0.60)$

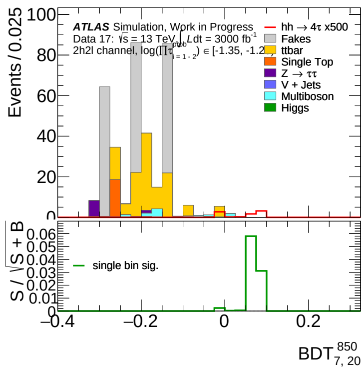


(g) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-0.60, -0.45)$

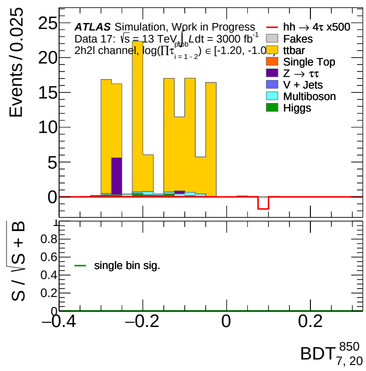


(h) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-0.45, -0.30)$

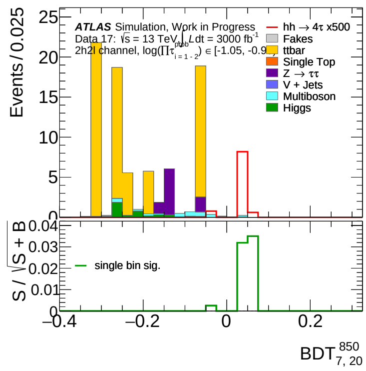
Figure A.15: Signal and background distributions for each $\log\left(\prod \tau^{\text{prob}}_{i}\right)$ intervals for the $3h1\ell$ channel. Each plot corresponds to a row in figure A.14(a).



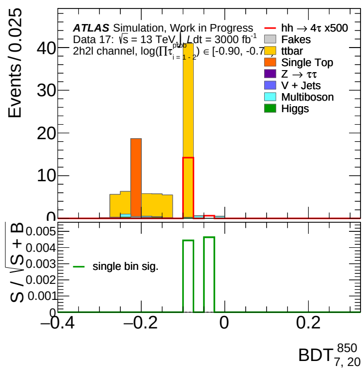
(a) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-1.35, -1.20)$



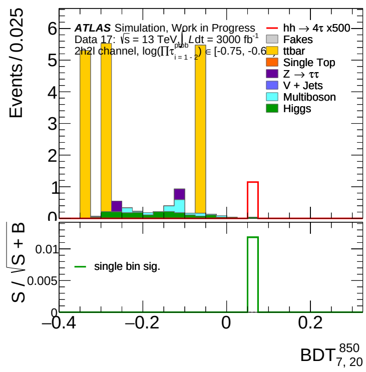
(b) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-1.20, -1.05)$



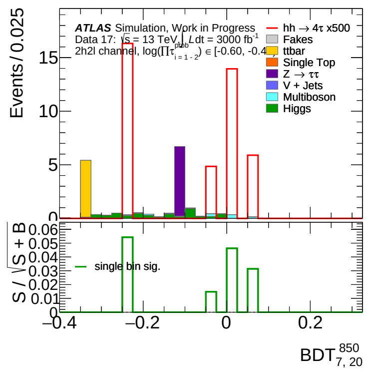
(c) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-1.05, -0.90)$



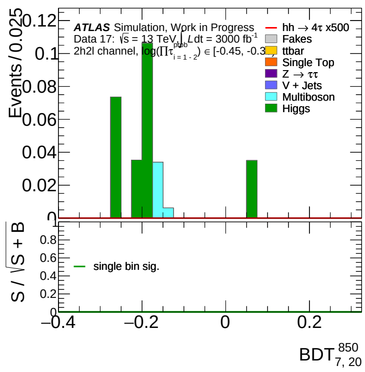
(d) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-0.90, -0.75)$



