GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Fakultät für Physik

Bachelor's Thesis

Studien zur Signaloptimierung in der Suche nach Higgs-Paarproduktion im $bb\tau\tau$-Zerfallskanal mit dem ATLAS-Detektor

Signal Extraction Studies in the Search for Higgs Boson Pair Production in the $bb\tau\tau$ Decay Channel with ATLAS

prepared by

André Wilhahn

from Eschwege

at the II. Physikalischen Institut

Thesis number:	II.Physik-UniGö-BSc-2020/05
Thesis period:	11th May 2020 until 31st August 2020
First referee:	Prof. Dr. Stan Lai
Second referee:	Priv.Doz. Dr. Jörn Große-Knetter

Abstract

Diese Bachelorarbeit beschäftigt sich mit der Suche nach sowohl nicht-resonanter, als auch resonanter Higgs-Paarproduktion im $b\bar{b}\tau\tau$-Zerfallskanal. Hierfür werden $139\,\mathrm{fb}^{-1}$ an Kollisionsdaten verwendet, die vom Atlas Experiment am Lhc an der Europäischen Organisation für Kernforschung (Cern) während der zweiten großen Datenaufnahme, dem Run 2, aufgenommen wurden. Die Schwerpunktsenergie betrug während des Run 2 $\sqrt{s} = 13\,\mathrm{TeV}$. Weiterhin liegt der Fokus bei beiden Analysen (sowohl bei der Suche nach nicht-resonanter als auch bei der nach resonanter Higgs-Paarproduktion) auf dem hadronischen Zerfall der beiden Tauonen.

Für eine bessere Untersuchung des $HH \to b\bar{b}\tau\tau$ Zerfallsprozesses wurde ein Boosted Decision Tree verwendet, um eine bessere Unterscheidung zwischen Signal und Hintergrundprozessen zu gewährleisten: In der nicht-resonanten Analyse wurden zwei verschiedene Sets von Hyperparametern verwendet, auf deren Basis der Boosted Decision Tree trainiert wurde. In der resonanten Analyse hingegen - in der nach einer skalaren Resonanz mit schmaler Zerfallsbreite gesucht wurde - wurde eine festgelegte Menge an Hyperparametern verwendet, jedoch wurde das Training des Boosted Decision Trees zum einen auf Basis einer einzigen Resonanzmasse durchgeführt und zum anderen auf Basis dreier benachbarter Resonanzmassen. Die Ergebnisse dieser beiden Analysen werden in dieser Arbeit vorgestellt werden.

Abstract

This bachelor thesis presents studies on signal extraction in searches for non-resonant and resonant Higgs boson pair production in the $b\bar{b}\tau\tau$ final state. It is based on $139\,\mathrm{fb}^{-1}$ of collision data that was recorded by the Atlas experiment in the Run 2 data-taking period from 2015 to 2018 at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ at the Lhc at Cern. Furthermore, this thesis focuses on the hadronic decay mode ($\tau_{\mathrm{had}}\tau_{\mathrm{had}}$) of both tau leptons in the final state.

For a better investigation of the $HH \to b\bar{b}\tau\tau$ process, a boosted decision tree was trained on Run 2 data in order to improve the signal-to-background separation in both the non-resonant and the resonant analysis: In the non-resonant analysis, the boosted decision tree was trained on two sets of hyperparameters and the performance of both sets is compared. In the resonant analysis, the boosted decision tree was trained on a single set of hyperparameters, but in one case on a single resonance mass point and in a second case on the target signal and two neighbouring resonance mass points, which are also compared.

Contents

1 Introduction **3**

2 Theoretical Background **5**

2.1 The Standard Model 5

2.2 Electroweak Symmetry Breaking 8

2.2.1 The Higgs Mechanism 8

2.2.2 Higgs Self-Coupling and Pair Production in the Standard Model . . 10

2.3 Beyond the Standard Model 13

2.3.1 Shortcomings of the Standard Model 13

2.3.2 Two Higgs Doublet Models 14

2.3.3 Resonant Higgs Pair Production 15

3 Experimental Setup **17**

3.1 The Large Hadron Collider 17

3.2 The ATLAS detector 18

3.2.1 The Inner Detector 20

3.2.2 The Calorimeter 21

3.2.3 The Muon Spectrometer and the Toroid Magnets 23

3.2.4 The Trigger System 24

4 Higgs Pair Production in the $bb\tau\tau$ final state **25**

4.1 Previous Search Results at ATLAS 25

4.2 Search for $HH \to bb\tau\tau$ with the Full Run-2-Dataset 28

4.2.1 Event Selection and Search Strategy 29

4.2.2 Kinematic Distributions 30

5 Signal Extraction using Boosted Decision Trees **37**

5.1 Introduction to Boosted Decision Trees 37

5.2 The BDT for the non-resonant analysis 38

5.3 The BDT for the resonant analysis 41

5.3.1 Training on a single mass point 41

5.3.2 Training on adjacent mass points 41
5.4 Results 42

6 Conclusion and Outlook **47**

Acknowledgements

Writing this bachelor thesis at the II. Physikalisches Institut has been a great experience and gave me a deep insight into the work life of an experimental particle physicist, but was also very nerve-wracking and ambitious project. I, for sure, would not have been able to do this research on my own if it was not for so many people to support me on this journey. This is why I want to thank the following persons:

First, I want to thank my friends and especially my family for their never-ending support in any possible way. The last three and a half months definitely had their ups and downs, during investigations for my research topic as well as the generally difficult time period in which I wrote this thesis. Thank you all so much for believing in me when I could not and for showing your understanding whenever I was only thinking about this thesis and nothing else.

My second big thanks goes to Petar Bokan and Andrés Melo. I cannot remember how many times I had to ask both of you a stupid question, or how many times I asked you for help because I had problems with code that was not working, or how often you had to explain to me the most trivial concepts a dozen times, only because my brain was (again) overloaded by the amount of information. Whatever the problem was, you guys were always there for me and I cannot put into words how thankful I am for your patience, calmness, and nerves of steel.

My last and biggest thanks goes to Prof. Dr. Stan Lai. During the fifth semester, I was really unsure whether I would find a topic for a bachelor thesis that I would like. Thinking in retrospect about the idea to ask you for a bachelor thesis project, this was by far the best idea that I could have come up with. Your supervision, your integration of us bachelor students into your working group, your continuous support, your down-to-earthness and many other laudable actions and attributes cannot be taken for granted and I am looking forward continuing to work with you in the near future.

1 Introduction

Over the last 100 years, physicists have expanded our understanding of the universe tremendously. Initiated by scientists like Albert Einstein, Max Planck, Niels Bohr and many more, physicists made giant leaps in expanding the mathematical structure on which modern physics is based on. The four fundamental forces of the universe - gravity, electromagnetism, the weak and the strong nuclear forces - can all be described with astonishing precision by two theories: Albert Einstein's general theory of relativity, which describes gravity as curvature of spacetime [1], and the Standard Model (SM), which incorporates quantum chromodynamics, quantum electrodynamics and the theory of weak nuclear interactions [2–4]. However, the SM seems to be an incomplete theory. Even though it proved itself as extremely accurate in describing many phenomena related to elementary particles, there are still questions that cannot be answered by the SM. Furthermore, only two of the four fundamental interactions have been unified by now: the weak nuclear force and electromagnetism, which are combined into the theory of electroweak interactions. Thus, the search for a more fundamental theory is one of the most important tasks for theoretical and experimental particle physicists today.

Investigation of resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau_{\text{had}}\tau_{\text{had}}$ final state is presented in this thesis. The Higgs boson, the most important constituent of the SM for this research, was discovered in 2012 by the ATLAS and CMS Collaboration at the LHC at CERN with a mass of approximately 125 GeV [5, 6]. A deep study of the Higgs boson is crucial for probing the electroweak symmetry breaking [7–9], which is an essential part of the SM. Furthermore, many theories beyond the SM predict heavy resonances that could decay into pairs of Higgs bosons.

This thesis is based on $139\,\text{fb}^{-1}$ collision data recorded by the ATLAS experiment at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ between the years 2015 and 2018 at the LHC at CERN. *Boosted decision trees* (BDTs) are trained on this data to improve the signal-to-background separation in both the resonant and non-resonant analysis.

The structure of this thesis is as follows: In Chapter 2, the theoretical background behind the SM and electroweak symmetry breaking will be discussed, where the Higgs mechanism and Higgs self-coupling in the SM will be outlined in detail. A discussion of the shortcomings of the SM as well as an introduction to two Higgs doublet models and resonant Higgs boson pair production follow after that. Chapter 3 describes the LHC and the ATLAS detector, while Chapter 4 describes Higgs boson pair production in the $b\bar{b}\tau\tau$ final state, where previous ATLAS results and the search for $HH \to b\bar{b}\tau_{\text{had}}\tau_{\text{had}}$ decays will be presented. Chapter 5 will outline signal extraction studies with BDTs. An introduction to BDTs will be given, followed by the results of the (non-)resonant analysis. A conclusion and an outlook will be given in Chapter 6.

2 Theoretical Background

In the following section, an introduction to the Standard Model will be given, as well as a discussion of electroweak symmetry breaking. In particular, the Higgs mechanism will be presented because this process explains how certain elementary particles obtain their masses. In addition to this, Higgs self-coupling and consequently Higgs boson pair production will be further explained, since they are the fundamental processes for the research of this thesis. Furthermore, the shortcomings and deficits of the SM as well as theories beyond the SM (with the focus on two Higgs doublet models and resonant Higgs boson pair production) will be outlined.

2.1 The Standard Model

The SM is currently the best known description of the elementary particles that form all visible matter in the observable universe, along with the interactions between these elementary particles (except gravity). There are 17 elementary particles in total in the SM, which can be divided into three different groups with individual internal quantum numbers. A depiction of the particle content of the SM is given in Figure 2.1.

The first group of particles consists of six *quarks*. Quarks are *fermions*, i.e. spin-$\frac{1}{2}$ particles. They can be separated into three generations, where two distinct quark flavours always form one generation: The first and lightest generation contains up and down quarks, the second generation charm and strange quarks, while the third and heaviest generation consists of top and bottom quarks. It is also worth mentioning, since the heavier quarks tend to be unstable, that quarks with higher masses (from higher generations) decay into lower mass quarks (in lower generations), which is why lower mass quarks are more common in nature than higher mass quarks.

Quarks carry a colour charge (red, green or blue) which allows them to couple to the strong nuclear force. Furthermore, quarks are also able to interact via electromagnetism because of their electric charge: Up, charm, and top quarks carry an electric charge of

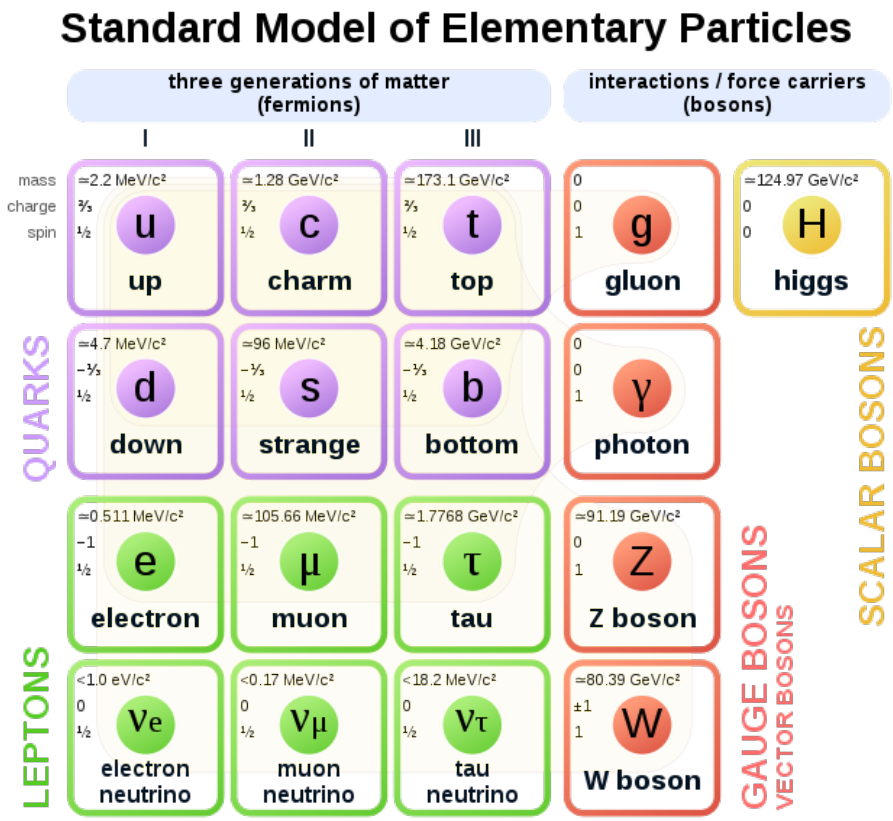


Figure 2.1: The Standard Model of particle physics. For each elementary particle, three quantum numbers are shown in the top left corner of every square: The highest one describes the mass of the particle in eV, the middle one is the electric charge of the particle in units of the elementary charge e, and the lowest one describes the spin of the particle.

$+\frac{2}{3}$, while down, strange, and bottom quarks carry an electric charge of $-\frac{1}{3}$. Moreover, quarks can interact via the weak nuclear force as well because of their half integer weak isospin which makes them the only elementary particles that couple to each of these three fundamental forces (up-type quarks possess a weak isospin of $+\frac{1}{2}$, whereas down-type quarks have a weak isospin of $-\frac{1}{2}$). In addition, there also exist antiquarks which are identical to their "normal" counterparts, but with the exception that they are oppositely charged in terms of electric and colour charge (the anticolour charges are called antired, antigreen, and antiblue).

Furthermore, quarks are subject to a physical phenomenon called *colour confinement*. This is caused by their strong nuclear potential which increases linearly as quarks diverge from each other. Consequently, to minimise their potential energy, quarks never appear

solely in nature, but only in groups of at least two. Bound states of quarks are called *hadrons*. Hadrons mostly appear in terms of *baryons*, which consist of three quarks with individual colour charges, *antibaryons*, which consist of three antiquarks with individual anticolour charges, and *mesons*, which are basically quark-antiquark pairs where the antiquark carries the anticolour charge that corresponds to the colour charge of the quark. As a result, hadrons always occupy colourless states.

Colour confinement also causes particle *jets* in experiments in high energy physics. If two or more initial quarks diverge very rapidly from each other within a particle detector, quark-antiquark pairs will emerge in order to decrease the potential energy between the initial quarks. As these new quark-antiquark pairs also travel apart from each other, more quark-antiquark pairs will arise. The production of these particle-antiparticle pairs increases exponentially, creating particle showers within a detector. These particle showers are called "jets".

Six *leptons* form the second group of particles in the SM. Leptons are fermions too, and can also be divided into three generations (similar to quarks): The electron and the electron neutrino (first generation), the muon and the muon neutrino (second generation), and the tau lepton and the tau neutrino (third generation). Electrons, muons, and tau leptons carry an electric charge of -1 and a weak isospin of $-\frac{1}{2}$, but no colour charge, whereas neutrinos neither have an electric nor a colour charge, but only a weak isospin of $+\frac{1}{2}$. Thus, electrons, muons, and tau leptons can couple electromagnetically and weakly, while neutrinos only interact via the weak nuclear force. Finally, just as each quark has an oppositely charged antiquark, each (charged) lepton has an (oppositely charged) antilepton with the same intrinsic properties as its counterpart.

The third and final group of all particles in the SM are the *elementary bosons*, where the term "boson" signifies an integer spin, i.e. a spin-1 or spin-0 particle. Four of the five elementary bosons are vector bosons and possess spin-1. They are the force-carriers of the SM where each vector boson mediates a certain force: The gluon is the carrier of the strong nuclear force and it possesses a colour charge too. Due to this, the gluon can also couple to itself and, just as quarks, it cannot be free due to colour confinement. The photon is the mediator of the electromagnetic force. It is electrically neutral and thus it does not couple to itself. Finally, the W and Z bosons are the gauge bosons of the weak nuclear force and the only massive elementary bosons. In contrast to the Z boson, the W boson has an integer weak isospin of ± 1, since it mediates between the fermions of the

SM (the Z boson has a weak isospin of 0). The W boson also carries an electric charge (either $+1$ or -1), whereas the Z boson is electrically neutral. The fifth elementary boson is the Higgs boson, the quantum of the Higgs field. Although the Higgs boson has an integer spin too, just like the other gauge bosons, it is a spin-0 particle in contrast to the other four spin-1 gauge bosons.

Gravity is not included in the SM. The coupling strength of gravitation is orders of magnitude weaker than those of the other forces, and its effects can be ignored in high energy interactions.

2.2 Electroweak Symmetry Breaking

2.2.1 The Higgs Mechanism

Experiments and measurements have shown that neither the Z nor the W boson are massless, but are even among the heaviest of all particles in the SM [10, 11], although the SM does not predict massive gauge bosons. The W boson has a mass of 80.4 GeV, the Z boson is slightly heavier with a mass of 91.2 GeV [12]. This is a contradiction that can be solved with the help of the *Higgs mechanism*.

The Higgs mechanism introduces a doublet of complex scalar fields of the form [13]

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} , \tag{2.1}$$

which allows considering a *Lagrange density* or *Lagrangian*

$$\mathcal{L} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \tag{2.2}$$

that is invariant under the gauge symmetry group of the SM. Here, $V(\Phi)$ is the Higgs potential given by

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

To have a finite minimum, λ must be larger than zero. However, μ^2 can be chosen arbitrarily. $\mu^2 > 0$ will result in a local minimum at $\Phi = 0$ (no symmetry breaking). The choice $\mu^2 < 0$, however, will yield an infinite amount of ground states located in a circle

around the symmetry axis of the potential given by:

$$\Phi^\dagger \Phi = -\frac{\mu^2}{2\lambda} =: \frac{\nu^2}{2} \,. \tag{2.4}$$

The shape of the Higgs potential is depicted in Figure 2.2 for this case, which has a non-zero *vacuum expectation value* for the Higgs field.

