GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Master's Thesis

Messung der Tau-Trigger-Effizienz und Untersuchung der CP-Eigenschaften des Higgs-Bosons im Ditau-Zerfallskanal mit dem ATLAS-Detektor

Measurement of the tau trigger efficiency and investigation of the CP nature of the Higgs boson in the ditau decay channel with the ATLAS detector

prepared by

Kieran Amos

from Stemwede

at the II. Physikalischen Institut

Thesis number:	II.Physik-UniGö-MSc-2019/03
Thesis period:	22nd October 2018 until 21st October 2019
First referee:	Prof. Dr. Stan Lai
Second referee:	Prof. Dr. Arnulf Quadt

Abstract

Experimente mit dem Atlas-Detektor, an denen Tau-Leptonen im Endzustand auftreten, verwenden das Tau-Triggersystem für die Online-Auswahl. Dieses Triggersystem muss mit der hohen instantanen Luminosität zurechtkommen, die im Jahr 2018 erreicht wurde. In dieser Arbeit wird die Tag-and-Probe-Methode zur Bestimmung der Tau-Triggereffizienz mit $Z \to \tau\tau$ Ereignissen erörtert. Für jedes Ereignis ist ein einzelnes Myon (Tag) erforderlich, und die Effizienz des Tau-Triggers wird dann aus dem Anteil der Ereignisse berechnet, bei denen der zugehörige hadronisch zerfallenden Tau-Kandidat (Probe) den Trigger passiert. Die Abhängigkeit der Effizienz vom Transversalimpuls und der Pseudorapidität des Tau-Kandidaten sowie die durchschnittliche Anzahl der Wechselwirkungen pro Überkreuzung werden dargestellt. Dieser Bericht diskutiert die Herausforderungen bei der Implementierung der Methode und dokumentiert Versuche zur Behebung von Fehlern bei der Modellierung und zur zukünftigen Verbesserung der Leistung.
Die zweite Hälfte der Arbeit gibt einen ersten Einblick in die Untersuchung der CP-Eigenschaft der VBF-Produktion des Higgs-Bosons im vollständig hadronischen $H \to \tau\tau$-Zerfallskanal unter Verwendung der *optimal observable* Methode mit dem vollständigen Run-2-Datensatz mit dem Atlas Detektor. Die für die spätere Analyse verwendete Event-Selektion wird untersucht und die Verteilung der kinematischen Variablen wie der der Methode des Optimalen Observablen und dafür relevanten Variablen wird beobachtet. Schließlich werden mehrere Signalproben analysiert, die unterschiedlichen Graden der CP-Verletzung entsprechen.

Abstract

Experiments with the Atlas detector involving tau leptons in the final state use the tau trigger system for the online selection. This trigger system must accommodate the high instantaneous luminosity achieved during the Lhc's run in 2018. In this thesis, the tag-and-probe method with $Z \to \tau\tau$ events used to determine the tau trigger efficiency will be discussed. For each event, a single muon (tag) is required, and the tau trigger efficiency is then calculated from the fraction of events where the accompanying hadronic tau decay candidate (probe) passes the trigger. The dependency of the efficiency on the transverse momentum and the pseudorapidity of the tau candidate as well as on the average number of interactions per bunch-crossing is presented. The thesis discusses the challenges involved in implementing the method, and documents attempts to resolve mismodelling issues and improve the performance for the future.
The second half of the thesis provides a first look at the investigation into the CP nature of the vector boson fusion production of the Higgs boson in the fully hadronic $H \to \tau\tau$ decay channel using the optimal observable method with the full Run 2 dataset with the Atlas detector. The event selection used for the later analysis is examined, and the distribution of kinematic variables such as the Optimal Observable and its dependent variables are investigated. Finally multiple signal samples corresponding to different degrees of CP violation are analysed.

Contents

1. Introduction

Tau leptons are important final state particles in many physics analyses. They provide the most precise measurement of the Yukawa coupling of the Higgs boson to fermions in the Standard Model (SM) [1], and are often preferred as the decay channel in Higgs pair production or Beyond the Standard Model (BSM) heavy resonances [2].
The first half of this report documents the measurement of tau trigger efficiencies and scale factors using the $Z \to \tau\tau$ Tag and Probe method using the ATLAS detector. In many analyses, the hadronic decays of tau leptons are used with a branching ratio of 65%, and a high acceptance of the associated trigger is required to extract the most from these events. Similarly, precise measurements of these trigger efficiencies are vital for many analyses involving tau leptons.
In the Tag and Probe method of determining the efficiency, $Z \to \tau\tau$ events are used where one tau lepton decays to a muon ($\tau \to \mu\nu_\mu\nu_\tau$) and the other decays hadronically ($\tau \to$ hadrons + ν_τ). Any discrepancies between the measurements in data and simulated Monte-Carlo (MC) events are accounted for using correction factors which are provided to ATLAS data analyses with the tau lepton final states to improve the trigger modelling. The data used was collected with the ATLAS detector during Run 2 in 2017 at $\sqrt{s} = 13\,\text{TeV}$ with a total integrated luminosity of $43.8\,\text{fb}^{-1}$.
The second part of the report provides a very early look at the ATLAS measurement of CP invariance in the vector boson fusion (VBF) production of the Higgs boson in the ditau decay channel $H \to \tau\tau$ with the final publication scheduled for 2021. Here, the full Run 2 dataset from 2015-2018 is used, corresponding to a total integrated luminosity of $140\,\text{fb}^{-1}$. As before, the focus in this report is on the fully hadronic decay channel. The full analysis uses the *optimal observable*[3] method where an observation of this variable calculated from leading-order VBF matrix elements allows for a direct measurement of CP invariance. The shape of the distribution is used to obtain a limit on a CP-odd extension to the otherwise CP-even SM matrix element for VBF Higgs production. This contribution is quantified by the parameter $\tilde{d}$. The previous Run-2 analysis using the 2015 and 2016 datasets at $36.1\,\text{fb}^{-1}$ established a value of $-0.090 < \tilde{d} < 0.035$ at 68% CL [4].
For now the focus is on the kinematic distributions and event yields at preselection level as