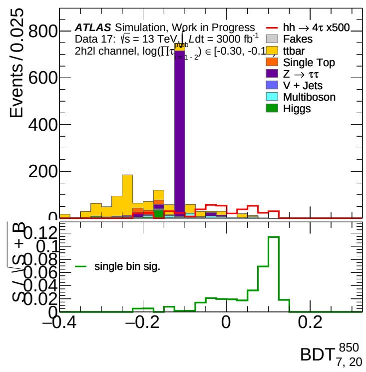
(e) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-0.75, -0.60)$



(f) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-0.60, -0.45)$



(g) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-0.45, -0.30)$



(h) $\log\left(\prod \tau^{prob}_{i=1-3}\right) \in [-0.30, -0.15)$

Figure A.16: Signal and background distributions for each $\log\left(\prod \tau_i^{\text{prob}}\right)$ intervals for the $2h2\ell$ channel. Each plot corresponds to a row in figure A.14(b).

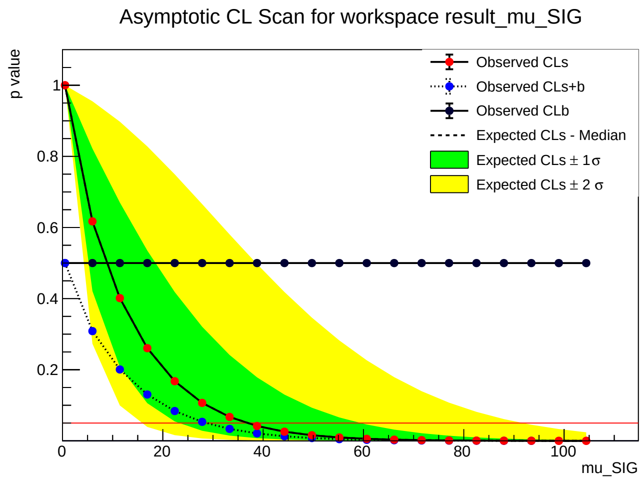


Figure A.17: The p value of the null-hypothesis as a function of the relative signal strength for an integrated luminosity of 36.1 fb^{-1}. The full phase space is utilised by dividing the $\log\left(\prod \tau_i^{\text{prob}}\right)$ into eight intervals for the leptonic channels. As the data was assumed to coincide with the expectation (Asimov data), the observed limit coincides with the expected limit. An upper limit of the signal cross section in comparison to the expected SM cross section of $\mu_{\text{SIG}} = \sigma_{\text{ggF}}/\sigma_{\text{ggF}}^{\text{SM}} = 36.9^{+22.5}_{-13.5}$ is observed for this fit configuration.

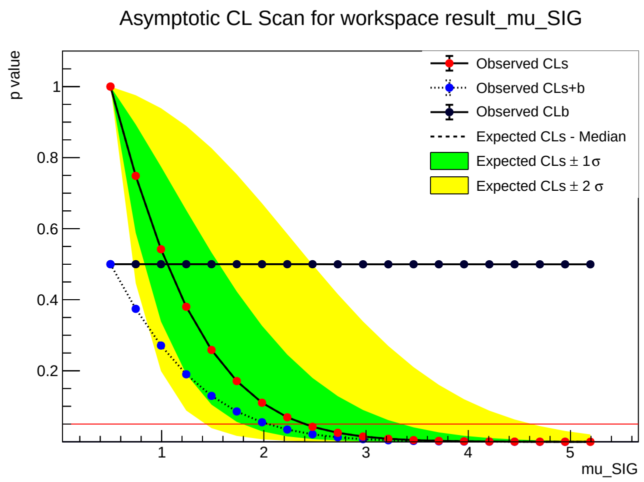


Figure A.18: The p value of the null-hypothesis as a function of the relative signal strength for an integrated luminosity of 3 000 fb^{-1}. The full phase space is utilised by dividing the $\log\left(\prod \tau_i^{\text{prob}}\right)$ into eight intervals for the leptonic channels. As the data was assumed to coincide with the expectation (Asimov data), the observed limit coincides with the expected limit. An upper limit of the signal cross section in comparison to the expected SM cross section of $\mu_{\text{SIG}} = \sigma_{\text{ggF}}/\sigma_{\text{ggF}}^{\text{SM}} = 2.40^{+0.95}_{-0.61}$ is observed for this fit configuration.