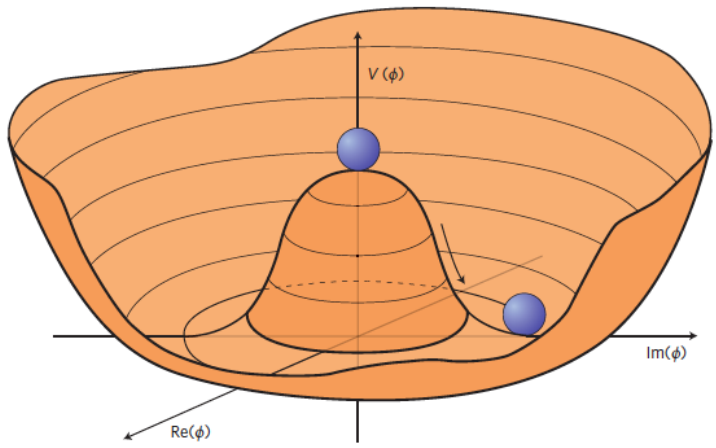


Figure 2.2: A depiction of the shape of the Higgs potential $V(\Phi)$ for the case $\mu^2 < 0$ [14].

Since there are infinitely many possible ground states, one can choose the ground state such that it lies on the real axis of the given coordinate system without any loss of generality, thus setting Φ to be entirely real. Furthermore, the electric neutrality of the ground state requires ϕ^+ to be equal to zero. Hence, $\phi_1 = \phi_2 = \phi_4 = 0$. After expanding Φ about the ground state, it can be written as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix} \,, \tag{2.5}$$

which corresponds to the unitary gauge. The selection of a specific ground state now breaks the symmetry of the system (thus the name *spontaneous symmetry breaking*). Substituting the unitary gauge into the Lagrangian will now yield terms which describe a new massive gauge field as well as the interaction with itself:

$$\mathcal{L} \supset \underbrace{\lambda \nu^2 h^2}_{\text{massive h scalar (Higgs field)}} \underbrace{+\lambda \nu h^3 + \frac{1}{4}\lambda h^4}_{\text{self-interaction of h}} \,. \tag{2.6}$$

The terms that are cubic and quartic in h describe the triple and quartic Higgs boson self-interaction. The term quadratic in h corresponds to the mass term of the Higgs boson at leading order:

$$m_h = \sqrt{2\lambda}\nu\,. \tag{2.7}$$

There will also be further terms incorporated into the Lagrangian that represent the coupling between the Higgs field and other gauge fields, which turn out to be the W and the Z boson. From these terms, both their masses can be determined which are identified as

$$m_W = \frac{1}{2} g_W \nu\,, \tag{2.8}$$

$$m_Z = \frac{1}{2} \frac{g_W}{\cos(\theta_W)} \nu\,, \tag{2.9}$$

where g_W is the coupling constant of the SU(2) interaction and θ_W is the Weinberg angle of the electroweak interaction.

The Higgs mechanism also explains how all SM fermions acquire masses. Their coupling to the Higgs fields is called the *Yukawa coupling* and generates masses that can be expressed as follows:

$$m_f = \frac{g_f \nu}{\sqrt{2}}\,. \tag{2.10}$$

Here, g_f represents the Yukawa coupling constant. Finally, it is worth to notice that the vacuum expectation value of the Higgs field can be determined through measurements. It is approximately equal to 246 GeV [15].

2.2.2 Higgs Self-Coupling and Pair Production in the Standard Model

As already outlined in Section 2.2.1, the Higgs boson has the ability to couple to itself, either in a triple or a quartic self-interaction. This is illustrated by the Feynman vertices in Figure 2.3.

These self-couplings play an important role in non-resonant Higgs boson pair production. However, Higgs boson pair production has a very low cross-section, even smaller than single Higgs production ($\sigma_{HH} = 31.05\,\text{fb}$ at $\sqrt{s} = 13\,\text{TeV}$ compared to the single Higgs production cross-section $\sigma_H = 43.92\,\text{pb}$ [16, 17]), making it extremely difficult to detect this process in a particle detector. Nevertheless, investigating Higgs boson pair produc-

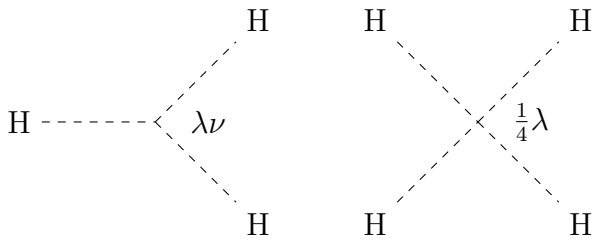


Figure 2.3: The Feynman diagrams for the triple and quartic self-interaction of the Higgs boson with their corresponding coupling strengths.

tion is important if one wants to study the self-coupling strength of the Higgs boson. At the LHC, pairs of Higgs bosons are dominantly produced in *gluon-gluon fusion* (ggF). This process is shown in Figure 2.4.

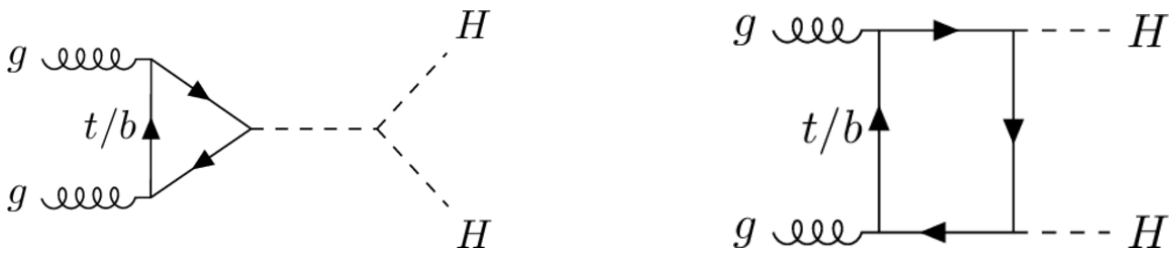


Figure 2.4: The leading order Feynman diagrams for Higgs boson pair production by ggF in the SM [18].

The left diagram in Figure 2.4 shows a Feynman diagram with a triangular top- or bottom-quark loop. The virtual Higgs boson decays into two final state Higgs bosons, completing the pair production. The right Feynman diagram depicts another way of producing Higgs pairs, namely via a heavy quark loop. This Feynman diagram, however, does not involve the Higgs self-coupling. Both production modes are called *non-resonant* Higgs pair production. In theory, the SM coupling constant of the triple Higgs self-interaction is given by:

$$g_{\mathrm{HHH}} = \frac{m_H^2}{2\nu} \,. \tag{2.11}$$

Putting this prediction to the test has become one of the major tasks in the search for di-Higgs production because an experimental confirmation of the Higgs self-interaction constant would be another proof for the correctness of the Higgs mechanism. The destructive interference, though, between the two Feynman diagrams shown above causes the low cross-section for Higgs boson pair production and, thus, impedes its experimental

observation.

However, there are also theories that predict larger cross-sections for the di-Higgs coupling as the SM does. If experiments showed that the actual value differs from the SM prediction, then this could be a hint to a more fundamental theory of particle interactions and new possible elementary particles.

There are also other production modes for Higgs boson pair production (for instance *vector boson fusion*). Their cross-sections are even smaller than the cross-section for ggF at the LHC, making them almost negligible in the observation of Higgs boson pair production. Figure 2.5 shows the predicted cross-section for Higgs boson pair production as a function of the centre-of-mass energy $\sqrt{s}$. In addition, Figure 2.6 shows experimental results from ATLAS and CMS that set limits on the cross-section of non-resonant Higgs boson pair production as a function of κ_λ, which is the ratio between the hypothesised Higgs self-coupling strength and the self-coupling strength that is predicted by the SM.

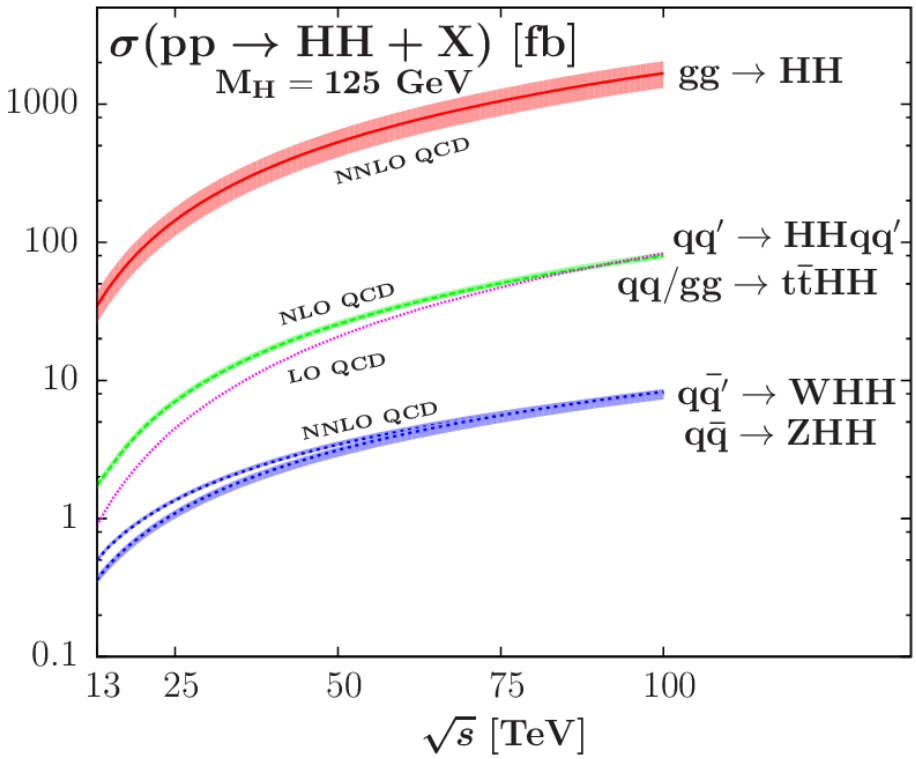


Figure 2.5: The cross-sections for different Higgs boson pair production modes as functions of the centre-of-mass energy, including their higher-order corrections [19].

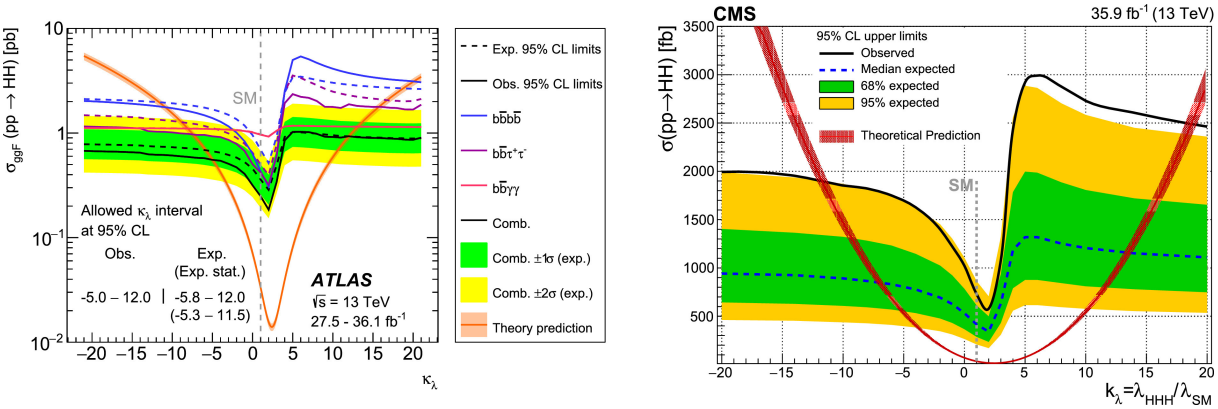


Figure 2.6: Cross-section limits of non-resonant Higgs boson pair production as a function of the ratio between the hypothesised Higgs self-coupling strength and the self-coupling strength predicted by the SM denoted as κ_λ. The left diagram was provided by the ATLAS experiment, the right one by CMS [20].

2.3 Beyond the Standard Model

2.3.1 Shortcomings of the Standard Model

Over the last decades, physicists have come across questions and problems that could not be explained by the theories of the day. Tied to particle physics and the SM, physicists realised that the SM cannot be a fundamental theory because it still leaves too many questions unanswered. A few of these unanswered problems are presented in the following.

A prominent example deals with a hypothesised new type of matter, called dark matter [21, 22], which is supposed to interact only via gravity with "normal" matter. The concept of dark matter is used to explain why galaxies are not torn apart by their own centrifugal forces and why the orbit velocities of stars within a galaxy stay nearly constant, even close to the edge of the disc where the orital velocities should decrease. Furthermore, the dark matter problem has not only become quite famous among astrophysicists, but also for particle physicists because experimental evidence for a dark matter particle could not only solve the rotation speed problem of galaxies, but could also lead to further theoretical expansion of the SM (for example by supersymmetric theories [23]), but only if R-parity is preserved [23] which is an important concept because otherwise supersymmetric theories would allow couplings that violate conservation of baryon and lepton numbers.

There are, of course, other problems closer related to the SM and its incompleteness, for example the baryon asymmetry of the universe [24]. As far as scientists can tell, the observable universe is made of "normal" matter and with hardly a trace of antimatter.

This implies that there must have been a large extent of CP-violating processes in the early universe that caused the huge imbalance between matter and antimatter [25]. However, the SM does not provide processes that include a sufficient amount of CP-violation in order to create such a big asymmetry.

Another unexplained phenomenon from the SM are neutrino oscillations [26]. These weakly interacting particles have been observed to oscillate in their flavour eigenstates, meaning that they can periodically change their "original" flavour for short periods of time due to phase differences in their wave functions, violating conservation of lepton flavour. This process does not happen instantaneously, which means that the neutrinos must experience time which, furthermore, indicates that they are moving slower than the speed of light in vacuum, making them massive particles which is experimentally confirmed [27]. The SM, however, does not predict massive neutrinos which also highlights its incompleteness.

2.3.2 Two Higgs Doublet Models

Many approaches have been taken to extend the SM, one of them being two Higgs doublet models (2HDMs). In Section 2.2.1, a complex doublet, now called Φ_1 instead of Φ, of scalar fields was added to the Lagrangian in order to explain the Higgs mechanism and electroweak symmetry breaking. This resulted in a spin-0 boson, namely the Higgs boson, that is now an inherent part of the SM. However, 2HDMs predict additional scalar bosons similar to the SM Higgs boson by adding another Higgs doublet Φ_2 to the Lagrangian. This leads to five scalar particles in total: The SM Higgs boson h, a pseudoscalar A, a neutral scalar H and two charged scalars $H^{\pm}$.

An example of a 2HDM is the minimal supersymmetric Standard Model (MSSM) [23]. It would not fully solve, for instance, the problem of baryogenesis mentioned in Section 2.3.1, but four additional scalar bosons would at least provide more sources of CP-violation that could have caused the imbalance of matter and antimatter in the early universe [28]. However, the MSSM could, for instance, solve the *hierarchy problem* [29] which deals with the question why the weak nuclear force is many orders of magnitude stronger than gravity and why the measured Higgs mass is so small, since large quantum corrections should make the Higgs mass very large. The MSSM provides explanations how the Higgs boson mass can be protected from large quantum corrections [29].

2.3.3 Resonant Higgs Pair Production

Some theories, such as the MSSM, predict heavy particles that could decay into pairs of Higgs bosons as shown in Figure 2.7.

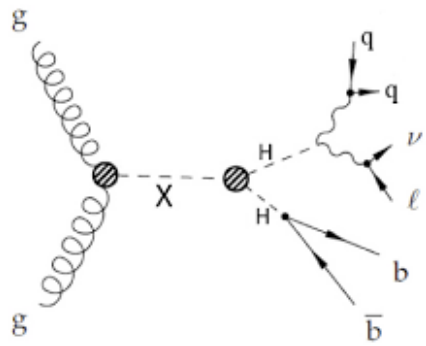


Figure 2.7: The leading order Feynman diagram for resonant Higgs boson pair production initiated by ggF with the $b\bar{b}q\bar{q}\ell\nu$ final state [30].

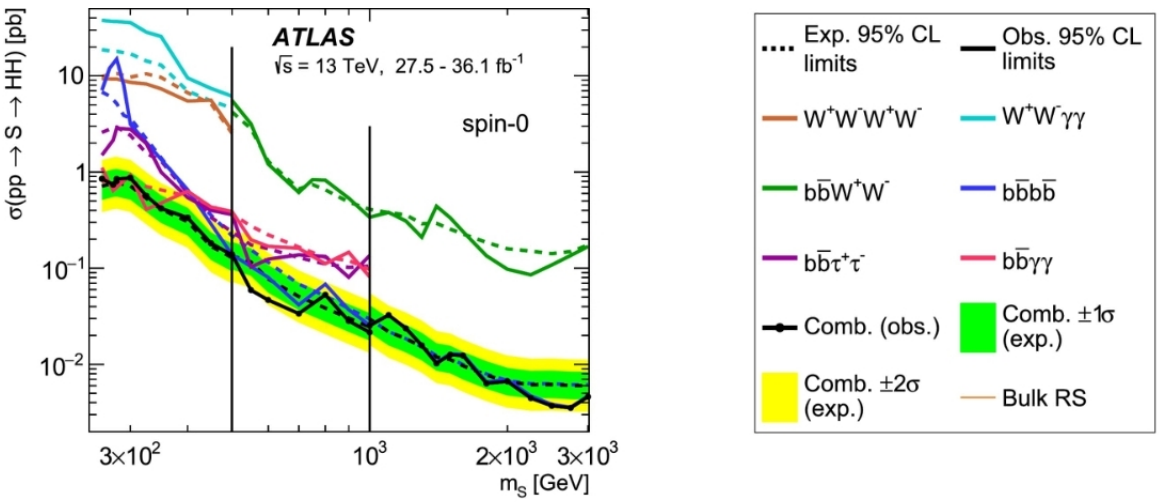


Figure 2.8: Current cross-section limits for resonant Higgs pair production as a function of resonance mass from ATLAS [31].