well as the generation of MC signal samples with varying levels of CP-odd contribution. This thesis is divided as follows: Chapter 2 details the Standard Model as well as the Higgs mechanism, the tau lepton and the theory behind the CP measurement and its implications for BSM scenarios. Chapter 3 gives a brief overview of the LHC and the ATLAS detector as well as the trigger system and the techniques used for modelling physical processes within the detector. For the tau trigger measurements, the event selection and background modelling employed are explained in Chapter 4 and the first results are presented in Chapter 5. Attempts to improve the Background modelling and identify issues in the decay to three charged pions are documented in Chapter 5.2. On the $H \to \tau\tau$ CP measurement, an overview of the full analysis is given in Chapter 6 and the work so far is presented in Chapter 7. Finally, both topics are summarised in Chapter 8.

2. Theory

2.1. The Standard Model of Particle Physics

The Standard Model (SM) attempts to provide a complete description of the building blocks of our universe as being comprised of fundamental particles and forces. Developed over the course of the last half-century, it currently provides the best known theoretical description of the building blocks of our universe with its successes ranging from the confirmation of the quark model to the astonishing accuracy of predictions made in Quantum Electrodynamics which have been verified in countless experiments. The 2012 discovery of the Higgs boson by the ATLAS and CMS experiments at CERN completed the Standard Model [5, 6].

The particles present in the SM can be divided into the twelve spin-$\frac{1}{2}$ fermions seen in Figure 2.1 from which all matter is comprised and the spin-1 gauge bosons with which they interact to give rise to the fundamental forces [7]. In the first few decades of the 20^{th} century, it was discovered that bulk matter is made up of atoms, consisting of protons (p) and neutrons (n) in the atomic core and surrounded by lighter electrons (e^-). A great breakthrough came with the discovery that the proton and neutron are not fundamental particles, but are comprised of *up*-quarks (u) and *down*-quarks (d). Along with the electron and the electron neutrino (ν_e), whose discovery was needed to account for missing momentum in beta decays, these quarks make up the first generation of fermions. As experiments began probing ever higher energies the existence of a second and third generation was uncovered with heavier fermions that are otherwise identical to those of the first generation. For the quarks the additions are (c), (s), (t) and (b) denoting the *charm*-quark, *strange*-quark, *top*-quark and *bottom*-quark respectively. The remaining fermions are referred to as leptons. The first generation electron (e^-) is complemented by the muon (μ^-) and the tau lepton (τ^-) along with their associated neutrinos (ν_e), (ν_μ), (ν_τ). Generally, fermions of higher generations quickly decay into first generation members, which explains the observation that everything in nature is built out of the lighter fermions. Finally, each fermion has an associated anti-particle that is identical in mass and lifetime but has opposite charges and other fundamental properties. For example the

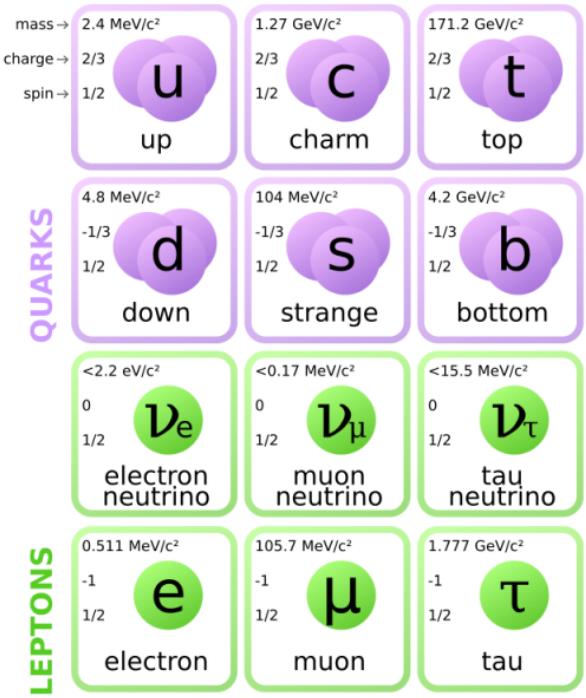


Figure 2.1.: The twelve fundamental fermions of the Standard Model along with their basic properties.

electron is paired with the positron (e^+), and each neutrino (ν) and quark (q) is mirrored by an anti-neutrino ($\bar{\nu}$) and an anti-quark ($\bar{q}$).
Along with their associated mass, charge and spin, all particles can be classified by the forces they experience. The four known fundamental forces are shown in Table 2.1. The

Force	Relative Stength	Boson		Mass [GeV]
Strong	1	Gluon	(g)	0
Electromagnetism	10^{-3}	Photon	(γ)	0
Weak	10^{-8}	W Boson	($W^{\pm}$)	80.4
		Z Boson	(Z)	91.2
Gravity	10^{-37}	Graviton?	(G)	0

Table 2.1.: The four fundamental forces and their associated bosons.

quarks are the only fermions to experience the strong nuclear force, while all fermions except the charge-less neutrinos undergo electromagnetic interactions. By contrast all fermions are found to react via the weak nuclear force. Note that gravity is not included in the Standard Model, and can be neglected in High Energy Physics experiments as it is over 30 orders of magnitude weaker than the other forces. Each force is mediated through the exchange of particles with integer spins called bosons and can be described using a

local gauge symmetry. The simplest example is the electromagnetic force with symmetry U(1), mediated by the photon (γ), resulting in the coupling strength being proportional to a single parameter, the electric charge Q. The theory of the electromagnetic interaction is known as Quantum Electrodynamics (QED). The other two forces are the strong nuclear force and the weak nuclear force.