Bibliography

[1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579 (cit. on pp. 3, 8).

[2] S. Weinberg, *Elementary particle theory of composite particles*, Phys. Rev. **130** (1963) 776 (cit. on p. 3).

[3] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264 (cit. on pp. 3, 8).

[4] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367 (cit. on pp. 3, 8).

[5] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys. Lett. **B47** (1973) 365 (cit. on pp. 3, 6).

[6] B. Povh et al., *Teilchen und Kerne*, 9th ed., ISBN: 978-3-642-37821-8, Springer, 2014 (cit. on p. 3).

[7] G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv: `1207.7214 [hep-ex]` (cit. on pp. 3, 11, 68).

[8] S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv: `1207.7235 [hep-ex]` (cit. on pp. 3, 11, 68).

[9] G. Aad et al., *Combined Measurement of the Higgs Boson Mass in pp Collisions at* $\sqrt{s} = 7$ *and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (19 2015) 191803, URL: `https://link.aps.org/doi/10.1103/PhysRevLett.114.191803` (cit. on pp. 3, 11).

[10] N. MissMJ et al., *Standard Model of Elementary Particles modified version*, visited on 09/04/2019, URL: `https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles_modified_version.svg` (cit. on p. 4).

[11] M. Tanabashi et al., *Review of Particle Physics*, Phys. Rev. D **98** (3 2018) 030001, URL: `https://link.aps.org/doi/10.1103/PhysRevD.98.030001` (cit. on pp. 3, 7, 11, 27, 28, 34).

[12] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Reading, USA: Addison-Wesley, 1995, ISBN: 9780201503975, 0201503972, URL: `http://www.slac.stanford.edu/~mpeskin/QFT.html` (cit. on pp. 5, 7–9).

[13] M. D. Schwartz, *Quantum Field Theory and the Standard Model*, Cambridge University Press, 2014, ISBN: 1107034736, 9781107034730, URL: `http://www.cambridge.org/us/academic/subjects/physics/theoretical-physics-and-mathematical-physics/quantum-field-theory-and-standard-model` (cit. on p. 5).

[14] C. G. Callan, *Broken Scale Invariance in Scalar Field Theory*, Phys. Rev. D **2** (8 1970) 1541, URL: https://link.aps.org/doi/10.1103/PhysRevD.2.1541 (cit. on p. 6).

[15] K. Symanzik, *Small distance behaviour in field theory and power counting*, Communications in Mathematical Physics **18**.3 (1970) 227, ISSN: 1432-0916, URL: https://doi.org/10.1007/BF01649434 (cit. on p. 6).

[16] C. N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (1 1954) 191, URL: https://link.aps.org/doi/10.1103/PhysRev.96.191 (cit. on p. 6).

[17] M. Thomson, *Modern particle physics*, New York: Cambridge University Press, 2013, ISBN: 9781107034266, URL: http://www-spires.fnal.gov/spires/find/books/www?cl=QC793.2.T46::2013 (cit. on pp. 7–9, 26).

[18] C. S. Wu et al., *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (4 1957) 1413, URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413 (cit. on p. 8).

[19] M. Goldhaber, L. Grodzins and A. W. Sunyar, *Helicity of Neutrinos*, Phys. Rev. **109** (3 1958) 1015, URL: https://link.aps.org/doi/10.1103/PhysRev.109.1015 (cit. on p. 8).

[20] G. 't Hooft, *Renormalization of Massless Yang-Mills Fields*, Nucl. Phys. **B33** (1971) 173 (cit. on p. 8).