Resonant Higgs boson pair production plays a major role in the search for new particles that could extend the SM beyond its current limitations [32]. In this process, an on-shell heavy particle, for example a heavy Higgs boson, emerges from ggF and eventually decays, resulting in a pair of SM Higgs bosons (the heavy particle does not necessarily have to couple to the Higgs pair directly. Other virtual particles, that mediate between the heavy particle and the Higgs bosons, can also be considered in this decay). In the search

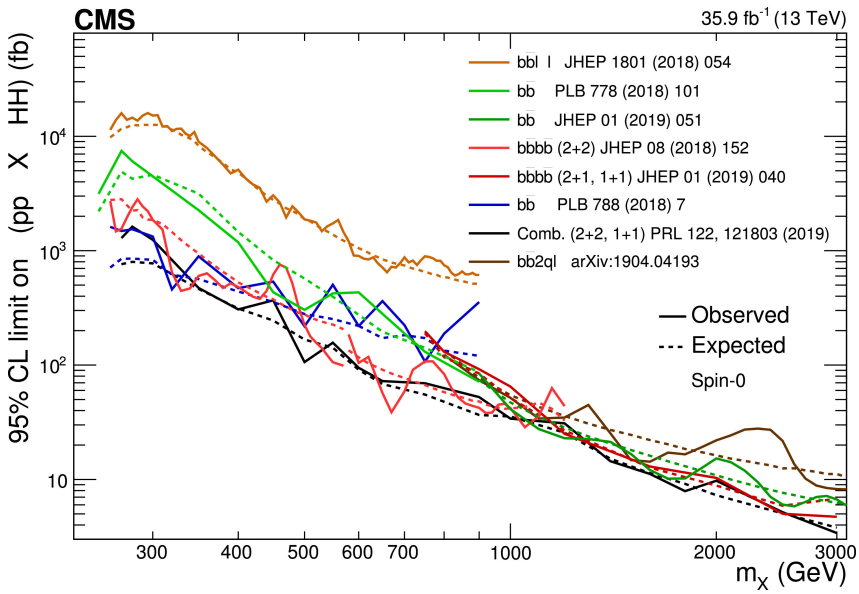


Figure 2.9: Current cross-section limits for resonant Higgs pair production as a function of resonance mass from CMS [20].

for resonant Higgs boson pair production presented in this thesis, only resonances with masses above 250 GeV are searched for since this is the threshold for on-shell Higgs boson pair production. Results from searches for resonant Higgs pair production are shown in Figure 2.8 and Figure 2.9. These searches set limits on the production cross-section of a spin-0 narrow-width resonance times the branching ratio for its decay into pairs of Higgs bosons.

3 Experimental Setup

Large particle accelerators are an essential and inevitable ingredient for high energy physics and the production of Higgs boson pairs. Their sizes are continuously growing, reaching even higher centre-of-mass energies and making much deeper investigations of particle interactions possible. In this Chapter, the Large Hadron Collider (LHC) will be introduced, which is currently the world's largest particle accelerator located in Geneva at the European Organization for Nuclear Research (CERN), as well as the ATLAS experiment on which this thesis is based upon.

3.1 The Large Hadron Collider

The LHC is a circular particle accelerator, more precisely a synchrotron, with a circumference of approximately 27 km. Its collider tunnels were built 50 m up to 175 m below the surface in order to protect its two beam pipes from cosmic radiation and to keep high energy radiation underground. In these two beam pipes, particles are accelerated in opposite directions up to velocities that are close to the speed of light. In order to achieve this, over a thousand dipole magnets force these particles onto circular trajectories and nearly 400 quadrupole magnets focus the beams such that the particle collisions at the four collision points occur in as small beam spots as possible. In addition, superfluid ^{4}He with a temperature of 1.9 K is used to cool these magnets down to their operating temperature. A depiction of the accelerator system is shown in Figure 3.1, where the four pre-accelerators (LEIR, BOOSTER, PS and SPS) are visible too. These pre-accelerators gradually accelerate the particle beams before they enter the LHC, where collisions take place.

Four large detectors are installed at various collision points of the LHC, each belonging to a different experiment: The ATLAS and CMS experiments are general-purpose experiments. In both experiments, various investigations are carried out, for instance a deep investigation of the Higgs boson, the search for physics beyond the SM, etc. The other two major experiments are LHCb, where precision measurements of B-hadrons are inves-

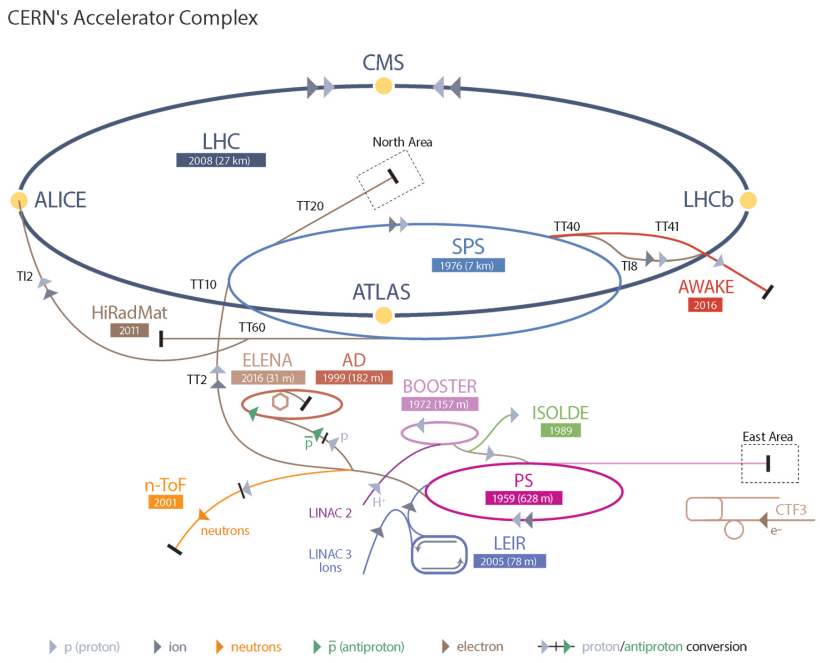


Figure 3.1: The LHC, all its pre-accelerator systems and the intersection points of the four major experiments (ATLAS, CMS, LHCb and ALICE) [33].

tigated, and ALICE, where quark-gluon plasmas are studied in detail.

Since its completion and first usage in 2010, two major data-taking periods took place at the LHC. The first one - Run 1 - began in 2010, where mainly protons with a centre-of-mass energy of 7 TeV were used in *colliding beam experiments* to collect data. In 2012 a centre-of-mass energy of 8 TeV was reached and in 2013 the LHC was shut down and installations began to prepare the LHC for higher centre-of-mass energy scales. Until then, 28 fb^{-1} of data had been taken in total (5.5 fb^{-1} at $\sqrt{s} = 7\,\mathrm{TeV}$ and 22.5 fb^{-1} at $\sqrt{s} = 8\,\mathrm{TeV}$). The second run - Run 2 - started in 2015 with a centre-of-mass energy of 13 TeV and ended in 2018, when the second long shutdown began. Over these four years, approximately 139 fb^{-1} of data had been taken. Today, an upgrade to a centre-of-mass energy of 14 TeV is in progress. The restart of the LHC is planned for the year 2021.

3.2 The ATLAS detector

The ATLAS (**A** **T**oroidal **L**HC **A**pparatu**S**) experiment is one of the major four experiments at the LHC at CERN. It is a multi-purpose experiment, meaning that the focus of this experiment spans several fields of study of particle physics. The detector was built

in such a way that it covers as much space as possible around the beam pipes and the collision point while being able to detect all kinds of particles in colliding beam experiments. Furthermore, the ATLAS detector has a cylindrical shape and consists of different concentric detector layers as well as various types of end-caps covering the ends of this cylindrically shaped apparatus (see Figure 3.2).

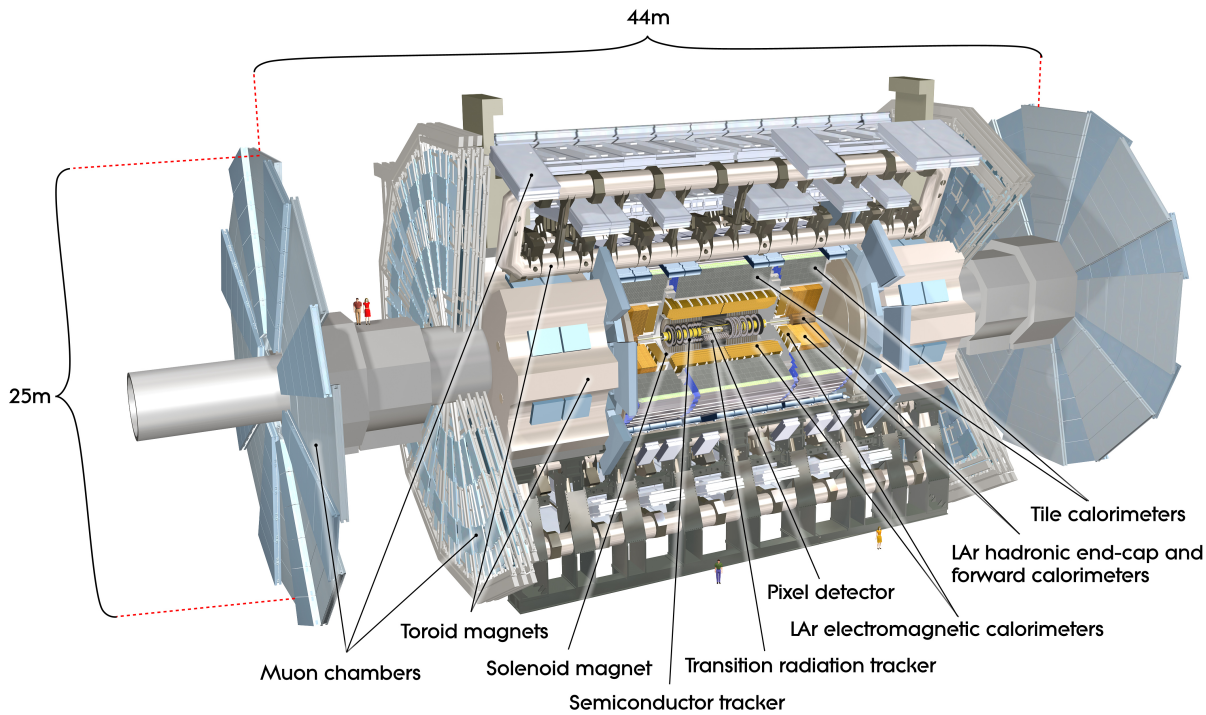


Figure 3.2: The structure of the ATLAS detector at the LHC at CERN. All parts of the inner detector, the calorimeters, the muon chamber, and the magnet system are labelled in this picture [34].

ATLAS uses right-handed cylindrical coordinates in order to describe the area around the intersection point of the beam pipes. In this case, the point of collision serves as the origin of the chosen coordinate system and the beam line as the z-axis. The x-axis points in the direction of the centre of the LHC ring and the y-axis points straight upwards. Furthermore, ATLAS uses two angles to describe this setup: ϕ is the azimuthal angle and θ is the polar angle that is measured with respect to the beamline. From this setup, two important quantities can be defined: the *pseudorapidity*, denoted as η, which is defined as

$$\eta := \ln\left(\tan\left(\frac{\theta}{2}\right)\right) , \tag{3.1}$$

and the *rapidity* y:

$$y := \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) . \tag{3.2}$$

Furthermore, a third variable can be defined from these two quantities, which is a measure for the distance between the momentum vectors of two particles in the detector:

$$\Delta R := \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} . \tag{3.3}$$

The difference between two rapidity values is invariant under Lorentz-boosts in the direction of the beamline (z-direction). The difference of two pseudorapidity values, however, is only approximately Lorentz-invariant under boosts in the z-direction [35]. Other important physical quantities are the *transverse momentum*

$$p_T = \sqrt{p_x^2 + p_y^2} , \tag{3.4}$$

the *transverse energy*

$$E_T = E\sin(\theta) , \tag{3.5}$$

and the *transverse mass*

$$m_T = E^2 - p_z^2 , \tag{3.6}$$

where E is the total energy of the particle, m its invariant mass, and p_i the component of the momentum parallel to the respective axis ($i = x, y, z$).

3.2.1 The Inner Detector

As already mentioned, the ATLAS detector is a combination of different detector subsystems, built to detect as many particles as possible and to measure them as precisely as possible. Therefore, a brief introduction into every subdetector system is given in the following.

The *Inner Detector* (ID) is located closest to the beamline. It has a radius of 1.2 m and a total length of 6.2 m and consists of three different particle trackers (see Figure 3.3). The innermost tracker, the *Pixel Detector*, is made of 1744 modules, each covering an area of $10\,\mathrm{cm}^2$ around the beamline and consisting of smaller readout channels (pixels), where each of them covers an area of $50 \times 400\,\mu\mathrm{m}^2$. There are over 80 million pixels in total in the Pixel Detector.

Outside the Pixel Detector is the *Semi-Conductor Tracker* (SCT) which is similarly struc-

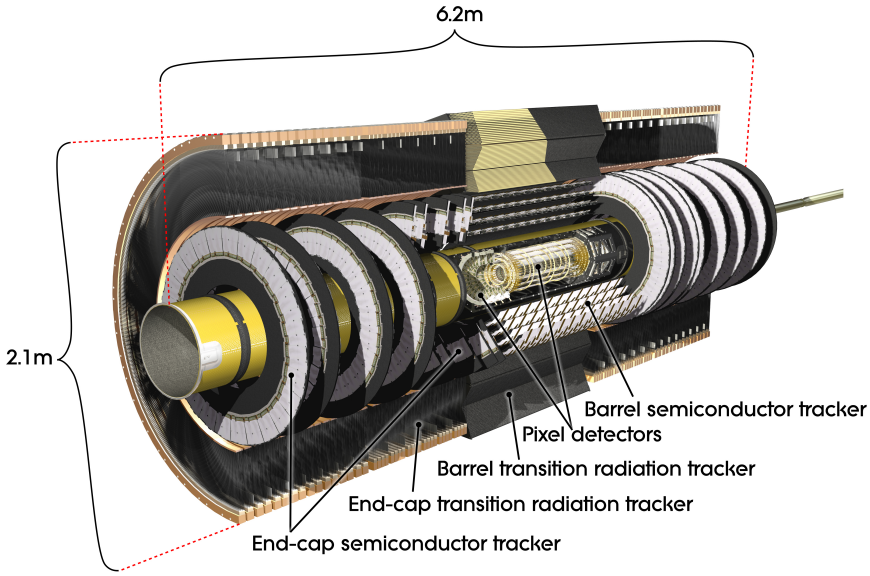


Figure 3.3: The ID of the ATLAS detector and all its subsystems [36].

tured to the Pixel Detector, except that thin strips of silicon instead of pixels are built within the tracker. It is divided into the barrel SCT and the end-cap SCT, which covers large values of pseudorapidity. In total, the SCT covers an area of $60\,\mathrm{m}^2$. The outermost tracker is the *Transition Radiation Tracker* (TRT; also divided into a barrel TRT and an end-cap TRT) which consists of small drift tubes, each filled with gas that becomes ionised by transition radiation, which is a type of electromagnetic radiation that is emitted by charged particles entering the tubes and passing through inhomogeneous media. The detached electrons from the ionised gas will travel towards a grid of wires due to an external electric potential, inducing measurable currents.

Surrounding the ID, a central solenoid produces a 2 T strong magnetic field that bends electrically charged particles, making it possible to measure their momenta. The η coverage and the relative p_T resolution of the ID are both listed in Table 3.1.

3.2.2 The Calorimeter

Two independent calorimeter systems are part of the ATLAS detector: the *electromagnetic calorimeter* (ECAL; used for detecting electromagnetically interacting particles) and the *hadronic calorimeter* (HCAL; used for detecting strongly interacting particles). A picture of the entire calorimeter system is shown in Figure 3.4. Both systems are *sampling* calorimeters, meaning that alternating layers of active material and absorbent material are part of their systems. The active material causes particle showers and the absorber

detects the incoming signals. Furthermore, both the ECAL and the HCAL are divided into subsystems: A large barrel surrounding the beamline and end-caps covering the bases of the barrel.

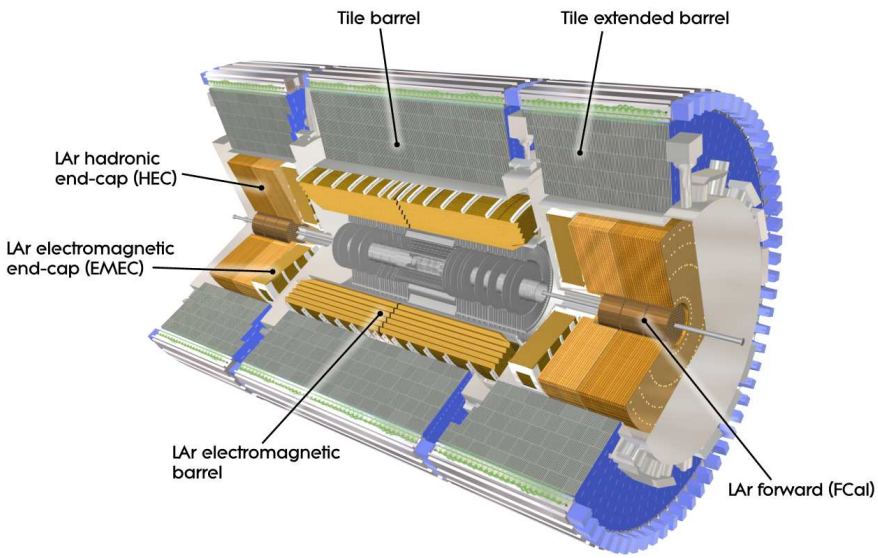


Figure 3.4: The ATLAS calorimeter system with all its subsystems (ECAL and HCAL), used to detect electromagnetically and hadronically interacting particles [34].