2.1.1. The Strong Nuclear Force

The theory of the strong nuclear force is called Quantum Chromodynamics or QCD. Due to its strength, this force will tend to dominate particle interactions whenever possible. To take part in an interaction using the strong force, a particle must carry an appropriate charge known as colour charge, where the three different "colours", labelled r, g and b, are the orthogonal states of the corresponding SU(3) symmetry. The leptons do not carry a colour charge and consequently quarks are the only fermions that are involved in strong interactions. The strong nuclear force is mediated by the gluon (g), a massless spin-0 boson like the photon.
An important experimental observation in QCD is that neither quarks nor gluons have ever been observed individually. This is explained by the hypothesis of *colour confinement* which arises from gluon self interactions and states that only bound colourless states of quarks can ever be observed as free particles. The common permissible states are referred to as mesons $|q\bar{q}\rangle$, baryons $|qqq\rangle$ and antibaryons $|\bar{q}\bar{q}\bar{q}\rangle$.
In Quantum Field Theory, a process between initial and final particle states can be examined using the sum of all possible Feynman diagrams that correspond to these states. Although there are infinitely many diagrams for each possible process, their individual contribution is related to the number of vertices with each vertex, providing a factor $\alpha_S(q^2)$ known as the coupling constant which is dependent on the square of the momentum transferred via the interaction q^2 and is specific to the force involved. The coupling constant for the strong force α_S decreases considerably with q^2. This is known as *asymptotic freedom* and as a result, for low energies, α_S approaches unity making a perturbative approach for bound hadrons and low-energy jets impossible.

2.1.2. The Electro-Weak Force

The weak interaction differs from the other two forces of the Standard Model in a number of ways. The W^+, W^- and Z bosons that mediate it are massive particles with their own decay modes and are capable of coupling to all fundamental fermions. Additionally, only the weak force is found to violate charge parity (CP) [8]. As a consequence of parity

violation, charged-current W bosons only couple to left chiral particle states and right chiral anti-particle states.

Despite its small strength compared to the strong and electromagnetic forces, this mechanism plays an important role in many decays as it is the only one capable of changing particle flavour via an interaction vertex involving the W boson. Each lepton is paired with another in an isospin doublet and can transform into each partner via exchange of a W boson. The charged leptons are always paired with their corresponding neutrinos of their flavour. In principle this is also true for the quarks, which are paired within their respective generations. However, the weak eigenstates of quarks as elements of iso-spin doublets differs from their mass eigenstates. The mass eigenstates are determined by each fermions coupling to the Higgs boson, generating mass and thus determining the composition of the bound states that make up the observable baryons and mesons. The relationship between weak and mass eigenstates is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix which can be used to determine the precise strength of each individual coupling:

$$\begin{pmatrix} \mathrm{d}' \\ \mathrm{s}' \\ \mathrm{b}' \end{pmatrix} = \begin{pmatrix} \mathrm{V_{ud}} & \mathrm{V_{us}} & \mathrm{V_{ub}} \\ \mathrm{V_{cd}} & \mathrm{V_{cs}} & \mathrm{V_{cb}} \\ \mathrm{V_{td}} & \mathrm{V_{ts}} & \mathrm{V_{tb}} \end{pmatrix} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \\ \mathrm{b} \end{pmatrix}.$$

Here the weak eigenstates d', s' and b' are expressed as linear combinations of the observed mass states d, s and b. The square of the absolute value of each element is then proportional to the strength of each coupling between the W and the two quarks. Because the off-diagonal elements are non-zero, generational mixing is allowed for quarks. In experiments, the diagonal elements describing reactions within the same generation are found to dominate.

One of the ultimate goals of particle physics is to incorporate the individual theories for the forces into a single unifying theory. This has been achieved for the weak and electromagnetic forces in the Glashow-Salam-Weinberg (GSM) Model [9–11]. The $\mathrm{SU}(2)_L$ symmetry from the weak interaction is extended with a modified version of electromagnetic symmetry to make $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. The electric charge Q and the third component of weak isospin $I_W^{(3)}$ make up the electroweak charge known as hypercharge $Y = 2(Q - I_W^{(3)})$.

2.2. The Higgs Mechanism

The Standard Model Lagrangian is required to be invariant under local gauge transformations. This is easily achieved for the massless photon and gluon, but the introduction of mass-terms into the Lagrangians for the massive weak bosons as well as the fermions

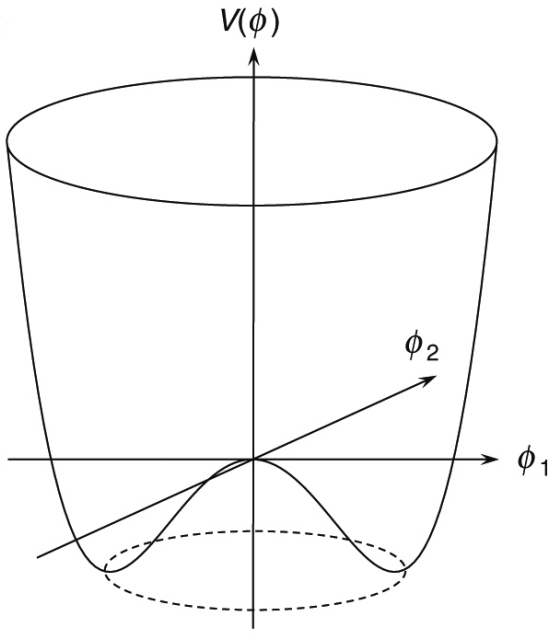


Figure 2.2.: The Higgs potential $V(\phi)$ for a complex, scalar field with $\lambda > 0$ and $\mu^2 < 0$.

breaks the gauge symmetry. The situation can be remedied in a process known as *spontaneous symmetry breaking*, where a field is introduced that consists of a weak isospin doublet of two complex scalar fields:

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

The resulting Lagrangian includes terms describing the Higgs potential:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2.$$

For $V(\phi)$ to have a minimum that corresponds to the vacuum state, it is necessary that $\lambda > 0$. The form of the potential for a single complex scalar field is shown in Figure 2.2 for $\mu^2 < 0$.