[21] G. 't Hooft, *Renormalizable Lagrangians for Massive Yang-Mills Fields*, Nucl. Phys. **B35** (1971) 167 (cit. on p. 8).

[22] G. 't Hooft and M. J. G. Veltman, *Regularization and Renormalization of Gauge Fields*, Nucl. Phys. **B44** (1972) 189 (cit. on p. 8).

[23] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (16 1964) 508, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508 (cit. on p. 9).

[24] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (9 1964) 321, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321 (cit. on p. 9).

[25] F. Tanedo, *Why do we expect a Higgs boson? Part II: Unitarization of Vector Boson Scattering*, (strongly modified), visited on 12/04/2019, URL: http://cph-theory.persiangig.com/90.11.26-3.htm (cit. on p. 9).

[26] Y. Fukuda et al., *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81** (8 1998) 1562, URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562 (cit. on p. 11).

[27] Q. R. Ahmad et al., *Measurement of the Rate of $\nu_e + d \to p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87** (7 2001) 071301, URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.071301 (cit. on p. 11).

[28] Q. R. Ahmad et al., *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (1 2002) 011301, URL: https://link.aps.org/doi/10.1103/PhysRevLett.89.011301 (cit. on p. 11).

[29] J. Rojo et al.,
The PDF4LHC report on PDFs and LHC data: results from Run I and preparation for Run II,
Journal of Physics G: Nuclear and Particle Physics **42**.10 (2015) 103103,
URL: https://doi.org/10.1088%2F0954-3899%2F42%2F10%2F103103 (cit. on p. 11).

[30] LHC Higgs Cross Section Working Group (2012-2013),
Standard Model Higgs boson decay branching ratios, (modified), visited on 14/04/2019,
URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections
(cit. on p. 13).

[31] G. Bernardi, M. Carena and T. Junk, *Higgs Bosons: Theory and Searches*, visited on 14/04/2019, 2012, URL: http://pdg.lbl.gov/2012/reviews/rpp2012-rev-higgs-boson.pdf
(cit. on p. 13).

[32] LHC Higgs Cross Section Working Group, *Higgs cross sections and decay branching ratios*, visited on 29/04/2019,
URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG
(cit. on pp. 13, 14, 21).

[33] G. Aad et al.,
Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector,
Eur. Phys. J. **C75**.10 (2015) 476, [Erratum: Eur. Phys. J.C76,no.3,152(2016)],
arXiv: 1506.05669 [hep-ex] (cit. on p. 12).

[34] 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. **C76**.1 (2016) 6, arXiv: 1507.04548 [hep-ex] (cit. on p. 12).

[35] V. Khachatryan et al.,
Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV,
Eur. Phys. J. **C75**.5 (2015) 212, arXiv: 1412.8662 [hep-ex] (cit. on p. 12).

[36] V. Khachatryan et al., *Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV*, Phys. Rev. **D92**.1 (2015) 012004,
arXiv: 1411.3441 [hep-ex] (cit. on p. 12).

[37] G. Aad et al.,
Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ *and 8 TeV*,
JHEP **08** (2016) 045, arXiv: 1606.02266 [hep-ex] (cit. on p. 12).

[38] M. Aaboud et al., *Cross-section measurements of the Higgs boson decaying into a pair of* τ*-leptons in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*,
Phys. Rev. **D99** (2019) 072001, arXiv: 1811.08856 [hep-ex] (cit. on pp. 12, 68).

[39] T. Plehn and M. Rauch, *The quartic higgs coupling at hadron colliders*,
Phys. Rev. **D72** (2005) 053008, arXiv: hep-ph/0507321 [hep-ph] (cit. on p. 14).

[40] T. Binoth et al., *Multi-Higgs boson production in the Standard Model and beyond*,
Phys. Rev. **D74** (2006) 113008, arXiv: hep-ph/0608057 [hep-ph] (cit. on p. 14).