The innermost component is the LAr electromagnetic barrel which is wrapped around the magnetic solenoid of the ID. It uses liquid argon as active material (hence "LAr") and lead as absorber. The LAr electromagnetic end-caps (EMECs) cover large values of pseudorapidity at both ends of the electromagnetic barrel and also use a liquid-argon-lead combination as active/absorbent material. Around this, the tile (extended) barrel envelops the electromagnetic barrel. It is the part of the HCAL where scintillating plastic tiles are used as active material and steel as absorber. The hadronic end-caps (HECs) are located on top of the EMECs. They are LAr calorimeters too, but use copper instead of lead as absorbent material. Finally, a LAr *Forward Calorimeter* (FCAL) is located within each end-cap in order to cover very high values of pseudorapidity. Each FCAL consists of three modules: The first one, made of copper, is used for electromagnetic measurements, the second and third one, made of tungsten, are used for hadronic detections.

Since argon has a boiling temperature of 87.15 K, it has to be cooled down continuously. For that reason, every part of the LAr calorimeter is housed within its own cryostat, constantly ensuring that the argon stays in its liquid phase (the ECAL and the end-caps

including the FCAL each have their own cryostat). Table 3.1 shows the relative energy resolution and the pseudorapidity coverage for both the ECAL and the HCAL. Altogether, the calorimeter has a diameter of 8 m and a total length of 12 m along the beam axis.

3.2.3 The Muon Spectrometer and the Toroid Magnets

The *muon spectrometer* (or *muon chamber*) is used for muon detection. It is the outermost layer of the ATLAS detector, has a similar functionality to the tracking chambers in the ID, begins at a radius of 4.25 m, and extends out to 22 m in diameter, making it the largest part of the entire ATLAS detector. It incorporates four different subsystems: First of all, there are *Thin Gap Chambers* located at each end of the detector which are used for triggering and second coordinate measurements (the second coordinate points in the the non-bending direction of the muon trajectory). The *Resistive Plate Chambers* fulfil the same purpose in the central region of the detector. Furthermore, there are *Monitored Drift Tubes* filled with carbon dioxide and gaseous argon and used for track and momentum measurements. Last but not least, *Cathode Strip Chambers* make high precision measurements of the muon coordinates at the ends of the detector.

A large air-core barrel toroid magnet and two smaller end-cap magnets are used for bending the muon tracks (for $|\eta| < 1.4$, the barrel toroid is used, for $1.4 < |\eta| < 1.6$ a combination of barrel and end-cap fields, and for $1.6 < |\eta| < 2.7$, only the end-cap magnets are used). Deflected by the Lorentz force, charged muons will travel in circular trajectories, making momentum measurements possible. The relative resolution and the η coverage of the muon spectrometer are shown in Table 3.1.

Table 3.1: The relative resolution for transverse momentum and energy as well as the η coverage of each ATLAS detector component [34]. E and p_T are both measured in GeV.

Detector component	**Required resolution**	**η coverage**	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\,\%\, p_T \oplus 1\,\%$	± 2.5	± 2.5
EM calorimetry	$\sigma_E/E = 10\,\%/\sqrt{E} \oplus 0.7\,\%$	± 3.2	± 2.5
Hadronic calorimetry (jets)			
barrel and end-cap	$\sigma_E/E = 50\,\%/\sqrt{E} \oplus 3\,\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\,\%/\sqrt{E} \oplus 10\,\%$	$3.1 < \|\eta\| < 4.9$	$3.1 < \|\eta\| < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\,\%$ at $p_T = 1\,\text{TeV}$	± 2.7	± 2.4

In order to create magnetic fields with a strength of approximately 4 T, three large toroid

magnets are used in the ATLAS detector: one around each end-cap and one around the barrel. Each toroid consists of eight coils, that are symmetrically and radially arranged around the beamline. To make it possible that these magnets create fields of such a great strength, the barrel toroid coils are cooled down by eight cryostats (one for each coil) and each end-cap is housed in its own cryostat.

3.2.4 The Trigger System

The last important part of the ATLAS detector is the trigger system. It is divided into two subtrigger systems: The *Level-1 hardware trigger* (L1) and the software based *High-Level Trigger* (HLT). L1 searches for leptons, photons, neutrinos (i.e. missing transverse energy), and jets, which all must have high transverse momenta. In addition, L1 defines Regions-of-Interest (RoI) with respect to ϕ and η where it identified interesting events. It decides within 2.5 µs or less after the event occurred whether it is worth keeping for further analysis. This makes it possible that the L1 can save up to 10^5 events per second. The data is then sent to the HLT where it is refined (based of the given RoIs and their information). It performs a detailed analysis in two different ways: Either by making use of data from small and isolated regions of the detector or by examining the whole event for selected detector layers. As a result, the HLT selects events with a rate of 1000 events per second which are then stored and saved for later offline analyses.

4 Higgs Pair Production in the $bb\tau\tau$ final state

The following chapter introduces the $HH \to b\bar{b}\tau\tau$ decay channel in Higgs boson pair production. First of all, the decay channels of the Higgs boson as well as previous results from the ATLAS experiment will be explained. After that, event selection and search strategies will be outlined, followed by a presentation of kinematic distributions corresponding to the $b\bar{b}\tau\tau$ decay channel of non-resonant Higgs boson pair production.

4.1 Previous Search Results at ATLAS

The Higgs boson has a mass of around 125 GeV and a predicted lifetime of $\mathcal{O}(10^{-22})$, thus it must be detected through its decay products. The branching ratios (BRs) for different decay channels of the Higgs boson are shown in Figure 4.1 as a function of the Higgs mass. Since the Higgs boson has a specific and experimentally verified mass of around 125 GeV, one can obtain the BRs from this diagram for this particular value. They are listed in Table 4.1. The BRs of the most common di-Higgs decay modes are shown in Figure 4.2.

Table 4.1: The most common SM Higgs boson decay channels and their corresponding BRs. In this table, the Higgs mass is assumed to be equal to 125.1 GeV [16].

Decay channel	Branching ratio
$H \to b\bar{b}$	$5.75 \times 10^{-1}{}^{+3.2}_{-3.3}$
$H \to \tau^+\tau^-$	$6.30 \times 10^{-2}{}^{+5.7}_{-5.6}$
$H \to \mu^+\mu^-$	$2.19 \times 10^{-4}{}^{+5.9}_{-5.9}$
$H \to gg$	$8.56 \times 10^{-2}{}^{+10.2}_{-10.0}$
$H \to \gamma\gamma$	$2.28 \times 10^{-3}{}^{+4.9}_{-4.9}$
$H \to ZZ^*$	$2.67 \times 10^{-2}{}^{+4.2}_{-4.2}$
$H \to WW^*$	$2.16 \times 10^{-1}{}^{+4.2}_{-4.2}$

The $HH \to b\bar{b}\tau\tau$ decay channel, where one Higgs boson decays into a bottom-antibottom

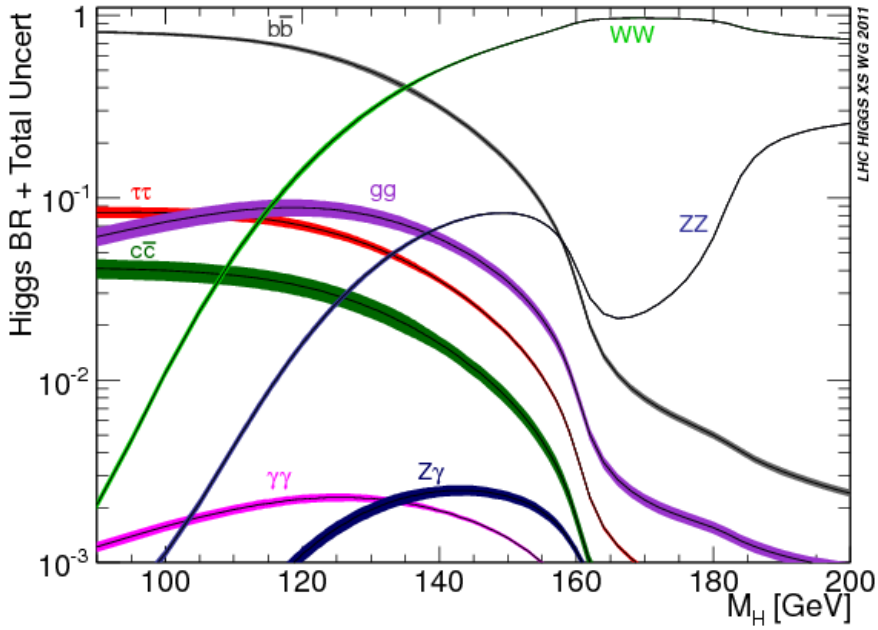


Figure 4.1: The BRs for different decay channels of the SM Higgs boson as a function of the Higgs mass [37].

	bb	WW	ττ	ZZ	γγ
bb	33%				
WW	25%	4.6%			
ττ	7.4%	2.5%	0.39%		
ZZ	3.1%	1.2%	0.34%	0.076%	
γγ	0.26%	0.10%	0.029%	0.013%	0.0005%

Figure 4.2: The BRs of the most common di-Higgs decay modes [17].

quark pair while the other one decays into a tau-antitau lepton pair, is considered in this thesis. It has a relatively large BR with 7.4 %. The final state bottom quarks of the di-Higgs decay can be detected as jets in a hadronic calorimeter.

The lifetime of both tau leptons is of $\mathcal{O}(10^{-15})$ and it is not possible to detect both of them directly in a detector (unlike electrons or muons) because each of them decays very quickly into a virtual W boson and a tau neutrino. Their decay products can be divided into three different decay modes.

For the $b\bar{b}\tau\tau$ final state, the first mode is the pure leptonic decay, where both W bosons (from the tau and the antitau lepton) decay into a lepton-neutrino pair ($WW \to \ell\nu\ell\nu$). Both charged leptons can easily be detected, unlike the neutrinos whose kinematics can only be measured through missing transverse energy E_T^{miss}.

The second decay mode is the pure hadronic decay of the two W bosons ($WW \to q\bar{q}q\bar{q}$). This implies four jets in the final state. The third possibility is a mixed leptonic-hadronic (or semi-leptonic) decay ($WW \to q\bar{q}\ell\nu$). The pure hadronic decay $\tau_{\text{had}}\tau_{\text{had}}$ has the largest BR with 41.9 % of all decay modes, followed by the semi-leptonic decays $\tau_e\tau_{\text{had}}$ and $\tau_\mu\tau_{\text{had}}$ both with 22.9 %. Consequently, the purely leptonic decays only have a total BR of 12.3 % [32] which is why they are less often considered. All BRs of the di-tau final state are summarised in Figure 4.3.

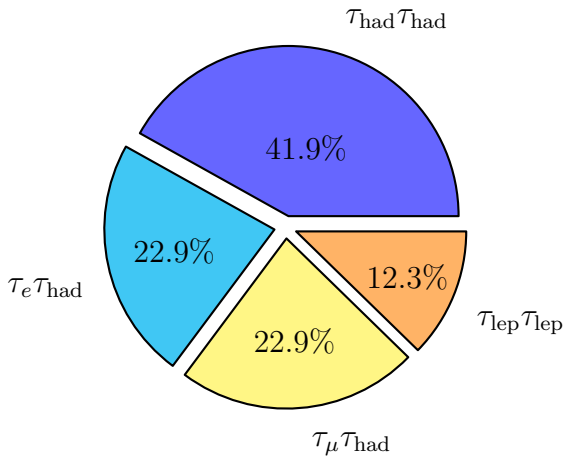


Figure 4.3: The BRs of the di-tau decay channels.

Compared to decay channels with higher BRs such as $HH \to b\bar{b}b\bar{b}$ or $HH \to b\bar{b}WW^*$, the $HH \to b\bar{b}\tau\tau$ decay represents a cleaner final state with less background. This is a very important feature in the search for non-resonant Higgs boson pair production and the measurement of the Higgs self-coupling strength as well as in the search for resonant

Higgs boson pair production and heavy resonances.

There have already been searches using ATLAS data from Run 2 that set limits on the production cross-section for the SM di-Higgs in the $b\bar{b}\tau\tau$ channel, considering the $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ decay modes. Table 4.2 summarises these results. Table 4.3 lists the corresponding results for several other channels.

Table 4.2: Observed and expected upper limits on the cross-sections times BR of non-resonant Higgs boson pair production in the $HH \to b\bar{b}\tau\tau$ decay channel at $\sqrt{s} = 13\,\text{TeV}$ [32]. $\pm 1\sigma$ represents the error of each observed limit.

		Observed	-1σ	Expected	$+1\sigma$
$\tau_{\text{lep}}\tau_{\text{had}}$	$\sigma(HH \to b\bar{b}\tau\tau)$ [fb]	57	49.9	69	96
	σ/σ_{SM}	23.5	20.5	28.4	39.5
$\tau_{\text{had}}\tau_{\text{had}}$	$\sigma(HH \to b\bar{b}\tau\tau)$ [fb]	40.0	30.6	42.4	59
	σ/σ_{SM}	16.4	12.5	17.4	24.2
Combination	$\sigma(HH \to b\bar{b}\tau\tau)$ [fb]	30.9	26.0	36.1	50
	σ/σ_{SM}	12.7	10.7	14.8	20.6

Table 4.3: Observed and expected upper limits on the cross-sections times BR for non-resonant Higgs boson pair production at $\sqrt{s} = 13\,\text{TeV}$ [17, 18, 38–40].

Channel	Observed [fb]	Expected [fb]
$\gamma\gamma b\bar{b}$	730	930
$b\bar{b}b\bar{b}$	110	1220
$\gamma\gamma WW^*$	730	930
$b\bar{b}\ell\nu\ell\nu$	1200	900

4.2 Search for $HH \to bb\tau\tau$ with the Full Run-2-Dataset

The following section provides a description of the search for $HH \to b\bar{b}\tau\tau$ decays with the full Run-2-dataset corresponding to $139\,\text{fb}^{-1}$. Here, the focus lies on the $\tau_{\text{had}}\tau_{\text{had}}$ decay channel. Its event selection will be outlined followed by a presentation of kinematic distributions.

4.2.1 Event Selection and Search Strategy

To find events that are suitable for an analysis of the $b\bar{b}\tau_{\text{had}}\tau_{\text{had}}$ decay channel, one has to apply event selection criteria. These requirements select events that are consistent with the $b\bar{b}\tau\tau$ topology.

$\tau_{\text{had}}\tau_{\text{had}}$ events first have to pass the following trigger selection [32, 41]: The first trigger applied is the *single tau trigger* (STT). Events in which the leading tau lepton (the tau lepton candidate with the higher transverse momentum) has a transverse momentum larger than $80-160\,\text{GeV}$ and fulfils the trigger tau ID criteria automatically pass the STT and become candidates for further event selection (the exact p_T-threshold for tau leptons depends on the data-taking period in which the event has been detected). Otherwise, they are checked by the *di-tau trigger* (DTT). In order to successfully pass the DTT, the leading tau lepton has to have $p_T > 35\,\text{GeV}$ and must pass the trigger ID criteria. The same ID requirements also apply for the sub-leading tau lepton (the tau lepton with the second highest transverse momentum), but its transverse momentum must only exceed $25\,\text{GeV}$. Furthermore, events checked by the DTT must also have an additional jet with $p_T > 25\,\text{GeV}$ if they were detected in 2015 or 2016. For the years 2017 and 2018, two versions of the DTT are used: One containing an additional jet with $p_T > 25\,\text{GeV}$ and $\Delta R_{\tau\tau} < 2.7$ and another with two additional jets that satisfy $p_T > 12\,\text{GeV}$. Jets originating from bottom quarks are identified using a multivariate discriminant that is trained against a background sample with $10\,\%$ jets that were initiated by charm quarks. The working point is chosen such that the average b-tagging efficiency is equal to $70\,\%$.

After passing the trigger requirements, events are subject to further selection criteria. Both tau lepton candidates must have opposite electric charge and satisfy $m_{\tau\tau} > 60\,\text{GeV}$ ($m_{\tau\tau}$ is reconstructed by using the *Missing Mass Calculator* or MMC [42]). Events containing an electron or muon candidate are discarded to make sure only those events with purely hadronic tau lepton decays are selected. Then, a distinction between STT and DTT events is made: STT leading tau leptons must pass loose tau lepton ID requirements and must satisfy $p_T > 100\,\text{GeV}$, $140\,\text{GeV}$ or $180\,\text{GeV}$, depending on the data-taking time period thresholds. Sub-leading tau lepton candidates, on the contrary, only must satisfy $p_T > 20\,\text{GeV}$. Every event is required have exactly two tau lepton candidates with $|\eta| < 2.5$. Also, every event requires at least two jets with $p_T > 45$ (20) GeV for the leading (sub-leading) jet.