In this case the minimum is given as

$$\phi^\dagger \phi = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda}.$$

Writing the Higgs Doublet in the unitary gauge and accommodating the massless photon leads to

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}$$

with $h(x)$ as the physical Higgs field.
The mass terms of the gauge bosons can be determined from the Lagrangian $(D_\mu \phi)^\dagger (D^\mu \phi)$, where the ordinary derivatives have been replaced by the covariant derivatives of the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ local gauge symmetry

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + i g_W \mathbf{T} \cdot \mathbf{W}_\mu + i g' \frac{Y}{2} B_\mu$$

Here, g_W and g' are the coupling constants of the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ local gauge symmetry, $\mathbf{T}$ contains the generators of $\mathrm{SU}(2)$, Y is the hypercharge and W_μ and B_μ are the gauge boson fields. The four original degrees of freedom now correspond to the three gauge bosons of the electroweak theory as well as an additional scalar, spin-0 particle from the excitation of the Higgs field:

$$m_W = \frac{1}{2} g_W v, \quad m_Z = \frac{1}{2} v \sqrt{g_W^2 + g'^2}, \quad m_A = 0, \quad m_h = \sqrt{2\lambda} v.$$

The Higgs mechanism also generates the fermion masses with the exception of the neutrinos. The approach differs slightly for up- and down-type fermions but after spontaneous symmetry breaking with $\phi^0 = v + h(x)$, the Lagrangian for the fermion fields ψ gains the terms

$$\mathcal{L}_\psi = -\frac{g_\psi}{\sqrt{2}} v \bar{\psi} \psi - \frac{g_\psi}{\sqrt{2}} h \bar{\psi} \psi.$$

The coupling g_ψ is known as the Yukawa coupling and is not predicted by the Higgs mechanism, however the the first term is the coupling to the non-zero vacuum expectation value of the Higgs field which can be seen as a mass term with

$$g_\psi = \sqrt{2} \frac{m_\psi}{v}.$$

The second term describes the coupling of the fermion to the Higgs boson itself.
One of the main goals of the Lhc at Cern was the detection of the Higgs boson, and in July 2012 the two main experiments, Atlas and Cms, independently announced the detection of a particle with a mass of 125 GeV that was subsequently shown to be consistent with the SM prediction of the Higgs boson [5, 6]. The current best value for the mass of the Higgs boson is $m_h = 125.10 \pm 0.14\,\mathrm{GeV}$ [12].

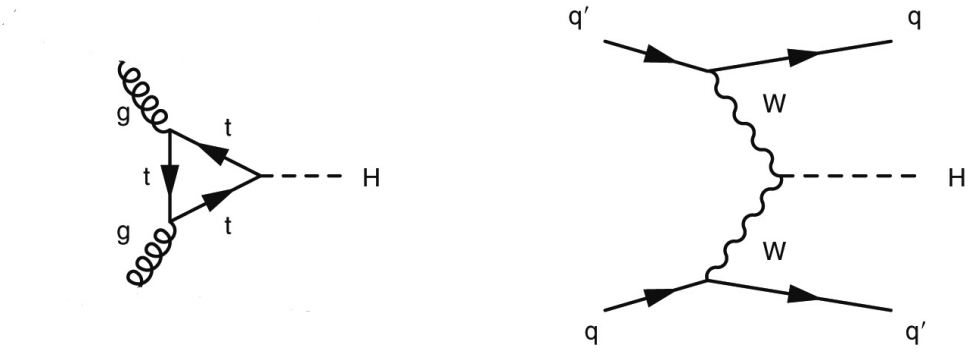


Figure 2.3.: The two most common Feynman diagrams for the production of the Higgs boson.

2.2.1. Higgs Production and Decay Modes

The detection of Higgs bosons is complicated both by the high energy required and the low cross section compared to the background events. The LHC was specifically designed with this in mind and operates at a high centre of mass energy at the luminosity frontier. The two most common Feynman diagrams for the production of a Higgs boson are shown in Figure 2.3. In the first case, known as *gluon-gluon fusion* (ggF), two gluons from the collection of virtual quarks and gluons within the colliding hadrons create a Higgs boson via a virtual top loop. In the second case, *vector boson fusion* (VBF), the Higgs boson is created directly from the annihilation of two W or Z bosons that are radiated by initial quarks in interacting protons. While the GGF cross section is much larger, the identification of the Higgs boson signal in this production mode is complicated by the large QCD background. For the VBF process, the scattered quarks from the colliding hadrons are expected to propagate in the beam direction, and thus the Higgs boson signal can be more easily separated from the relevant backgrounds. The cross sections for both processes at $\sqrt{s} = 13\,\text{TeV}$ and $m_h = 125\,\text{GeV}$ are shown in Table 2.2. Included are the cross sections for Higgs Radiation (WH, ZH) and ttH and bbH channels where a Higgs boson is created in association with a top or bottom quark pair.

The Higgs boson can potentially decay into all particles that have mass with the exception of the heavier top quark, but the coupling strength is proportional to the mass of the involved particles. The most relevant branching ratios for the observed 125 GeV Higgs boson are shown in Table 2.3. The difficulty of resolving decays involving jets made the $H \to q\bar{q}$ modes unlikely candidates for the Higgs discovery with the possible exception of $H \to b\bar{b}$ where the mesons containing b-quarks can be identified by the secondary vertices created at the point of their decay. In the case of the $\mathrm{W^+W^-}$ or $\tau^+\tau^-$ modes decaying leptonically, the undetectable neutrinos hamper the energy resolution. As a result, the

Process	Cross section [pb]
ggF	43.92
VBF	3.748
WH	1.380
ZH	0.9753
ttH	0.5085
bbH	0.5116

Table 2.2.: Cross sections for Higgs boson production at $\sqrt{s} = 13\,\text{TeV}$ and $m_H = 125\,\text{GeV}$ [13]

Decay mode	Branching ratio [%]
$H \to b\bar{b}$	57.8
$H \to WW^*$	21.6
$H \to gg$	8.6
$H \to \tau^+\tau^-$	6.4
$H \to c\bar{c}$	2.9
$H \to ZZ^*$	2.7
$H \to \gamma\gamma$	0.2

Table 2.3.: The predicted branching ratios for a Higgs boson with mass $m_H = 125\,\text{GeV}$ [13].