[41] E. Accomando et al., "Physics at the CLIC multi-TeV linear collider", *Proceedings, 11th International Conference on Hadron spectroscopy (Hadron 2005): Rio de Janeiro, Brazil, August 21-26, 2005*, 2004, arXiv: hep-ph/0412251 [hep-ph], URL: http://weblib.cern.ch/abstract?CERN-2004-005 (cit. on p. 14).

[42] G. Degrassi, M. Fedele and P. P. Giardino, *Constraints on the trilinear Higgs self coupling from precision observables*, JHEP **04** (2017) 155, arXiv: 1702.01737 [hep-ph] (cit. on p. 14).

[43] M. Gorbahn and U. Haisch, *Indirect probes of the trilinear Higgs coupling: $gg \to h$ and $h \to \gamma\gamma$*, JHEP **10** (2016) 094, arXiv: 1607.03773 [hep-ph] (cit. on p. 14).

[44] G. Degrassi et al., *Probing the Higgs self coupling via single Higgs production at the LHC*, JHEP **12** (2016) 080, arXiv: 1607.04251 [hep-ph] (cit. on p. 14).

[45] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, CERN Yellow Reports: Monographs, 869 pages, 295 figures, 248 tables and 1645 citations. Working Group web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG, 2016, URL: https://cds.cern.ch/record/2227475 (cit. on p. 14).

[46] E. W. N. Glover and J. J. van der Bij, *HIGGS BOSON PAIR PRODUCTION VIA GLUON FUSION*, Nucl. Phys. **B309** (1988) 282 (cit. on p. 14).

[47] G. D. Kribs and A. Martin, *Enhanced di-Higgs Production through Light Colored Scalars*, Phys. Rev. **D86** (2012) 095023, arXiv: 1207.4496 [hep-ph] (cit. on p. 14).

[48] M. Aaboud et al., *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117**.18 (2016) 182002, arXiv: 1606.02625 [hep-ex] (cit. on pp. 14, 20).

[49] M. Cepeda et al., *Higgs Physics at the HL-LHC and HE-LHC*, (2019), arXiv: 1902.00134 [hep-ph] (cit. on pp. 14, 15, 87, 92).

[50] M. J. Dolan, C. Englert and M. Spannowsky, *Higgs self-coupling measurements at the LHC*, JHEP **10** (2012) 112, arXiv: 1206.5001 [hep-ph] (cit. on p. 14).

[51] M. Aaboud et al., *Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. Lett. **121**.19 (2018) 191801, [Erratum: Phys. Rev. Lett.122,no.8,089901(2019)], arXiv: 1808.00336 [hep-ex] (cit. on pp. 14, 81, 89, 92).

[52] M. Aaboud et al., *Search for pair production of Higgs bosons in the $b\bar{b}b\bar{b}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **01** (2019) 030, arXiv: 1804.06174 [hep-ex] (cit. on pp. 14, 68, 89, 92).

[53] M. Aaboud et al., *Search for Higgs boson pair production in the $\gamma\gamma b\bar{b}$ final state with 13 TeV pp collision data collected by the ATLAS experiment*, JHEP **11** (2018) 040, arXiv: 1807.04873 [hep-ex] (cit. on pp. 14, 89, 92).

[54] M. Aaboud et al., *Search for Higgs boson pair production in the $\gamma\gamma WW^*$ channel using pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. **C78**.12 (2018) 1007, arXiv: 1807.08567 [hep-ex] (cit. on pp. 14, 89, 90, 92).

[55] M. Aaboud et al., *Search for Higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **04** (2019) 092, arXiv: 1811.04671 [hep-ex] (cit. on pp. 14, 89, 90, 92).

[56] M. Aaboud et al., *Search for Higgs boson pair production in the $WW^{(*)}WW^{(*)}$ decay channel using ATLAS data recorded at $\sqrt{s} = 13$ TeV*, JHEP **05** (2019) 124, arXiv: 1811.11028 [hep-ex] (cit. on pp. 14, 89, 90, 92).