The pseudorapidity requirements as well as the loose ID criteria for both tau lepton candi-

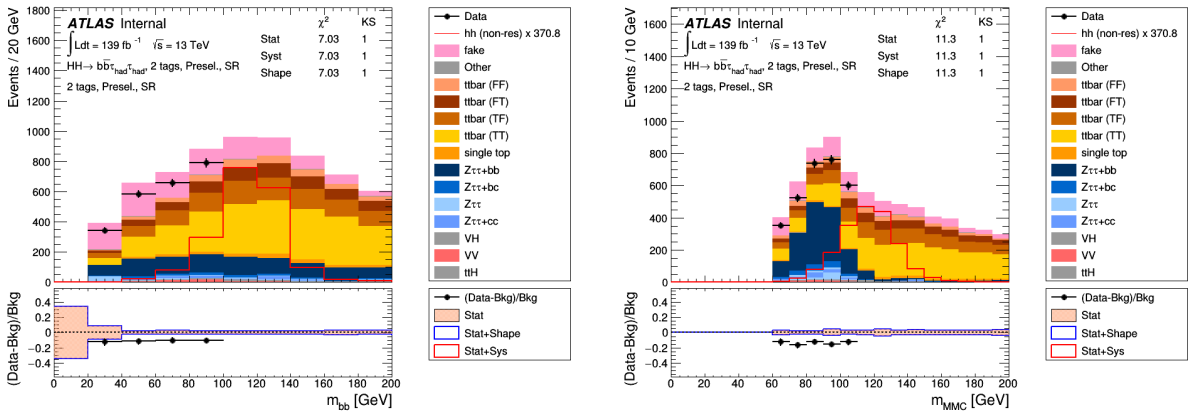
dates are the same for DTT events, but with $p_T^{\text{leading }\tau} > 40\,\text{GeV}$ and $p_T^{\text{sub-leading }\tau} > 30\,\text{GeV}$. Lastly, at least two jets with $p_T > 80$ (20) GeV for the leading (sub-leading) jet are required to pass the DTT event selection for data from 2015 and 2016. Two additional jets with $p_T > 45$ (45) GeV or $p_T > 80$ (20) GeV and $\Delta R_{\tau\tau} < 2.5$ are required for the 2017 and 2018 data-taking period, depending on the version of the DTT that the event passed.

4.2.2 Kinematic Distributions

After the event selection, different physical quantities (e.g. invariant masses of particle systems, ΔR, $\Delta\phi$, etc.) are exploited to further separate the signal process from background. A few examples are shown in Figure 4.4 for $HH \to b\bar{b}\tau_{\text{had}}\tau_{\text{had}}$. These histograms display event distributions in the signal region with two b-tagged jets (2-b-tag SR) of: invariant masses for the di-b-jet system (Figure 4.4(a)), the di-tau system (Figure 4.4(b)), and the di-Higgs system (Figure 4.4(c)). They correspond to the non-resonant analysis and their signal distributions represent the non-resonant Higgs boson pair production signal (80 % signal blinding was applied while modelling these distributions). It is clearly visible that the invariant mass distributions for the di-tau and the di-b-jet systems peak at a value slightly less than 125 GeV for the signal process. The invariant di-Higgs mass distribution peaks at approximately 400 GeV to 450 GeV.

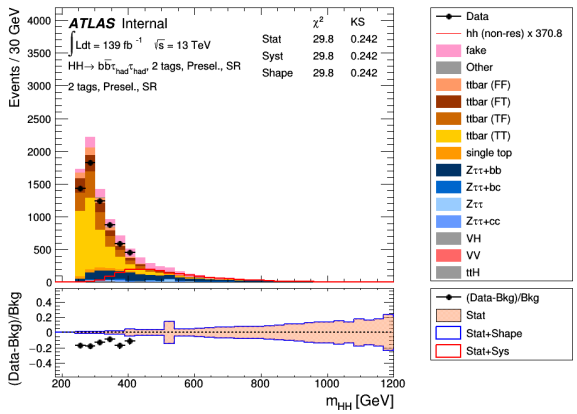
In addition to the number of signal events in Figure 4.4, there are also many background processes that can produce $b\bar{b}\tau_{\text{had}}\tau_{\text{had}}$ final states. The most common background processes are described below.

Firstly, fake-tau leptons are a major background in the $b\bar{b}\tau_{\text{had}}\tau_{\text{had}}$ decay channel. These originate from multijet events that emerge from QCD background processes with misidentified tau lepton candidates. A data-driven approach in a background enriched same sign control region (SS CR) is used to estimate this background, where the "same sign" refers to the fact that each event in this CR contains two reconstructed tau leptons whose electric charges have the same sign. This region is used to obtain an estimation for the amount of fake-tau leptons in the opposite sign signal region (OS SR). For this, one has to find the total number of reconstructed tau leptons within the SS CR. This is done by simulating events with Monte-Carlo simulations (MC) and then by "subtracting" these samples from the data. Since multijets cannot be simulated with MC, this subtraction will yield the number of multijets/fake-tau leptons in the data. This number is then split into two smaller sets (one contains tau leptons that passed the ID criteria and one those that failed them). A *fake factor* (FF) can then be calculated by dividing the number of



(a)

(b)



(c)

Figure 4.4: The invariant masses of the di-Higgs, the di-b-jet and the di-tau systems for non-resonant Higgs boson pair production. Note, that the total area covered by the thin, red line (marking signal events) has been multiplied with a scaling factor such that it equals 20 % of the area covered by background events. The main background processes are fake-tau leptons (pink), $t\bar{t}$ and single top production (orange, yellow, and brown), $Z \to \tau\tau$+jets (dark and light blue), double vector boson production (pale red), single Higgs boson production with either a $t\bar{t}$ pair or a vector boson (light grey), and negligible processes summarised as "Other" (dark grey). Furthermore, 80 % signal blinding was used while modelling these histograms.

passed tau leptons by the number of those that failed:

$$FF = \frac{N^{\text{SS, CR}}_{\text{pass ID}}}{N^{\text{SS, CR}}_{\text{fail ID}}} \,. \tag{4.1}$$

It is important to mention that this division is done in bins of tau lepton p_T. A multiplication of this FF with the number of failed tau leptons in the OS SR then yields

$$N_{\text{fake}}^{\text{OS, SR}} = FF \cdot N_{\text{fail ID}}^{\text{OS, SR}} , \tag{4.2}$$

which is a measure for the amount of fake-tau leptons in the SR.

Another major background in the $b\bar{b}\tau_{\text{had}}\tau_{\text{had}}$ channel are $t\bar{t}$ and single top production. $t\bar{t}$ production is divided into four categories: one where the leading and the sub-leading tau leptons are both real-tau leptons (or "true-tau leptons", hence the "TT" in Figure 4.4), one where they are both fake ("FF", not the same "FF" as for fake factor), one where the leading tau lepton is fake and the sub-leading is real ("FT"), and vice versa ("TF"). These processes are estimated using MC simulation.

The last large background contributors are the production of a single Z boson, which can eventually decay into a tau-antitau lepton pair and also creates b- and/or c-quark jets, double vector boson production, and the single Higgs boson production in combination with either a vector boson or a $t\bar{t}$ pair. All other background processes are smaller and they are summarised as "Other".

Lastly, further distributions for important variables are shown in the following figures because these variables (including those in Figure 4.4) provide the best separation between signal and background events. The signal shown in these figures also corresponds to the non-resonant SM di-Higgs production. Figure 4.5 shows azimuthal angles $\Delta\phi$ between the bottom quarks, the tau leptons, and the tau neutrinos of the di-Higgs decay, as well as Figures 4.6(a) and 4.6(b). The other subfigures of Figure 4.6 show distances ΔR between the same particles. The position in ϕ of the missing transverse energy E_T^{miss} with respect to the tau lepton decay products ($E_T^{\text{miss}}\phi$ centrality) and the "stransverse" mass M_{T2}^W as well as the transverse mass M_T^W between the leading tau lepton and E_T^{miss}, the transverse momenta of the leading and sub-leading tau leptons ($p_T^{\tau_0}$ and $p_T^{\tau_1}$), and the radius $X_{\text{m,bb}\tau\tau}$ of a "circle" within the di-tau and di-b-jet mass plane, in which the majority of signal events are located (this radius goes to 0 for better separation power), are shown in Figure 4.7. The mathematical definitions of $E_T^{\text{miss}}\phi$ centrality, M_{T2}^W, M_T^W, and $X_{\text{m,bb}\tau\tau}$ are as follows:

$$E_T^{\text{miss}}\phi \text{ centrality} = \frac{A+B}{\sqrt{A^2+B^2}} \tag{4.3}$$

with

$$A = \frac{\sin(\phi_{E_T^{\text{miss}}} - \phi_{\tau_2})}{\sin(\phi_{\tau_1} - \phi_{\tau_2})} \tag{4.4}$$

$$B = \frac{\sin(\phi_{\tau_1} - \phi_{E_T^{\text{miss}}})}{\sin(\phi_{\tau_1} - \phi_{\tau_2})}\,, \tag{4.5}$$

$$M_{T2}^{W} = \min_{p_T^{\tau_0} + p_T^{\tau_1} = p_T^{\text{miss}}} \{\max(m_T^0, m_T^1)\} \tag{4.6}$$

with

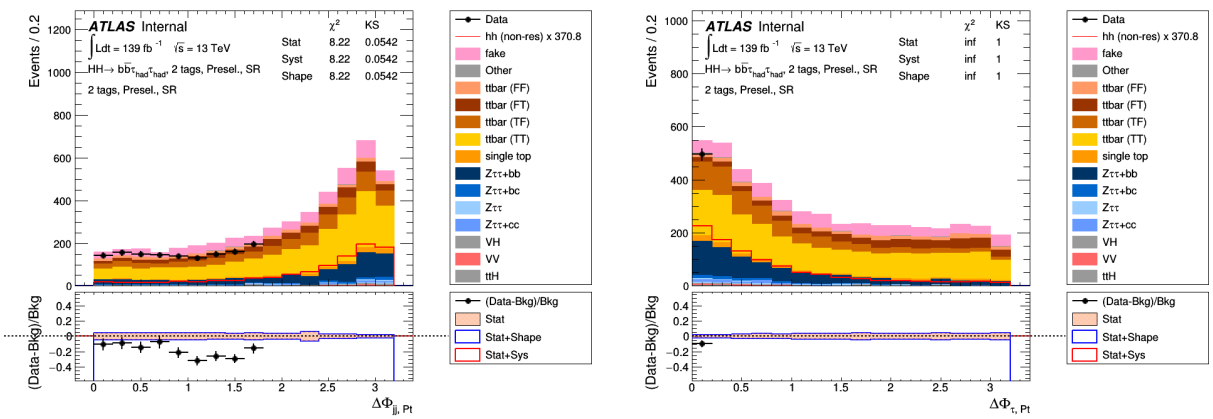
$$m_T = \sqrt{m_b^2 + m_\tau^2 + 2(E_T^b E_T^\tau - p_T^b p_T^\tau)}\,, \tag{4.7}$$

$$m_T^W = \sqrt{2 p_T^\ell E_T^{\text{miss}} (1 - \cos(\Delta\phi))}\,, \tag{4.8}$$

and

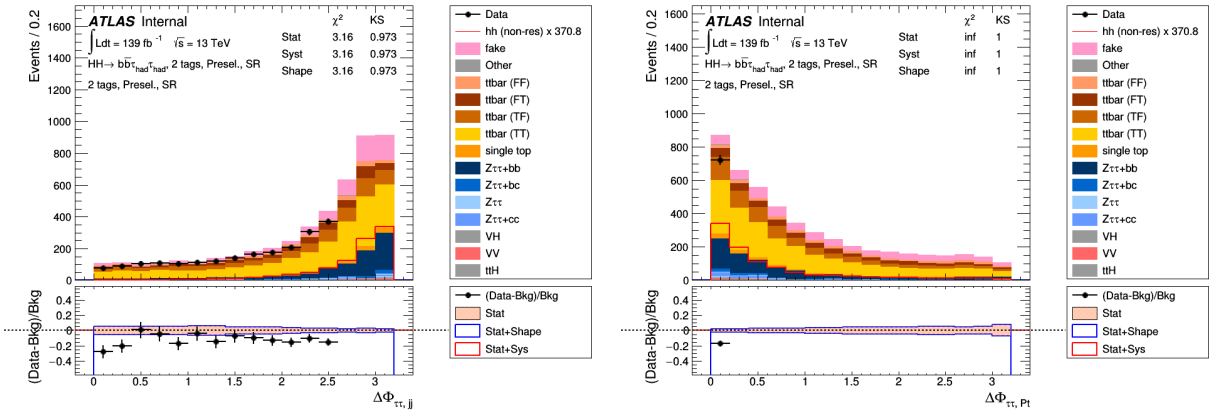
$$X_{\text{m,bb}\tau\tau} = \sqrt{\left(\frac{m_{\tau\tau} - 116\,\text{GeV}}{0.15 \times m_{\tau\tau}}\right)^2 + \left(\frac{m_{bb} - 112\,\text{GeV}}{0.15 \times m_{bb}}\right)^2}\,. \tag{4.9}$$

Finally, it is worth mentioning that all of these distributions (Figure 4.4 to 4.7) are shown after the preselection criteria were applied and that all kinematic variables and brief explanations of them are provided in Table 4.4.



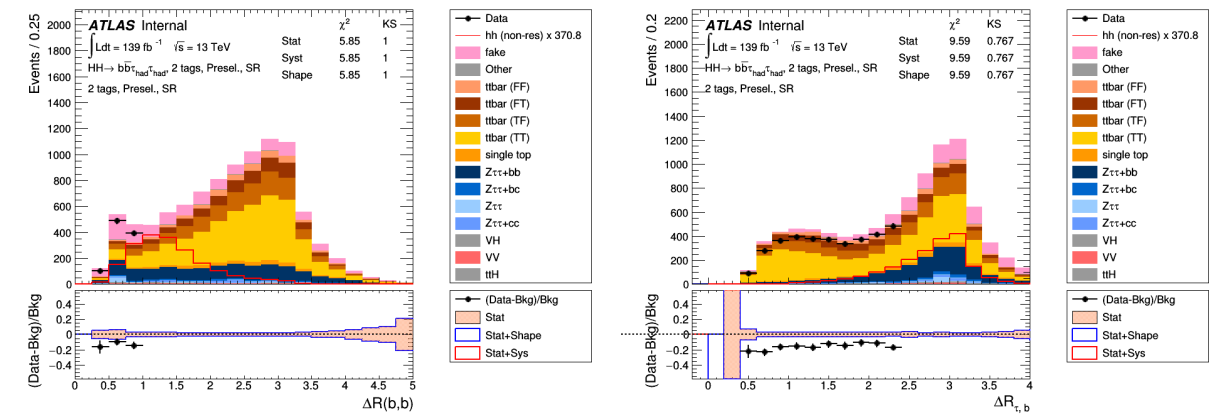
(a) (b)

Figure 4.5: The event distributions for $\Delta\phi_{\text{jj},p_T}$ and $\Delta\phi_{\tau,p_T}$. The red curves show the signal of the non-resonant Higgs boson pair production. Their scaling factor was chosen such that the area under each red curve is equal to 20 % of the total area covered by all background events. In addition, 80 % signal blinding was used while modelling these histograms.