Higgs Boson was observed using the much rarer top-loop induced $H \to \gamma\gamma$ channel and the $H \to ZZ^*$ channel where the Z bosons decay into four charged leptons. Since then the ATLAS and CMS detectors have produced evidence of $H \to \tau\tau$ [1] and $H \to b\bar{b}$ [14] decays.
So far only single Higgs production has been observed. This is consistent with the low cross section predicted to be 33.70 fb based on NNLO calculations [15] for pair produced Higgs bosons predicted by the Standard Model at $\sqrt{s} = 13\,\text{TeV}$. Beyond Standard Model (BSM) theories offer the possibility of observing significantly more di-Higgs events at current energies possibly as a result of a resonance from a heavier Higgs boson.

2.3. The Tau Lepton

The tau lepton (τ) is the third generation charged lepton, discovered in 1975 at the Stanford Linear Accelerator (SLAC), with a mass of ($M_\tau = 1776.91 \pm 0.12$) MeV [17], approximately 3500 times the mass of the electron. With a lifetime of just $\tau_\tau = 2.9 \cdot 10^{-13}\,\text{s}$, corresponding to a decay length of $l_\tau \approx 2\,\text{mm}$ at $E = 40\,\text{GeV}$, they typically decay too quickly to reach the active regions of the detector, requiring reconstruction from their

Decay mode	Branching ratio [%]	Classification
$e^-\bar{\nu}_e\nu_\tau$	17.82 ± 0.04	leptonic
$\mu^-\bar{\nu}_\mu\nu_\tau$	17.39 ± 0.04	
$h^-\nu_t au \geq 0$ neutrals	48.52 ± 0.11	hadronic, 1-prong
$h^-h^-h^+\nu_t au \geq 0$ neutrals	15.20 ± 0.06	hadronic, 3-prong
Other decay modes	≈ 1.07	

Table 2.4.: Classification of tau lepton decay modes and corresponding branching ratios. $h^\pm$ can be either $\pi^\pm$ or $K^\pm$. [16].

decay products.

Due to their high mass, tau leptons are the only leptons capable of decaying not only leptonically ($\tau \to \ell\nu_\ell\nu_\tau$, $\ell = \mu, e$) but also hadronically ($\tau \to \text{hadrons} + \nu_\tau$) in about 65% of all cases. The typical hadronic decay is comprised of either one or three charged hadrons making up 72% and 22% of these cases respectively and known as 1-prong and 3-prong decays. These are mostly pions $\pi^\pm$ and one or more neutral pions π^0 may also be produced. The hadrons make up the visual portion of the hadronic decay and are referred to as $\tau_{\text{had-vis}}$. An overview of the decay channels is shown in Figure 2.4

Quarks and gluons produce jets of hadrons which are the main objects that can be misidentified as tau leptons. The main variables used to discriminate against them are the width of the shower in the calorimeters, which is compromised of fewer particles and narrower for tau leptons in comparison to hadronic jets, the number of charged tracks assuming values other than one or three, and the lack of a displaced secondary vertex from the decay of the original tau lepton. Track reconstruction can be hampered by overlapping decay products leading to fewer observed prongs, or by preemptive neutral pion decay to photons ($\pi^0 \to \gamma\gamma$) leading to electron-positron pair production cascades within the detector which are prone to being misidentified as additional prongs.

Tau leptons are important final state particles at the LHC and are utilised in many physics analyses ranging from electro-weak precision measurements [18] to Higgs [1] and top quark [19] physics and BSM searches [20].

2.4. Beyond the Standard Model

Despite its successes, the Standard Model is far from a complete picture of reality. It is dependent on numerous parameters that must be determined from experiment, ranging from fermion masses to gauge coupling strengths and mixing angles between particle states. The origin of the values of these parameters is unknown, and is a prime motivator in the search for a theory of everything.

Such a theory must also be capable of answering further outstanding questions in physics. These include the addition of gravity and the reconciliation of quantum field theory and general relativity, as well as the reason for the primordial matter-antimatter asymmetry that gave rise to the matter-dominated universe we observe today. Today it is known that neutrinos have mass, but the mechanism by which they acquire it is still undetermined. Cosmological observations show an excess of matter in the universe that cannot be seen directly, but there is no particle in the Standard Model capable of fulfilling this role.
In the 1970's the electromagnetic and weak interactions were unified into the electro-weak interaction via the Glashow-Salam-Weinberg Model [11] which is today part of the Standard Model. A common goal in particle physics is to expand the electro-weak model to include the strong force, with supersymmetry being a prime candidate [21]. However, as of today there is little evidence to favour one theory from another with the accommodation of more experimental data vital for future progress.

2.4.1. Baryogenesis and Sakharov Conditions

As mentioned above, the observable universe today is dominated by matter. This is at odds with predictions from the Standard Model, which hypothesises that matter and anti-matter were created in equal amounts during the Big Bang and should be destroyed at the same rate in pair-annihilation. Proposed solutions that result in this *baryogenesis* consequently require physical laws that treat the two states differently.
The Sakharov conditions were proposed in 1967 by Andrei Sakharov as three conditions necessary for interactions to favour baryons over anti-baryons [22]:

- Baryon number violation
- CP-symmetry violation
- Interactions out of thermal equilibrium

CP violation is explained in detail in the next section. Baryon number violation in some particle interactions is clearly required to lead to an imbalance between matter and anti-matter. This is not known to occur in the Standard Model, but is accommodated for in extensions of the SM with processes such as proton decay. In the last condition Sakharov states that the rate of expansion of the universe must be large enough that thermal equilibrium between baryons and anti-baryons can at some point no longer be achieved, shutting down the asymmetry generation and leaving an imbalance in favour of baryons. In thermal equilibrium, states with the same energy are equally likely, and since particles and their counterparts have the same mass, this would lead to the probability of positive

baryon states being equal to their negative states with the system oscillating between the two.