[57] L. Di Luzio, R. Gröber and M. Spannowsky, *Maxi-sizing the trilinear Higgs self-coupling: how large could it be?*, Eur. Phys. J. **C77**.11 (2017) 788, arXiv: 1704.02311 [hep-ph] (cit. on p. 15).

[58] A. Einstein, *Die Feldgleichungen der Gravitation*, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 844-847. (1915), Provided by the SAO/NASA Astrophysics Data System, URL: http://adsabs.harvard.edu/abs/1915SPAW.......844E (cit. on p. 15).

[59] S. M. Carroll, *Spacetime and geometry: An introduction to general relativity*, 2004, ISBN: 0805387323, 9780805387322, URL: http://www.slac.stanford.edu/spires/find/books/www?cl=QC6:C37:2004 (cit. on p. 15).

[60] N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, (2018), arXiv: 1807.06209 [astro-ph.CO] (cit. on p. 16).

[61] A. Riotto, "Theories of baryogenesis", *Proceedings, Summer School in High-energy physics and cosmology: Trieste, Italy, June 29-July 17, 1998*, 1998 326, arXiv: hep-ph/9807454 [hep-ph] (cit. on p. 16).

[62] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012) 1, arXiv: 1106.0034 [hep-ph] (cit. on p. 17).

[63] S. Dimopoulos and H. Georgi, *Softly Broken Supersymmetry and SU(5)*, Nucl. Phys. **B193** (1981) 150 (cit. on p. 17).

[64] S. P. Martin, *A Supersymmetry primer*, (1997) 1, [Adv. Ser. Direct. High Energy Phys.18,1(1998)], arXiv: hep-ph/9709356 [hep-ph] (cit. on p. 17).

[65] Z. Chacko et al., *Natural little hierarchy from a partially goldstone twin Higgs*, JHEP **01** (2006) 126, arXiv: hep-ph/0510273 [hep-ph] (cit. on p. 17).

[66] R. Grober and M. Muhlleitner, *Composite Higgs Boson Pair Production at the LHC*, JHEP **06** (2011) 020, arXiv: 1012.1562 [hep-ph] (cit. on p. 17).

[67] J. Mrazek et al., *The Other Natural Two Higgs Doublet Model*, Nucl. Phys. **B853** (2011) 1, arXiv: 1105.5403 [hep-ph] (cit. on p. 17).

[68] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, Phys. Rev. Lett. **83** (1999) 3370, arXiv: hep-ph/9905221 [hep-ph] (cit. on p. 17).

[69] P. Basler, M. Mühlleitner and J. Wittbrodt, *The CP-Violating 2HDM in Light of a Strong First Order Electroweak Phase Transition and Implications for Higgs Pair Production*, JHEP **03** (2018) 061, arXiv: 1711.04097 [hep-ph] (cit. on p. 17).

[70] O. S. Brüning et al., *LHC Design Report*, vol. v.1 : the LHC Main Ring, CERN Yellow Reports: Monographs, Geneva: CERN, 2004, URL: https://cds.cern.ch/record/782076 (cit. on p. 19).

[71] O. S. Brüning et al., *LHC Design Report*, vol. v.2 : the LHC Infrastructure and General Services, CERN Yellow Reports: Monographs, Geneva: CERN, 2004, URL: https://cds.cern.ch/record/815187 (cit. on p. 19).

[72] M. Benedikt et al., *LHC Design Report*, vol. v.3 : the LHC Injector Chain, CERN Yellow Reports: Monographs, Geneva: CERN, 2004, URL: https://cds.cern.ch/record/823808 (cit. on p. 19).

[73] Weltmaschine, visited on 17/04/2019, URL: http://www.weltmaschine.de/service__material/mediathek/lhc/ (cit. on p. 19).