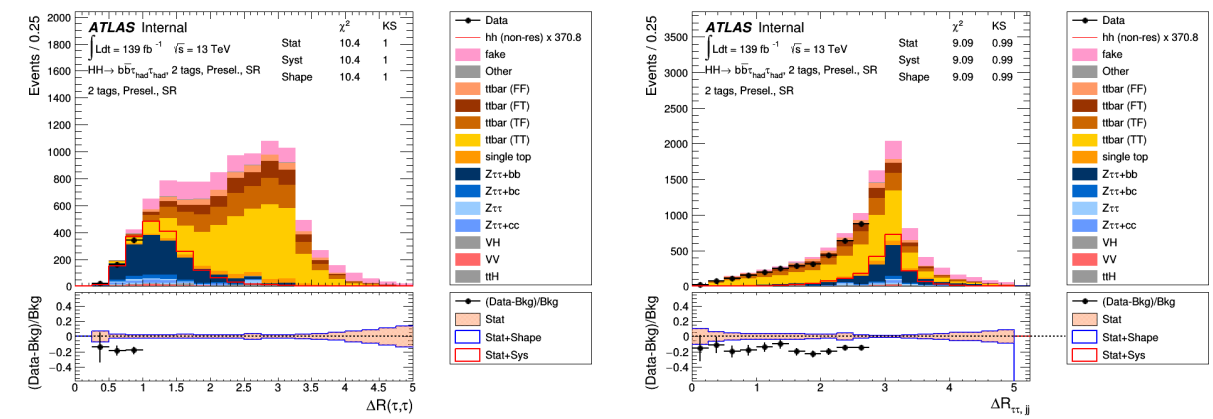
(a)

(b)



(c)

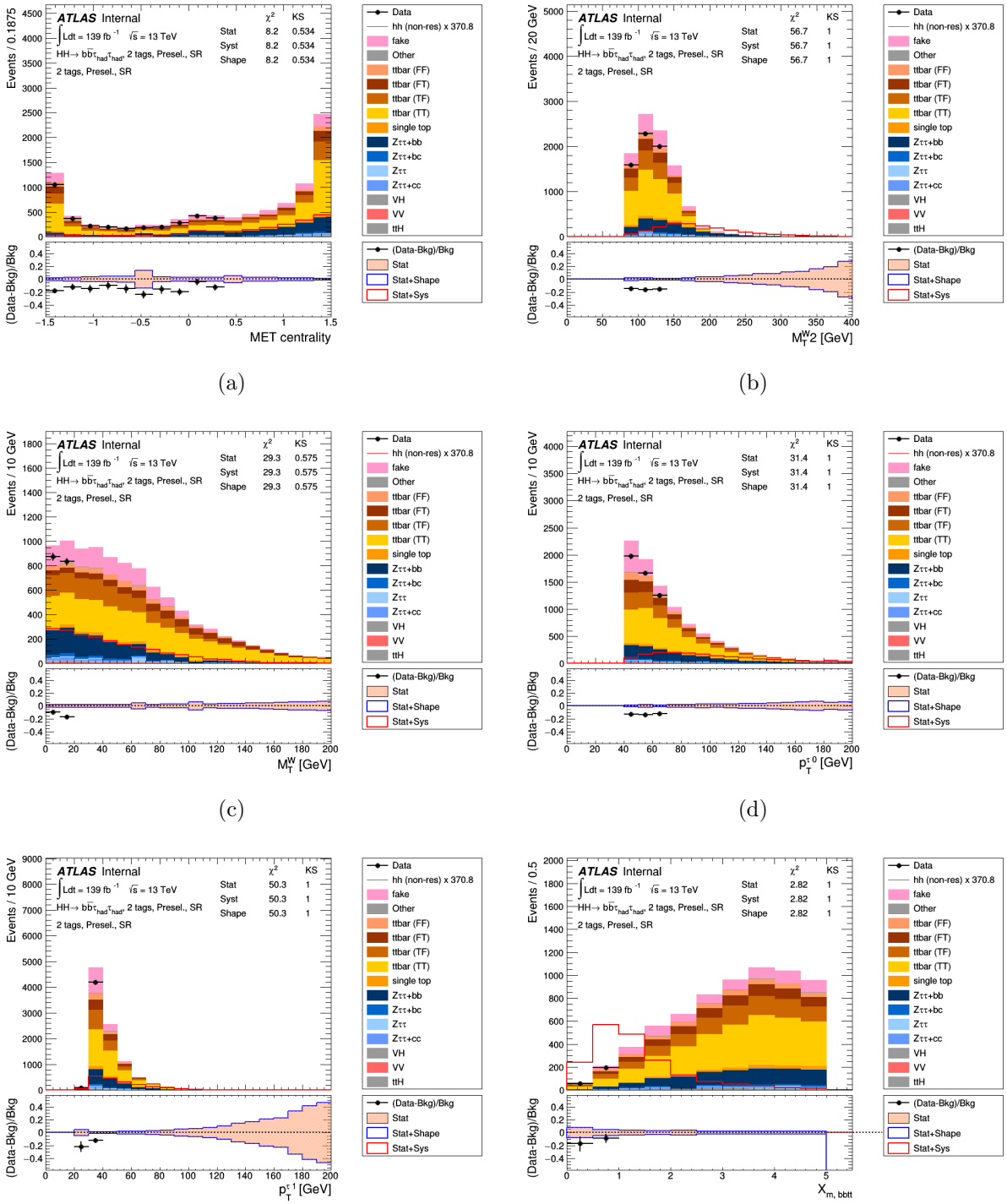
(d)



(e)

(f)

Figure 4.6: The event distributions for $\Delta\phi_{\tau\tau,\mathrm{jj}}$, $\Delta\phi_{\tau\tau,p_T}$, ΔR_{bb}, $\Delta R_{\tau,\mathrm{b}}$, $\Delta R_{\tau\tau}$, and $\Delta R_{\tau\tau,\mathrm{jj}}$. The red curves show the signal of the non-resonant Higgs boson pair production. Their scaling factor was chosen such that the area under each red curve is equal to 20 % of the total area covered by all background events. In addition, 80 % signal blinding was used while modelling these histograms.



(a) (b) (c) (d) (e) (f)

Figure 4.7: The event distributions for $E_T^{\text{miss}}\phi$ centrality, M_{T2}^W, M_T^W, $p_T^{\tau_0}$, $p_T^{\tau_1}$, and $X_{\text{m,bb}\tau\tau}$. The red curves show the signal of the non-resonant Higgs boson pair production. Their scaling factor was chosen such that the area under each red curve is equal to 20 % of the total area covered by all background events. In addition, 80 % signal blinding was used while modelling these histograms.

Table 4.4: Explanations of the kinematic variables used to model the previously shown kinematic distributions.

Variable	Explanation
$\Delta\Phi_{\mathrm{jj},p_T}$	angle between the di-b-jet system and $E_{\mathrm{T}}^{\mathrm{miss}}$
$\Delta\Phi_{\tau,p_T}$	angle between the leading tau lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$
$\Delta\Phi_{\tau\tau,\mathrm{jj}}$	angle between the di-tau and di-b-jet systems
$\Delta\Phi_{\tau\tau,p_T}$	angle between the di-tau system and $E_{\mathrm{T}}^{\mathrm{miss}}$
ΔR_{bb}	distance between the two b-jets
$\Delta R_{\tau,\mathrm{b}}$	distance between the leading tau lepton and the leading b-jet
$\Delta R_{\tau\tau}$	distance between the visible tau lepton decay products
$\Delta R_{\tau\tau,\mathrm{jj}}$	distance between the di-tau and di-b-jet systems
m_{bb}	invariant mass of the di-b-jet system
m_{HH}	invariant mass of the di-Higgs system
m_{MMC}	invariant mass of the di-tau system, calculated using the MMC
$E_{\mathrm{T}}^{\mathrm{miss}}\phi$ centrality	position in ϕ of the missing transverse energy with respect to the visible tau lepton decay products
$m_{\mathrm{T2}}^{\mathrm{W}}$	the largest mass of the parent particle that is compatible with the kinematic constraints which, in this case, is the top quark decaying into a bottom quark and a W boson (*stransverse mass*)
$m_{\mathrm{T}}^{\mathrm{W}}$	the transverse mass between the leading tau lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$
$p_{\mathrm{T}}^{\tau_0}$	the transverse momentum of the leading tau lepton
$p_{\mathrm{T}}^{\tau_1}$	the transverse momentum of the sub-leading tau lepton
$X_{\mathrm{m,bb}\tau\tau}$	radius of the "circle" in the m_{bb} and $m_{\tau\tau}$ mass plane

5 Signal Extraction using Boosted Decision Trees

The following chapter deals with signal extraction in the resonant and non-resonant analysis. The tool that is used for this are boosted decision trees. Thus, an introduction to BDTs is given in the next section. The usage of BDTs in the non-resonant and resonant analysis is then discussed. In the non-resonant analysis, two sets of BDT hyperparameters are compared. For the resonant analysis, a comparison between a BDT training on a single signal mass point and a training on three adjacent signal mass points is performed. The last section provides a summary of all results from both analyses.

5.1 Introduction to Boosted Decision Trees

A BDT is a binary structure that consists of many *decision trees* which are also binary structures. Decision trees consist of a *root node*, several *decision nodes*, and also *terminal nodes* or *leaves*: The root node is the starting point of the decision tree. From here, a whole set of events, used in an analysis, is split up into two smaller subsets in order to separate signal from background. This is done by simple conditions based on the values of kinematic variables that can either be true or false for a given event. Depending on the result, the event either becomes part of the first subset or the other. Each subset then reaches an individual decision node where the splitting is repeated based on the variable with the highest separation power. This process continues until a terminal node is reached, where a predefined number is assigned to each subset. Such a terminal node does not split the subset any further and represents the final classification into either signal or background events. There are many exit conditions that can end the separation process of a decision tree in order to reach a terminal node, for example a minimum number of events per node, a maximum depth of the tree, etc. A schematic diagram of a decision tree is shown in Figure 5.1.

A BDT is a collection of decision trees where events are separated many times by different

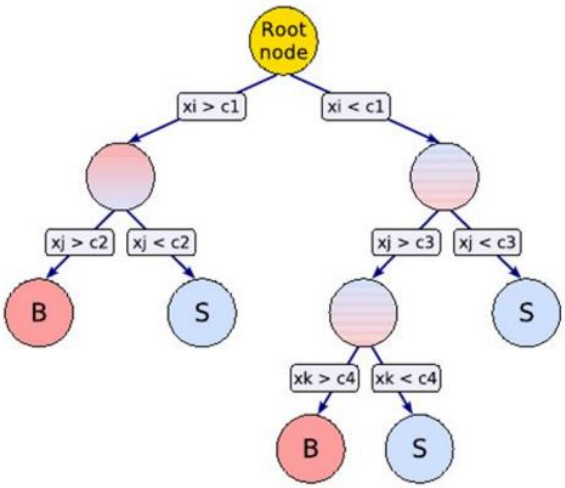


Figure 5.1: Schematic depiction of a decision tree [43].

decision trees: The first decision tree separates events based on the given variables and assigns weights to each event based on how well it could separate this specific event. The worse a decision tree can handle a particular event, the higher its event weight is going to be for the next decision tree, which is called *boosting*. The second tree then also separates events based on the same variables, but with the event weights from the former tree applied. The third tree does the same procedure with its new event weights from the former trees and so on. In the end, all decision trees are combined into a BDT that assigns a weight, known as the BDT score, to each event. High weights (close to 1) indicate signal-like events, while low BDT scores (close to -1) indicate background-like events. This method has the advantage that it will provide good signal-to-background separation results, even if one single tree does not perform well.

5.2 The BDT for the non-resonant analysis

In the non-resonant analysis, the same variables, whose kinematic distributions were shown in the previous chapter, were used to train a BDT because they make it possible to maximise the signal-to-background separation power of the BDT. In addition, two sets of hyperparameters were used to train a BDT (these sets are henceforth called "benchmark" and "tuned"). After the trainings, the performances of both BDTs were compared in order to find out which hyperparameters overtrain the BDT less and provide a better separation between the signal and background. A brief explanation of every used hyperparameter is given in the following. The exact preferences, that were used for each hyperparameter, are shown in Table 5.1.

Table 5.1: The BDT benchmark (obtained from [32]) and tuned hyperparameters (obtained from [44]) that were used for training. The benchmark hyperparameters were originally used as intial hyperparameters for the analysis in [32]. Later on, they were changed in an attempt to "tune" the BDT performance (hence *tuned*) and used in [44] to train a BDT.

Hyperparameter	Benchmark	Tuned
AdaBoostBeta	0.15	0.2
NTrees	200	250
MaxDepth	4	5
MinNodeSize	5 %	3 %
NCuts	100	100

Adaptive Boosting (AdaBoost) is the boosting algorithm that was used in this analysis. It aims at combining multiple weak classifiers to build one strong classifier by assigning high event weights to badly handled events and vice versa. Similarly, the better a classifier performs, the higher will be its assigned weight and vice versa. The next hyperparameter is AdaBoostBeta. In simplified terms, this quantity determines the *learning rate* of the AdaBoost algorithm.

The remaining hyperparameters are the *number of trees* (NTrees) used in the BDT, the *maximum number of decision nodes* an event can pass (MaxDepth), the *minimum number of events within a node* (MinNodeSize), and the *granularity* or the *number of grid points* of each variable range (NCuts). This last hyperparameter is used to set the size of the step in finding the cut value for a variable that has the highest separation gain [32].

In addition to these two sets of hyperparameters, the samples used to train the BDTs (signal and background) were both split into two subsets, too: One containing only events with even event numbers and one only containing those with odd event numbers (these are all simulated events). The BDT was trained separately on both of these sets and the resulting classifiers were then applied to their opposite counterpart of the dataset (the classifier that was trained on simulated even event numbers was applied to data events with odd event numbers and vice versa). Figure 5.2 shows the BDT distributions of each training (even/odd for benchmark/tuned hyperparameters) and Figure 5.3 shows the inverse background efficiency of each training as a function of the signal efficiency. At a signal efficiency of 65 %, the benchmark training yields inverse background efficiencies of 177 and 148 for even and odd event numbers, respectively, while the tuned training

provides 186 (even) and 170 (odd). These results show that both BDTs are consistent with each other.

Since the output of the BDT depends on the stopping condition and also on statistical fluctuations, it is possible to *overtrain* the BDT, which would lead to suboptimal results. Thus, it is important to check the degree of overtraining of BDTs. This is done by comparing the results of applying the BDT classifier on a training sample (e.g. events with odd event numbers) to the results of applying the BDTs on a testing sample (events with even event numbers). Based on such distributions, a Kolmogorov-Smirnov test (KS test) can be performed [45]. A KS value close to one means that both distributions are very similar, whereas a value close to zero implies a high probability that the distributions originate from different PDFs. Looking at Figure 5.2, the large KS values for background events show high similarities between the background distributions, whereas the KS values for signal events are rather small and imply larger differences between each pair of distributions.

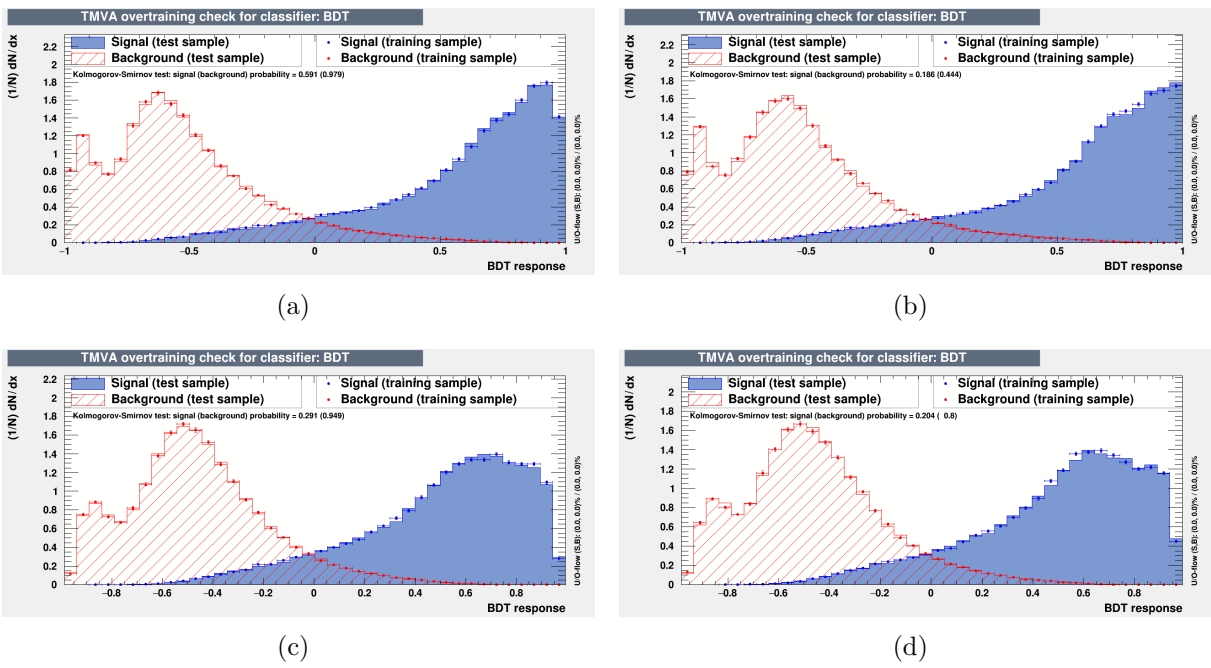
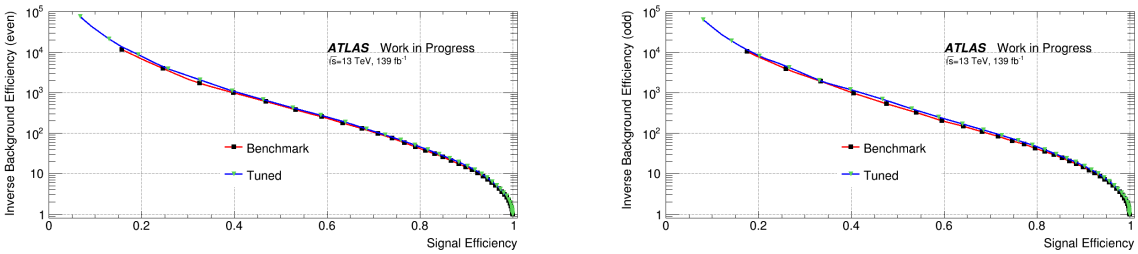


Figure 5.2: The BDT distributions for the benchmark (top row) and tuned (bottom row) hyperparameters as well as for even (left column) and odd (right column) event numbers. The KS values for signal and background events are shown in the top left corner of each subfigure (signal without parentheses and background with parentheses).



(a) (b)

Figure 5.3: The inverse background efficiencies for the even-event-number (left) and odd-even-number (right) classifier as a function of signal efficiency.

5.3 The BDT for the resonant analysis

5.3.1 Training on a single mass point

This resonant analysis aims to find signal events from the decay of an on-shell generic narrow-width scalar boson into two on-shell SM Higgs bosons. Since the mass of this hypothetical particle is unknown, the following set of 15 resonance mass points has been investigated: $m_X = \{251, 260, 280, 300, 325, 350, 400, 450, 500, 550, 600, 700, 800, 900, 1000\}$ GeV. To do so, a BDT was trained on each of these mass points for better signal-to-background separation (only the tuned hyperparameters were used in this analysis) and, similar to the non-resonant analysis, every training sample was split into two sets with either even or odd event numbers. The BDT was, again, trained on both of them and the classifiers for even events numbers were applied to odd event numbers and vice versa.

After the training was done, the inverse background efficiency was calculated at a signal efficiency of 65 %. This was done twice for every mass point: One time for the training on only even-event-number events and a second time on those with odd event numbers. The inverse background efficiency was then plotted as a function of resonance mass which is shown in Figure 5.4. A discussion of the performance follows in Section 5.4.