2.5. CP Symmetry and $\tilde{d}$

Charge conjugation and parity (CP) symmetry is the invariance of physical laws under the combined operations of replacing all particles with their antiparticle counterparts (charge conjugation) and inverting their spatial coordinates (parity). In 1956 parity violation was observed individually in the weak interaction using the beta decay of cobalt-60 [23]. This led to the proposal that P-violation could be compensated by additional C-violation to insure overall invariance under CP. However, in 1964, CP violation was also discovered [24], in the oscillations of neutral kaons and today it is known that the violation of this symmetry is one of the necessary Sakharov conditions. Adding time reversal leads to CPT symmetry which is strongly expected to be invariant for all physical phenomena via the CPT theorem [25]. CP violation thus implies the existence of T-violation.
The Standard Model contains three possible sources of CP violation. It was observed in the weak interaction in the Cabibbo-Kobayashi-Maskawa (CKM) matrix of the quark sector as described in Section 2.1.2. The violation in this mechanism is not strong enough to solve the problem of baryogenesis by itself. The Standard Model description of the strong interaction includes the possibility of CP violation [26], but measurements have failed to show any evidence for it. This is known as the strong CP problem and is one of the larger unsolved problems in particle physics. The final source of CP violation comes from the neutrino sector. Neutrino oscillations are described by the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, the equivalent of the CKM matrix and various ongoing and planned neutrino experiments are investigating the possible existence of a complex phase enabling CP violation [27].
The Standard Model does not provide enough CP violation to explain baryogenesis, and additional BSM mechanisms for CP violation remain an active research topic. This thesis investigates the possibility of CP violation in the VBF production of the Higgs boson. Applying either the charge or parity operator twice to an eigenstate much leave the state unchanged. The eigenvalues of C and P, and hence CP are thus ± 1. Observables comprised of these states can then be classified as CP-even or CP-odd respectively. In employing a variable that is CP-odd, it is possible to measure the interference between the SM and CP-odd couplings and thus perform a direct measurement of CP invariance without resorting to searching for increased event rates as with CP-even observables.
To describe this framework, an effective Lagrangian is used consisting of the Standard

Model Lagrangian expanded with CP-violating mass dimension six operators describing the interaction between the Higgs doublet and the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ electroweak gauge fields. The other Higgs interactions are assumed to be the same as in the Standard Model. After electroweak symmetry breaking, this effective Lagrangian can be written as [28]

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \tilde{g}_{HAA} H \tilde{A}_{\mu\nu} A^{\mu\nu} + \tilde{g}_{HAZ} H \tilde{A}_{\mu\nu} Z^{\mu\nu} + \tilde{g}_{HZZ} H \tilde{Z}_{\mu\nu} Z^{\mu\nu} + \tilde{g}_{HWW} H \tilde{W}^{+}_{\mu\nu} W^{-\mu\nu}$$

expressed in terms of the Higgs boson H, the photon A and the weak gauge bosons Z and $W^{\pm}$. The four coupling factors $\tilde{g}_{HVV}$ ($V = W^{\pm}, Z, A$) only have two degrees of freedom due to constraints introduced by $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ invariance. Written as the dimensionless parameters $\tilde{d}$ and $\tilde{d}_B$, these are:

$$\begin{aligned}
\tilde{g}_{HAA} &= \frac{g}{2m_W}(\tilde{d}\sin^2\theta_W + \tilde{d}_B \cos^2\theta_W) \\
\tilde{g}_{HAZ} &= \frac{g}{2m_W}\sin 2\theta_W(\tilde{d} - \tilde{d}_B) \\
\tilde{g}_{HZZ} &= \frac{g}{2m_W}(\tilde{d}\cos^2\theta_W + \tilde{d}_B \sin^2\theta_W) \\
\tilde{g}_{HWW} &= \frac{g}{m_W}\tilde{d}.
\end{aligned}$$

g is the weak charged coupling factor, m_W the mass of the W boson and θ_W the weak mixing angle.

While all these processes contribute to VBF production they cannot be distinguished experimentally and thus $\tilde{d} = \tilde{d}_B$ is assumed. This choice, while arbitrary, is consistent with previous analyses investigating CP invariance in $H \to WW$ and $H \to ZZ$ decays [29]. It results in a vanishing $HZ\gamma$ vertex:

$$\tilde{g}_{HAA} = \tilde{g}_{HZZ} = \frac{1}{2}\tilde{g}_{HWW} = \frac{g}{2m_W}\tilde{d} \quad \text{and} \quad \tilde{g}_{HAZ} = 0.$$

With this choice implemented the Lorentz structures of the Higgs to gauge boson coupling vertices in terms of the gauge boson momenta $p_{1,2}$ becomes

$$T^{\mu\nu}(p_1, p_2) = \sum_{V=W^{\pm},Z} \frac{2m_V^2}{v} g^{\mu\nu} + \sum_{V=W^{\pm},Z,\gamma} \frac{2g}{m_W}\tilde{d}\,\epsilon^{\mu\nu\rho\sigma} p_{1,\rho} p_{2,\sigma}.$$

The first terms describe the CP-even Standard Model coupling structure while the second term is CP-odd.