[74] R. Steerenberg, *LHC Report: Protons: mission accomplished*, visited on 17/04/2019, 2018, URL: https://home.cern/news/news/physics/lhc-report-protons-mission-accomplished (cit. on p. 20).

[75] ATLAS Experiment, *Total Integrated Luminosity at Run 2 (13 TeV pp data only)*, visited on 17/04/2019, 2018, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (cit. on p. 20).

[76] HL-LHC Project, *Project Schedule*, visited on 17/04/2019, URL: https://project-hl-lhc-industry.web.cern.ch/content/project-schedule (cit. on pp. 21, 92).

[77] G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on pp. 21–25).

[78] *ATLAS inner detector: Technical Design Report, 1*, Technical Design Report ATLAS, Geneva: CERN, 1997, URL: https://cds.cern.ch/record/331063 (cit. on p. 23).

[79] S. Haywood et al., *ATLAS inner detector: Technical Design Report, 2*, Technical Design Report ATLAS, Geneva: CERN, 1997, URL: https://cds.cern.ch/record/331064 (cit. on p. 23).

[80] K. Potamianos, *The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider*, PoS **EPS-HEP2015** (2015) 261, arXiv: 1608.07850 [physics.ins-det] (cit. on p. 23).

[81] A. Ahmad et al., *The Silicon Microstrip Sensors of the ATLAS SemiConductor Tracker*, tech. rep. ATL-INDET-PUB-2007-007. ATL-COM-INDET-2007-008. CERN-ATL-COM-INDET-2007-008. 1, CERN, 2007, URL: http://cds.cern.ch/record/1019885 (cit. on p. 23).

[82] E. Abat et al., *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance*, JINST **3** (2008) P02013 (cit. on p. 23).

[83] Y. Nakahama, *The ATLAS Trigger System: Ready for Run-2*, Journal of Physics: Conference Series **664**.8 (2015) 082037, URL: http://stacks.iop.org/1742-6596/664/i=8/a=082037 (cit. on p. 25).

[84] P. Jenni et al., *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*, Technical Design Report ATLAS, Geneva: CERN, 2003, URL: https://cds.cern.ch/record/616089 (cit. on p. 25).

[85] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on p. 26).

[86] G. Aad et al., *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11**.04 (2016) P04008, arXiv: 1512.01094 [hep-ex] (cit. on p. 27).

[87] T. A. collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, (2016) (cit. on p. 27).

[88] G. Aad et al., *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, Eur. Phys. J. **C74**.7 (2014) 2941, arXiv: 1404.2240 [hep-ex] (cit. on p. 27).

[89] A. Ezhilov, *Electrons in ATLAS: from Run 1 to Run 2*, tech. rep. ATL-PHYS-PROC-2015-180, 2015, URL: https://cds.cern.ch/record/2112121 (cit. on p. 27).

[90] G. Aad et al., *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s}$ *=13 TeV*, Eur. Phys. J. **C76**.5 (2016) 292, arXiv: 1603.05598 [hep-ex] (cit. on p. 28).

[91] H. Herde, *Muon reconstruction performance in ATLAS at Run-II*, tech. rep. ATL-PHYS-PROC-2015-121, CERN, 2015, URL: https://cds.cern.ch/record/2059849 (cit. on p. 28).

[92] M. J. Schultens, *Search for Supersymmetry with Hadronically and Leptonically Decaying Tau Leptons at the ATLAS Experiment*, Doktorarbeit: Universität Bonn, 2016, URL: https://web.physik.uni-bonn.de/group/view.php?&group=1&lang=de&c=t&id=78 (cit. on p. 28).

[93] Y. Sakurai, "The ATLAS Tau Trigger Performance during LHC Run 1 and Prospects for Run 2", *Proceedings, 2nd Conference on Large Hadron Collider Physics Conference (LHCP 2014): New York, USA, June 2-7, 2014*, 2014, arXiv: 1409.2699 [hep-ex], URL: http://www.slac.stanford.edu/econf/C140602.2/papers/1409.2699v1.pdf (cit. on p. 28).