5.3.2 Training on adjacent mass points

A second BDT training method that was used during this resonant analysis was the training on adjacent mass points, or more specifically: One training sample incorporated the training samples of three neighbouring mass points. For both ends of the energy range, in which all mass points lie, only a mass point pair was used to train the BDT (251 GeV together with 260 GeV and 900 GeV together with 1000 GeV). This method was used

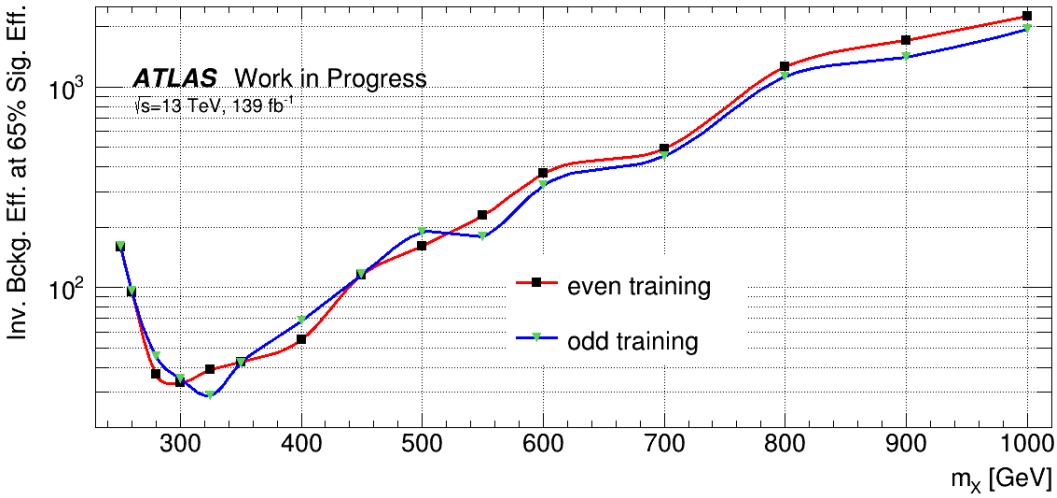


Figure 5.4: The inverse background efficiency at 65 % signal efficiency as a function of the resonance mass for the BDT trained on a single mass point, tuned hyperparameters, and even/odd event numbers, respectively.

because the reconstructed resonance mass variable, m_{HH}, is used in the training. Thus, it is ensured that the BDT is not overly sensitive to a resonance with a mass between some two tested values. In order to ensure the sensitivity to signals throughout the whole mass range, the training on adjacent mass points is performed, even though this reduces the performance of the BDT, evaluated exactly on the mass point on which it is trained.

The same methods and hyperparameters as mentioned before were used for this BDT training with training samples that were, again, split into even and odd event numbers. Furthermore, the inverse background efficiency was plotted, too, as a function of resonance mass for both sample types (Figure 5.5). The BDT distributions for this resonant training are shown in Figures 5.6 to 5.8 for each mass point.

5.4 Results

Figure 5.2 shows the BDT distributions of the non-resonant analysis and reveals that the BDT trained on benchmark hyperparameters as well as the BDT trained on tuned hyperparameters both show a good separation power for signal events and an even better one for background events. There is no indication of significant overtraining. Furthermore, Figure 5.3 shows the inverse background efficiencies for the training on even and odd event numbers as well as that the distributions of the inverse background efficiency as a

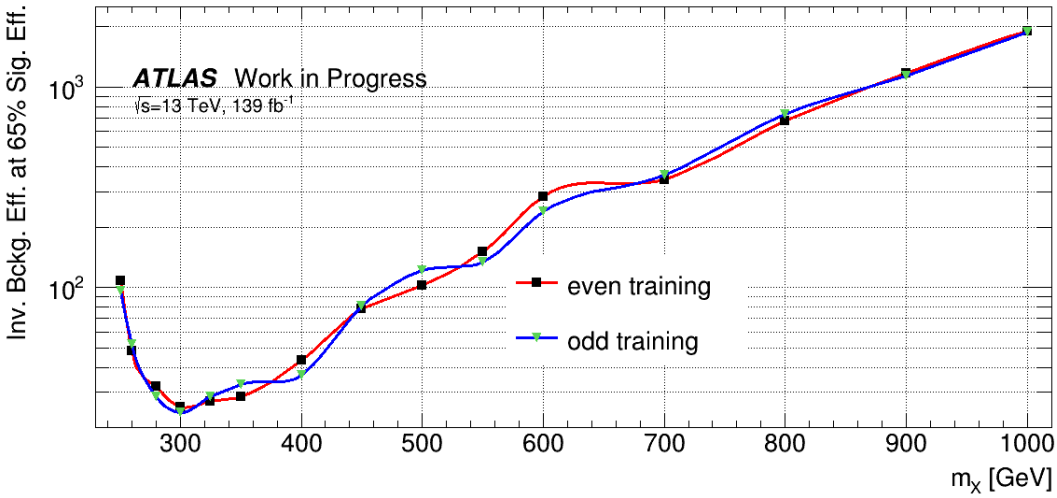
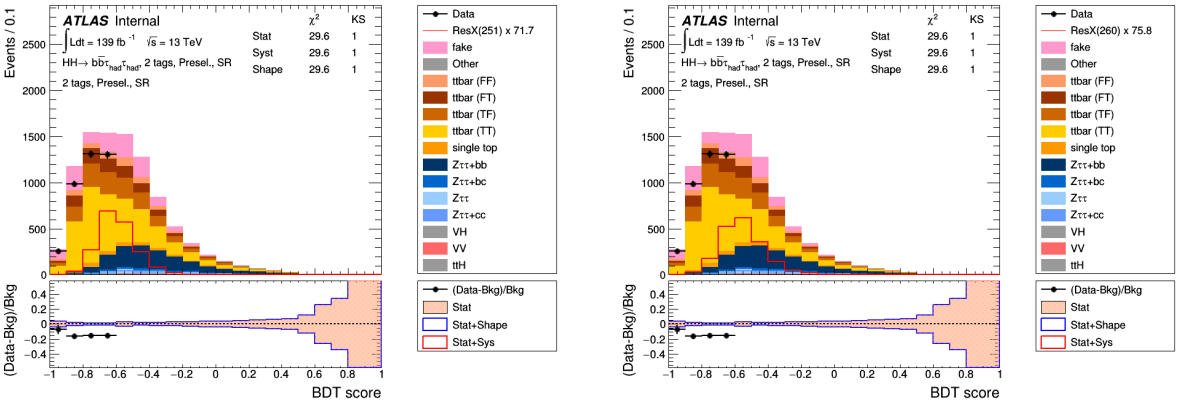


Figure 5.5: The inverse background efficiency at 65 % signal efficiency as a function of the resonance mass for the BDT trained on adjacent mass points, tuned hyperparameters, and even/odd event numbers, respectively.

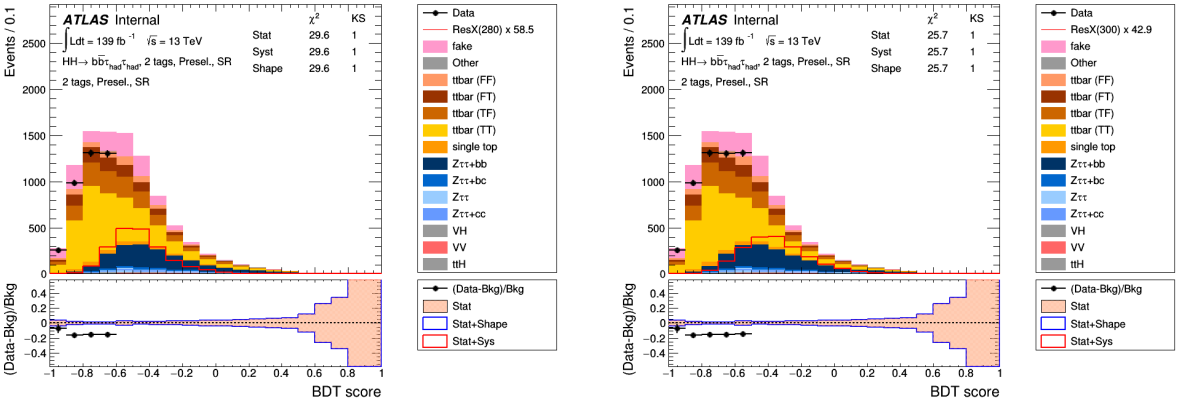
function of signal efficiency look very similar, for both even and odd event numbers as well as for the benchmark and the tuned hyperparameters. Thus, both BDTs are nearly equal in terms of signal-to-background separation power. The benchmark inverse background efficiencies are slightly lower than the tuned values, but this does not have a big impact on the performance of both BDTs.

Figures 5.4 and 5.5 show the inverse background efficiencies for both even and odd event numbers for the single-mass-point and adjacent-mass-points training, respectively. In addition, Figures 5.6 to 5.8 show the BDT distributions for the whole resonance mass range. The results of the resonant analysis can be summarised with the help of the previously mentioned figures as follows: Looking at Figures 5.4 and 5.5, both training methods seem to be approximately equal in signal-to-background separation power at a working point of 65 % signal efficiency (for even as well as for odd event numbers). The inverse background efficiency at a resonance mass of 251 GeV (158 for even event numbers and 161 for odd event numbers for the single mass training; 108 for even and 96 for odd even numbers for the adjacent mass points training) already indicates a good performance of the BDT. Figure 5.6(a), which shows the BDT distribution for a resonance mass of 251 GeV, supports this assumption since the kinematics at this resonance mass are very similar to the non-resonant Higgs boson pair production if the off-shell Higgs boson had a mass twice as large as the mass of an on-shell Higgs boson. A good performance is also seen for the case of resonance masses greater than or equal to 450 GeV (at 450 GeV the inverse background



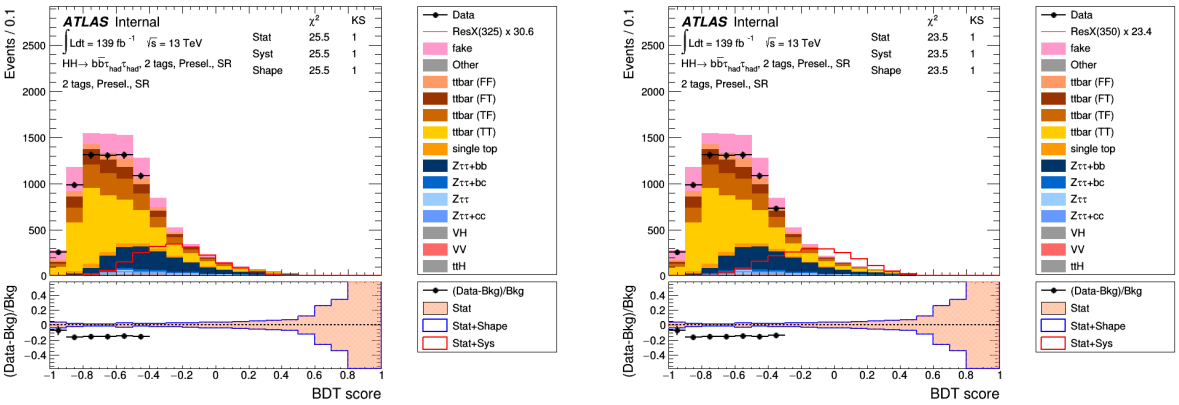
(a) The BDT distribution at 251 GeV.

(b) The BDT distribution at 260 GeV.



(c) The BDT distribution at 280 GeV.

(d) The BDT distribution at 300 GeV.

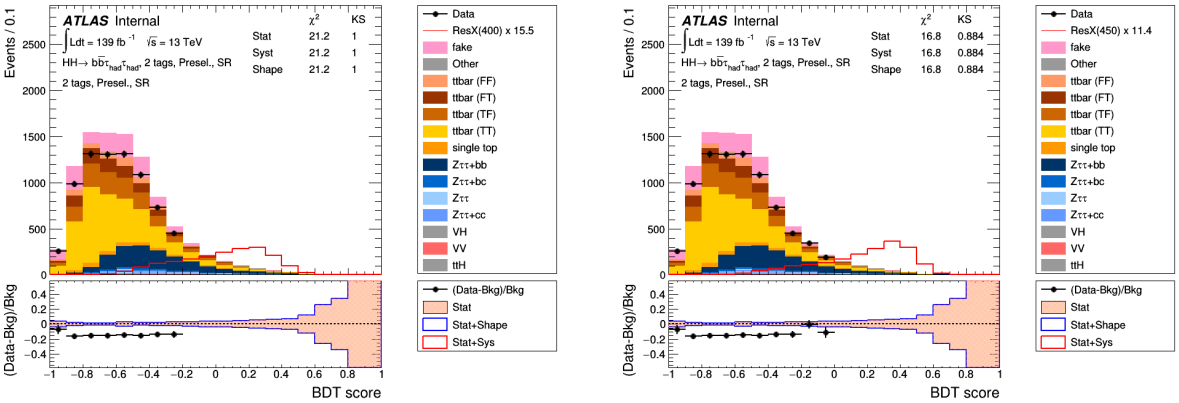


(e) The BDT distribution at 325 GeV.

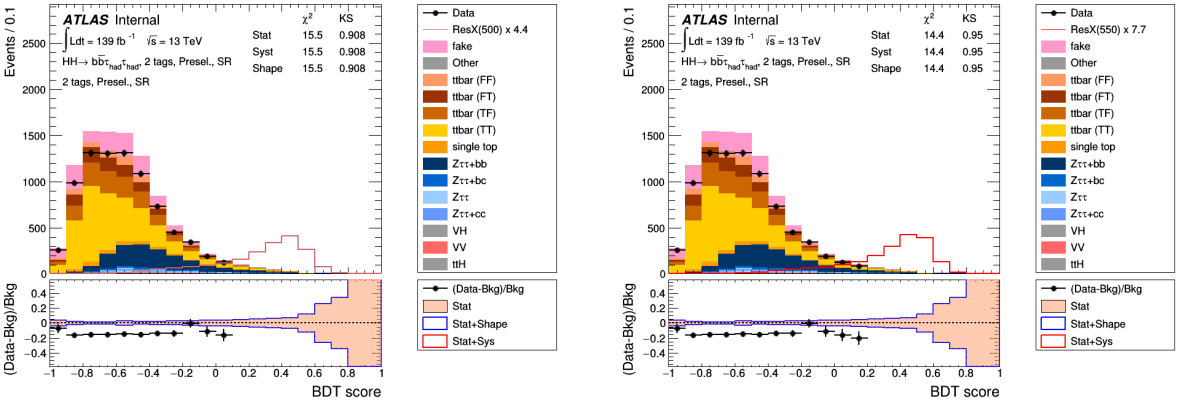
(f) The BDT distribution at 350 GeV.

Figure 5.6: The BDT distributions of the resonant analysis for resonance masses between 251 GeV and 350 GeV.

efficiencies are 115 (even) and 116 (odd) for the single mass training and 79 (even) and 80 (odd) for the adjacent mass points training). The inverse background efficiency reaches



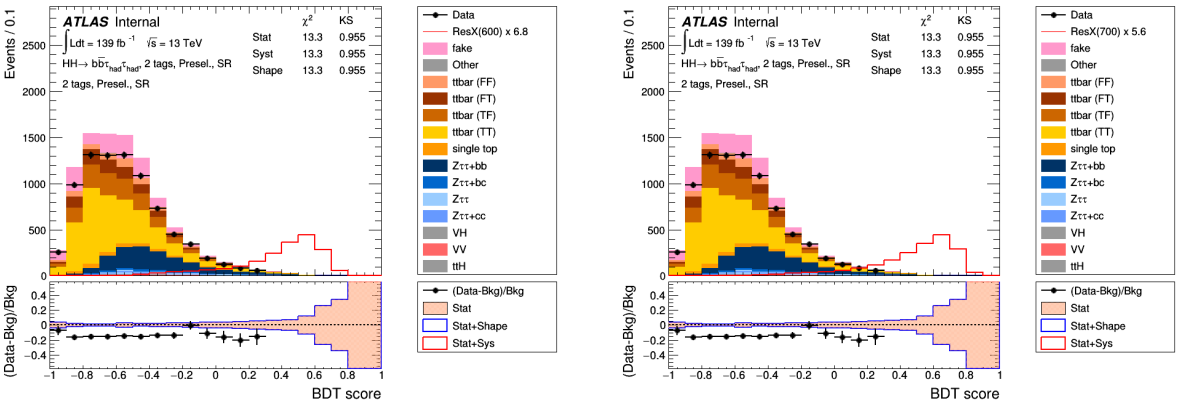
(a) The BDT distribution at 400 GeV.



(b) The BDT distribution at 450 GeV.

(c) The BDT distribution at 500 GeV.

(d) The BDT distribution at 550 GeV.

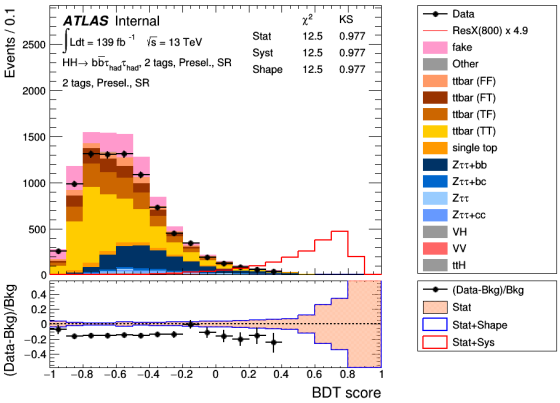


(e) The BDT distribution at 600 GeV.

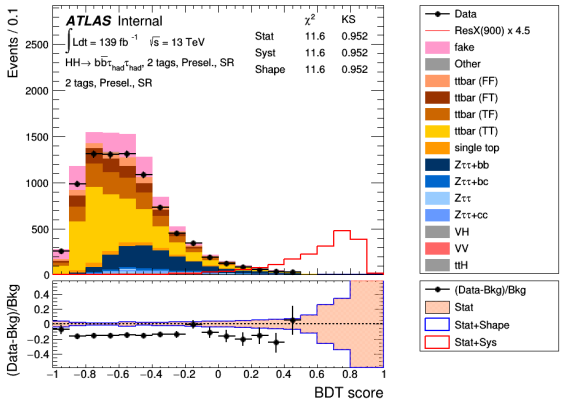
(f) The BDT distribution at 700 GeV.