This leads to the VBF production matrix element $\mathcal{M}$ that can also be split into a CP-even

SM part $\mathcal{M}_{\text{SM}}$ and a CP-odd matrix element $\mathcal{M}_{\text{CP-odd}}$ proportional to $\tilde{d}$:

$$\mathcal{M} = \mathcal{M}_{\text{SM}} + \tilde{d} \cdot \mathcal{M}_{\text{CP-odd}}.$$

The differential cross section containing the matrix element squared consequently has three contributions:

$$|\mathcal{M}|^2 = |\mathcal{M}_{\text{SM}}|^2 + \tilde{d} \cdot 2\text{Re}(\mathcal{M}^*_{\text{SM}}\mathcal{M}_{\text{CP-odd}}) + \tilde{d}^2 \cdot |\mathcal{M}_{\text{CP-odd}}|^2.$$

Only the second term in this equation is CP-odd, however, it will not contribute to the total cross section or event yield when CP-symmetric selection criteria are applied. The third term increases the total cross section but is not the focus of this analysis.

3. Experimental Setup

3.1. The Large Hadron Collider

In collider experiments particles are smashed together in intersecting beams, and their energy is used to create heavy particles. These can subsequently decay into lighter, stable particles. Each interaction is governed by the strength of the relevant force, and analysing the resultant final state in specialised detectors allows insights into their precise nature.
The Large Hadron Collider (LHC) at CERN is currently the most powerful high-energy collider ever built. After approval for the project was granted in 1994, construction began in the 26.7 km tunnel previously used by the former LEP Collider, greatly reducing the total cost of the project. The LHC also makes use of the injection chain originally used by the previous experiment where particles are first accelerated in as series of smaller accelerators before injection into the main beam. At completion in 2008 the LHC operated at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ with an increase to $\sqrt{s} = 8\,\text{TeV}$ in 2012. In May 2015, after a two-year pause for accelerator and detector upgrades it ran at $\sqrt{s} = 13\,\text{TeV}$, and as of 2019, the collider is in a long shutdown period in anticipation of resuming operation at $\sqrt{s} = 14\,\text{TeV}$ in 2021. The design luminosity of the LHC is $10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ [30].
In its primary mode of operation, proton beams are inserted using a linear accelerator and a series of smaller synchrotrons and are then accelerated in opposite directions around two overlapping rings. Both rings can be divided into eight straight sections and eight curved sections. Major limiting factors on the maximum obtainable energy are the 1232 dipole magnets with a field strength of up to 8.33 T, requiring cooling to 1.9 K and situated at the curved regions. At the maximum operating capacity, the beams themselves consist of 2808 bunches, each containing approximately $1.15 \cdot 10^{11}$ protons. The collisions occur at four points of the rings on the straight segments at intervals of 25 ns. Here the four main detectors are located. ATLAS and CMS are general-purpose detectors known for their discovery of the Higgs boson in 2012. They are complemented by the ALICE and LHCb detectors, which specialise in heavy-ion collisions and b/c-hadron CP violation, respectively.

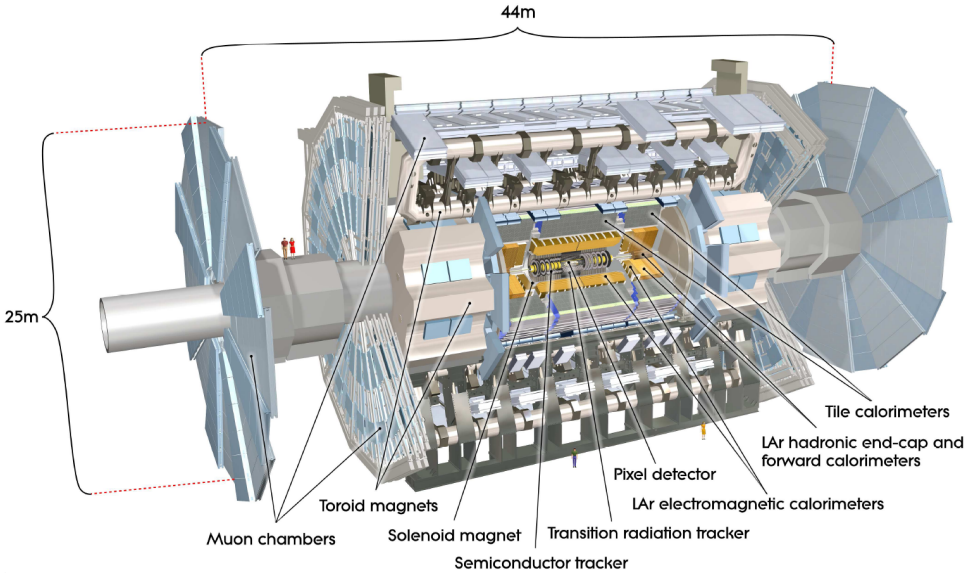


Figure 3.1.: An overview of the ATLAS detector.

3.2. The ATLAS Detector

The general-purpose ATLAS (A Toroidal LHC ApparatuS) detector is comprised of cylindrical layers and end-caps of various detector types constructed around a point where the beams intersect as shown in Figure 3.1. It has a length of 46 m, is 25 m wide and weighs approximately 7000 tons. Particular requirements during its design were the precise measurement of muons in the outer muon spectrometers and full reconstruction capabilities in the calorimeters resulting in additional detector layers in the end-caps to account for events with low transverse momentum [31]. The trajectories of particles are mostly described using cylindrical coordinates with the z-axis directed along the path of the colliding beams. In Cartesian coordinates a right-handed system is used with the x-axis pointing from the of the interaction point towards the centre of the LHC ring and the y-axis pointing upward. The main parameters derived from ATLAS coordinates to classify the particle trajectories in this experiment are:

1. The transverse momentum $p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}$.
2. The azimuthal angle, ϕ.
3. The pseudorapidity, $\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$, with θ the polar angle.
4. The Cartesian coordinate, z, along the beam line.