[94] G. Aad et al., *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at* $\sqrt{s}$*=8 TeV*, Eur. Phys. J. **C75**.7 (2015) 303, arXiv: 1412.7086 [hep-ex] (cit. on p. 29).

[95] *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, tech. rep. ATL-PHYS-PUB-2015-045, CERN, 2015, URL: https://cds.cern.ch/record/2064383 (cit. on pp. 29, 42, 43, 54).

[96] The ATLAS collaboration, *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 pp. 29, 39, 41–43).

[97] S. Höche, "Introduction to parton-shower event generators", *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, 2015 235, arXiv: 1411.4085 [hep-ph] (cit. on p. 30).

[98] A. Buckley et al., *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011) 145, arXiv: 1101.2599 [hep-ph] (cit. on p. 30).

[99] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250 (cit. on p. 30).

[100] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on p. 31).

[101] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph] (cit. on p. 31).

[102] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph] (cit. on p. 31).

[103] S. Frixione et al., *The MCaNLO 4.0 Event Generator*, (2010), arXiv: 1010.0819 [hep-ph] (cit. on pp. 31, 39).

[104] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. **C76**.4 (2016) 196, arXiv: 1512.01178 [hep-ph] (cit. on pp. 31, 39).

[105] C. Deutsch, *Identification and Classification of Hadronic Tau Lepton Decays in the ATLAS Experiment for Run 2 of the LHC*, Master thesis: Universität Bonn, 2017 (cit. on p. 31).

[106] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis with ROOT*, (2009), arXiv: physics/0703039, URL: https://root.cern.ch/tmva (cit. on pp. 31, 32, 81).

[107] Y. Freund and R. E. Schapire, "Experiments with a New Boosting Algorithm", *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*, ICML'96, Bari, Italy: Morgan Kaufmann Publishers Inc., 1996 148, ISBN: 1-55860-419-7, URL: http://dl.acm.org/citation.cfm?id=3091696.3091715 (cit. on p. 32).

[108] ATauLeptonAnalysiS, *xTauFramework*, visited on 21/05/2019, URL: https://gitlab.cern.ch/ATauLeptonAnalysiS/xTauFramework (cit. on p. 32).

[109] ATauLeptonAnalysiS, *HAPPy: Higgs Analysis Plotting in Python*, visited on 21/05/2019, URL: https://gitlab.cern.ch/ATauLeptonAnalysiS/HAPPy (cit. on p. 32).

[110] I. Antcheva et al., *ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization*, Comput. Phys. Commun. **180** (2009) 2499, arXiv: 1508.07749 [physics.data-an] (cit. on p. 32).

[111] M. Baak et al., *HistFitter software framework for statistical data analysis*, Eur. Phys. J. **C75** (2015) 153, arXiv: 1410.1280 [hep-ex] (cit. on pp. 32, 67, 88).

[112] M. Aaboud et al., *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. **C77**.7 (2017) 466, arXiv: 1703.10485 [hep-ex] (cit. on p. 40).

[113] G. Aad et al., *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, Eur. Phys. J. **C76**.5 (2016) 295, arXiv: 1512.05955 [hep-ex] (cit. on p. 40).

[114] D. Duschinger, *Track Classification - R21 Finalization*, tech. rep., 2017, URL: https://indico.cern.ch/event/615208/contributions/2481630/attachments/1415563/2167141/TauCP_TrackClassificationTaskforce_20170220.pdf (cit. on p. 41).

[115] ATLAS Collaboration, Tau Working Group, personal communication, 2019 (cit. on p. 43).

[116] ATLAS Collaboration, *Combination of searches for Higgs boson pairs in pp collisions at 13 TeV with the ATLAS experiment.*, (2018), URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2018-043/ (cit. on p. 90).

List of Figures

List of Tables