Figure 5.7: The BDT distributions of the resonant analysis for resonance masses between 400 GeV and 700 GeV.

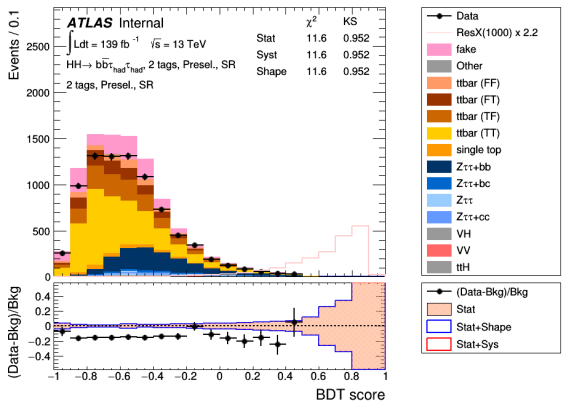
even larger values for higher resonance masses. Moreover, Subfigures 5.7(b) to 5.7(f) and Figure 5.8 show a continuous shift to the right in the signal distribution. This shows



(a) The BDT distribution at 800 GeV.



(b) The BDT distribution at 900 GeV.



(c) The BDT distribution at 1000 GeV.

Figure 5.8: The BDT distributions of the resonant analysis for resonance masses between 800 GeV and 1000 GeV.

clearly that the signal kinematics are easier to distinguish from background kinematics for higher resonance masses.

However, Figure 5.4 and 5.5 both also show a serious drop in the inverse background efficiency for resonance masses between 260 GeV and 400 GeV, reaching a minimum at 300 GeV or 325 GeV. The reason for this is that the signal kinematics are the most similar to background kinematics in this range of resonance masses. The $t\bar{t}$ production process especially impedes the separation of the resonance signal because it is the dominant background process in this decay channel and the invariant mass of a di-top system lies exactly in this range of energy. The BDT performance ideally needs to be improved in this region, unlike for other resonance mass points. In total, however, the trainings on adjacent mass points do not degrade the BDT performance by too much.

6 Conclusion and Outlook

A search for resonant and non-resonant Higgs boson pair production in the $HH \rightarrow b\bar{b}\tau_{\text{had}}\tau_{\text{had}}$ decay channel based on $139\,\text{fb}^{-1}$ of collision data recorded by the ATLAS experiment at the LHC at CERN during the Run 2 data-taking period at $\sqrt{s} = 13\,\text{TeV}$ was presented. BDTs were trained in both analyses in order to improve the signal-to-background separation.

In the non-resonant analysis, two sets of hyperparameters (*benchmark* and *tuned*) were used to train a BDT, each on two sets of events respectively: Events with even and events with odd event numbers. The resulting performances were compared and the investigation showed that both sets were suitable for the non-resonant analysis and nearly equal in terms of performance. Both BDTs yielded good signal-to-background separation power for events with both even and odd event numbers. This was supported by the inverse background efficiency that was plotted as a function of signal efficiency. All curves showed similar behaviour and a negligible difference in performance between each classifier. There is no sign of overtraining.

In the resonant analysis, a BDT was trained on only one set of hyperparameters (the tuned hyperparameters) in order to improve the separation power for a generic narrow-width scalar resonance signal and background. For this reason, 15 resonance masses were used as training points in this analysis. Furthermore, the BDT training was split into two methods: The first method was to train the BDT only on one resonance mass, the second one to train it on three neighbouring mass points. Both methods yielded similar results that showed good BDT separation power for a resonance mass of approximately twice the SM Higgs boson mass (251 GeV), but also for resonance masses bigger than or equal to 450 GeV. However, poorer results were achieved between 260 GeV and 400 GeV because of similar signal and background kinematics for this mass range.

By now, no statistically significant excess for Higgs boson pair production has been found in data. However, exclusion limits were set for the cross-section of Higgs boson pair pro-

duction by the ATLAS Collaboration, based on a partial Run-2 dataset. A production rate larger than 6.7 times the SM prediction was excluded. Furthermore, upper and lower limits were set on the Higgs self-coupling strength, which is constrained to $[-5.0, 12.1]$ times the SM expectation. No indications of resonant Higgs boson pair productions have been found and thus upper limits on their cross-sections are set. However, searches in the $HH \to b\bar{b}\tau\tau$ decay channel based on the full Run-2 dataset are already in process and might reveal new insights regarding Higgs boson pair production in the near future.

Future plans for the LHC are already designed and being planned. The *High-Luminosity Large Hadron Collider* project (HL-LHC) is supposed to start in 2027, making it possible for particle physicists to collect nearly ten times more data as they did between the years 2015 to 2018 during Run 2. With many more Higgs bosons produced and even more sensitive detectors, it should become possible to study non-resonant as well as resonant Higgs boson pair production far more deeply as it is now possible. Hence, multivariate discriminants and their performance in signal-to-background separation will very likely play an important role in the study of Higgs boson self-coupling as well as in the search for particles foreign to the SM and theories beyond the SM.

List of Figures

2.1 The Standard Model of particle physics. For each elementary particle, three quantum numbers are shown in the top left corner of every square: The highest one describes the mass of the particle in eV, the middle one is the electric charge of the particle in units of the elementary charge e, and the lowest one describes the spin of the particle. 6

2.2 A depiction of the shape of the Higgs potential $V(\Phi)$ for the case $\mu^2 < 0$ [14]. 9

2.3 The Feynman diagrams for the triple and quartic self-interaction of the Higgs boson with their corresponding coupling strengths. 11

2.4 The leading order Feynman diagrams for Higgs boson pair production by ggF in the SM [18]. 11

2.5 The cross-sections for different Higgs boson pair production modes as functions of the centre-of-mass energy, including their higher-order corrections [19]. 12

2.6 Cross-section limits of non-resonant Higgs boson pair production as a function of the ratio between the hypothesised Higgs self-coupling strength and the self-coupling strength predicted by the SM denoted as κ_λ. The left diagram was provided by the Atlas experiment, the right one by Cms [20]. 13

2.7 The leading order Feynman diagram for resonant Higgs boson pair production initiated by ggF with the $b\bar{b}q\bar{q}\ell\nu$ final state [30]. 15

2.8 Current cross-section limits for resonant Higgs pair production as a function of resonance mass from Atlas [31]. 15

2.9 Current cross-section limits for resonant Higgs pair production as a function of resonance mass from Cms [20]. 16

3.1 The Lhc, all its pre-accelerator systems and the intersection points of the four major experiments (Atlas, Cms, LHCb and Alice) [33]. 18

3.2 The structure of the Atlas detector at the Lhc at Cern. All parts of the inner detector, the calorimeters, the muon chamber, and the magnet system are labelled in this picture [34]. 19

3.3 The ID of the ATLAS detector and all its subsystems [36]. 21
3.4 The ATLAS calorimeter system with all its subsystems (ECAL and HCAL), used to detect electromagnetically and hadronically interacting particles [34]. 22

4.1 The BRs for different decay channels of the SM Higgs boson as a function of the Higgs mass [37]. 26
4.2 The BRs of the most common di-Higgs decay modes [17]. 26
4.3 The BRs of the di-tau decay channels. 27
4.4 The invariant masses of the di-Higgs, the di-b-jet and the di-tau systems for non-resonant Higgs boson pair production. Note, that the total area covered by the thin, red line (marking signal events) has been multiplied with a scaling factor such that it equals 20 % of the area covered by background events. The main background processes are fake-tau leptons (pink), $t\bar{t}$ and single top production (orange, yellow, and brown), $Z \to \tau\tau$+jets (dark and light blue), double vector boson production (pale red), single Higgs boson production with either a $t\bar{t}$ pair or a vector boson (light grey), and negligible processes summarised as "Other" (dark grey). Furthermore, 80 % signal blinding was used while modelling these histograms. 31
4.5 The event distributions for $\Delta\phi_{\mathrm{jj},p_T}$ and $\Delta\phi_{\tau,p_T}$. The red curves show the signal of the non-resonant Higgs boson pair production. Their scaling factor was chosen such that the area under each red curve is equal to 20 % of the total area covered by all background events. In addition, 80 % signal blinding was used while modelling these histograms. 33
4.6 The event distributions for $\Delta\phi_{\tau\tau,\mathrm{jj}}$, $\Delta\phi_{\tau\tau,p_T}$, ΔR_{bb}, $\Delta R_{\tau,\mathrm{b}}$, $\Delta R_{\tau\tau}$, and $\Delta R_{\tau\tau,\mathrm{jj}}$. The red curves show the signal of the non-resonant Higgs boson pair production. Their scaling factor was chosen such that the area under each red curve is equal to 20 % of the total area covered by all background events. In addition, 80 % signal blinding was used while modelling these histograms. 34
4.7 The event distributions for $E_T^{\mathrm{miss}}\phi$ centrality, M_{T2}^W, M_T^W, $p_T^{\tau_0}$, $p_T^{\tau_1}$, and $X_{\mathrm{m,bb}\tau\tau}$. The red curves show the signal of the non-resonant Higgs boson pair production. Their scaling factor was chosen such that the area under each red curve is equal to 20 % of the total area covered by all background events. In addition, 80 % signal blinding was used while modelling these histograms. 35

5.1 Schematic depiction of a decision tree [43]. 38

5.2 The BDT distributions for the benchmark (top row) and tuned (bottom row) hyperparameters as well as for even (left column) and odd (right column) event numbers. The KS values for signal and background events are shown in the top left corner of each subfigure (signal without parentheses and background with parentheses). 40
5.3 The inverse background efficiencies for the even-event-number (left) and odd-even-number (right) classifier as a function of signal efficiency. 41
5.4 The inverse background efficiency at 65 % signal efficiency as a function of the resonance mass for the BDT trained on a single mass point, tuned hyperparameters, and even/odd event numbers, respectively. 42
5.5 The inverse background efficiency at 65 % signal efficiency as a function of the resonance mass for the BDT trained on adjacent mass points, tuned hyperparameters, and even/odd event numbers, respectively. 43
5.6 The BDT distributions of the resonant analysis for resonance masses between 251 GeV and 350 GeV. 44
5.7 The BDT distributions of the resonant analysis for resonance masses between 400 GeV and 700 GeV. 45
5.8 The BDT distributions of the resonant analysis for resonance masses between 800 GeV and 1000 GeV. 46

List of Tables

3.1 The relative resolution for transverse momentum and energy as well as the η coverage of each ATLAS detector component [34]. E and p_T are both measured in GeV. 23

4.1 The most common SM Higgs boson decay channels and their corresponding BRs. In this table, the Higgs mass is assumed to be equal to 125.1 GeV [16]. 25

4.2 Observed and expected upper limits on the cross-sections times BR of non-resonant Higgs boson pair production in the $HH \to b\bar{b}\tau\tau$ decay channel at $\sqrt{s} = 13\,\text{TeV}$ [32]. $\pm 1\sigma$ represents the error of each observed limit. 28

4.3 Observed and expected upper limits on the cross-sections times BR for non-resonant Higgs boson pair production at $\sqrt{s} = 13\,\text{TeV}$ [17, 18, 38–40]. . 28

4.4 Explanations of the kinematic variables used to model the previously shown kinematic distributions. 36

5.1 The BDT benchmark (obtained from [32]) and tuned hyperparameters (obtained from [44]) that were used for training. The benchmark hyperparameters were originally used as intial hyperparameters for the analysis in [32]. Later on, they were changed in an attempt to "tune" the BDT performance (hence *tuned*) and used in [44] to train a BDT. 39

Bibliography

[1] A. Einstein, *Die Grundlage der allgemeinen Relativitätstheorie*, Annalen der Physik **47**, 769 (1916)

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19**, 1264 (1967)

[3] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nuclear Physics **22**, 579 (1961)

[4] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519**, 367 (1968)

[5] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716**, 1 (2012)

[6] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716**, 30 (2012)

[7] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13**, 508 (1964)

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

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

[10] ATLAS Collaboration, *Measurement of the W-boson mass in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, Eur. Phys. J. C **78**, 110 (2018)

[11] L3 Collaboration, *Measurement of the Z-boson mass using $e^+e^- \to Z\gamma$ events at centre-of-mass energies above the Z pole*, Phys. Lett. B **585**, 42 (2004)

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

[13] V. A. Bednyakov and N. D. Giokaris and A. V. Bednyakov, *On the Higgs mass generation mechanism in the Standard Model*, Physics of Particles and Nuclei **39**, 13 (2008)

[14] CMS Collaboration, *The Discovery of the Higgs Boson with the CMS Detector and its Implications for Supersymmetry and Cosmology*, Conference Proceedings Time and Matter (TAM2013) (2013), arXiv: 1309.0721 [hep-ph]

[15] S. F. Novaes, *Standard Model: An Introduction* (2000), arXiv: 0001283 [hep-ph]

[16] LHC Higgs Cross Section Working Group, S. Heinemeier, C. Mariotti, G. Passarino, R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties* (2013), arXiv: 1307.1347 [hep-ph]

[17] D. de Florian et al., *Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector* (2016)

[18] ATLAS Collaboration, *Search for non-resonant Higgs boson pair production in the $bb\ell\nu\ell\nu$ final state with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **801**, 135145 (2020)

[19] J. Baglio, A. Djouadi and J. Quevillon, *Prospects for Higgs physics at energies up to 100 TeV*, Rep. Prog. Phys. **79**, 116201 (2016)

[20] M. Gouzevitch and A. Carvalho, *A review of Higgs boson pair production*, Reviews in Physics **5**, 100039 (2020)

[21] C. Sivaram and Venkata Manohara Reddy. A, *Dark Matter, Dark Energy and Rotation Curves* (2007), arXiv: 0710.4384 [astro-ph]

[22] M. J. Rees, *Dark Matter: Introduction*, Phil. Trans. Roy. Soc. Lond. A: Mathematical, Physical and Engineering Sciences **361**, 2427 (2003)

[23] C. Csaki, *The Minimal Supersymmetric Standard Model (MSSM)*, Mod. Phys. Lett. A **11**, 599 (1996)

[24] G. C. Nayak, *Matter-Antimatter Asymmetry Of The Universe and Baryon Formation From Non-Equilibrium Quarks and Gluons* (2019), arXiv: 1909.05640v1 [physics.gen-ph]

[25] A. D. Dolgov, *CP violation in cosmology* (2005), arXiv: 0511213v2 [hep-ph]

[26] C. Giunti and M. Laveder, *Neutrino Mixing* (2003), arXiv: 0310238 [hep-ph]

[27] KATRIN Collaboration, *An improved upper limit on the neutrino mass from a direct kinematic method by KATRIN*, Phys. Rev. Lett. **123** (2019)

[28] H. Bossi and S. Chakdar, *A Symmetric Two Higgs Doublet Model* (2018), arXiv: 1810.13408 [hep-ph]

[29] C. H. Llewellyn Smith and G. G. Ross, *The real gauge hierarchy problem*, Phys. Lett. B **105**, 38 (1981)

[30] ATLAS Collaboration, *Search for Higgs boson pair production in the $b\bar{b}WW^*$ decay mode at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Journal of High Energy Physics **2019** (2019)

[31] ATLAS Collaboration, *Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Lett. B **800**, 135103 (2020)

[32] ATLAS Collaboration, *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**, 191801 (2018)

[33] E. Mobs, *The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019* (2019), CERN-GRAPHICS-2019-002

[34] ATLAS Collaboration, *Studies of the performance of the ATLAS detector using cosmic-ray muons*, Eur. Phys. J. C **71** (2011)

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

[36] ATLAS Collaboration, *The ATLAS Inner Detector operation, data quality and tracking performance* (2013), arXiv: 1303.3630 [physics.ins-det]

[37] A. Denner, S. Heinemeyer, I. Puljak, D. Rebuzzi and M. Spira, *Standard Model Higgs-Boson Branching Ratios with Uncertainties*, Eur. Phys. J. C **71** (2011)

[38] ATLAS Collaboration, *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*, Journal of High Energy Physics **2018** (2018)

[39] ATLAS Collaboration, *Search for Higgs boson pair production in the final state of $\gamma\gamma WW^*(\rightarrow \ell\nu jj)$ using 13.3 fb^{-1} of pp collision data recorded at $\sqrt{s}$ = 13 TeV with the ATLAS detector* (2016), ATLAS-CONF-2016-071

[40] ATLAS Collaboration, *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*, Phys. Rev. D **94**, 052002 (2016)

[41] A. Adhikary, S. Banerjee, R. K. Barman, B. Bhattacherjee and S. Niyogi, *Revisiting the non-resonant Higgs pair production at the LH-LHC*, Journal of High Energy Physics **2018(116)** (2018)

[42] A. Elagin, P. Murat, A. Pranko and A. Safonov, *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **654**, 481 (2011)

[43] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis* (2009), arXiv: 0703039 [physics.data-an]

[44] V. Grauer, *Investigation of Z+heavy flavour Background in the $HH \rightarrow b\bar{b}\tau^+\tau^-$ Search at the ATLAS Experiment*, II.Physik-UniGö-BSc-2019/03 (2019), Bachelor's thesis

[45] V. Sadhanala and Y. Wang and A. Ramdas and R. J. Tibshirani, *A Higher-Order Kolmogorov-Smirnov Test* (2019), arXiv: 1903.10083v1 [stat.ML]

Erklärung nach §13(9) der Prüfungsordnung für den Bachelor-Studiengang Physik und den Master-Studiengang Physik an der Universität Göttingen:

Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.
Darüberhinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestandenen Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den 12. Oktober 2020

(André Wilhahn)