For highly relativistic angles, η is a good approximation of the rapidity, y, differences of which are invariant under Lorentz transformations along the beam axis. The angular

separation of particles or jets i, j in the detector can be measured using

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

3.2.1. Inner Detector

The Inner Detector [32] is 7 m long with a radius of 1.15 m and surrounded by a 2 T magnetic field generated by a solenoid magnet which encompasses the other components. Building outwards from the central beam are the silicon pixel detectors, the silicon microstrip (SCT) trackers and the Transition Radiation Tracker (TRT). The purpose of the Inner Detector is to accurately measure charged particle tracks and their interaction vertices. The pixel detectors and the SCT make up the precision tracking detectors at the centre. Divided into a barrel region with concentric cylinders around the beam line and end-cap regions with disks perpendicular to the beam line they provide coverage up to $|\eta| < 2.5$. The pixel detector features approximately 80.4 million read-out channels with a minimum pixel size of $50 \times 400\,\mu\text{m}^2$ in $R - \phi \times z$. Combining the the information from the three pixel layers crossed per track leads to intrinsic accuracies of $10\,\mu\text{m}(R - \phi)$ and $115\,\mu\text{m}(z)$ in the barrel and $10\,\mu\text{m}(R - \phi)$ and $115\,\mu\text{m}(R)$ in the end-caps.
The SCT makes use of approximately 6.3 million read-out channels spaced so that each track crosses 8 strip layers. Since each strip only allows for a one-dimensional measurement, adjacent strips are placed at an angle to each other. The intrinsic accuracies are $17\,\mu\text{m}(R-\phi)$ and $580\,\mu\text{m}(z)$ for the barrel and $17\,\mu\text{m}(R-\phi)$ and $580\,\mu\text{m}(R)$ for the disks. The TRT provides tracking in the outer region up to $|\eta| < 2.0$. 4 mm diameter straw tubes combine to a total of $351,000$ readout channels with an intrinsic accuracy per straw of $130\,\mu\text{m}$ in the $(R - \phi)$ plane. While lacking the resolution of the inner precision trackers, these straw tubes make a significant contribution to the momentum measurement due to the approximate 36 hits per track and the long track length.

3.2.2. Calorimeters

Outside of the solenoid the calorimeters measure the energy deposited in higher-density materials [33]. Their total coverage extends to $|\eta| < 4.9$ with the electromagnetic (EM) calorimeter providing the same coverage as the inner detector, optimised for the detection of photons and electrons, and the hadronic calorimeter extending the coverage for recon-

struction of jets and missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. The EM calorimeter is divided into the barrel ($|\eta| < 1.475$) and the end-caps ($1.375 < |\eta| < 3.2$) with a thickness of > 22 and > 24 radiation lengths respectively. Granular liquid argon is used, inter-spaced with lead absorber plates and cooled by the same cryostats used for the solenoid.
The hadronic calorimeter mainly consists of a tile calorimeter with the barrel at $|\eta| < 1.0$ and extended barrels at $0.8 < |\eta| < 1.7$. Scintillator tiles make up the active material with steel used as the absorber. It extends from a radius of 2.28 m to 4.25 m, corresponding to 9.7 interaction lengths at $\eta = 0$. The hadronic calorimeter is completed by liguid argon end-caps at $1.5 < |\eta| < 3.2$ and by the liquid argon forward calorimeter with a coverage of $3.1 < |\eta| < 4.9$. The former overlaps with both the EM end-caps and the forward calorimeter in order to maintain sufficient material density in the transition regions. The latter is comprised of three high-density modules in order to limit the radiation in the muon chamber. The first uses copper and is optimised for electromagnetic measurements while the last two use tungsten to measure hadronic interactions.

3.2.3. Muon system

The majority of the detector by volume is made up of the muon system surrounding the calorimeters. Superconducting air-core toroid magnets deflect muons and their tracks are then measured by trigger and high-precision tracking chambers [34]. The barrel toroid ($|\eta| < 1.4$) is complemented by two end-cap magnets ($1.6 < |\eta| < 2.7$) with the transition region covered by combination of the magnetic fields. Each toroid is made up of 8 individual coils with the resulting magnetic field being mostly perpendicular to the muon tracks. The field strength varies considerably but the total bending power ranges from 1 to 7.5 Tm.
The track coordinates are measured by high-precision tracking chambers with Monitored Drift Tubes (MDT's) over most of the η coverage. Cathode Strip Chambers (CSC's) cover the $2.0 < |\eta| < 2.7$ region with its higher rate and background contamination. Trigger chambers (Resistive Plate Chambers (RPC's) in the barrel and Thin Gap Chambers (TGC's) in the end-cap regions) cover the range $|\eta| < 2.4$. These help with bunch-crossing identification, establish p_{T} thresholds and measure the track coordinate in the direction orthogonal to the tracking chambers.
An overview of the resolutions and angle coverages of the various detector components can be found in Table 3.1.

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

Table 3.1.: General performance of the ATLAS detector showing the resolutions and coverage angles of the various components. E and p_T are in GeV [35].

3.3. Monte Carlo Simulation

Performing physics analyses with the ATLAS detector requires detailed modelling of particle reactions in the Standard Model, as well as precise knowledge of the detector's response to these physics processes. To achieve this, simulated samples are created using Monte Carlo (MC) generators. To reach the stage where these events can be directly compared with real events recorded using the actual detector, multiple steps must be undertaken which are contained in the ATLAS simulation infrastructure [36].

The initial generation of the event is performed with matrix element calculations that correspond to the relevant Feynman diagrams. This includes all particles created from the initial interaction vertex, but also the products of promptly-decaying particles such as W and Z bosons that decay too rapidly ($c\tau < 10\,\text{mm}$) for the magnetic fields and geometry of the detector to have an effect. Common MC generators include SHERPA [37] and PYTHIA [38].

Quarks and gluons are never observed singularly. As a result of colour confinement, the field energy between the increasingly isolated partons grows until it is more energetically favourable for a new $q\bar{q}$ pair to be produced. This continuous process is known as a *parton shower* and results in sprays of particles. The resulting particles propagate at high energies through the detector and once the initial energy is too sparsely distributed to allow for additional parton creation they combine to form colour singlet hadrons. This process is known as *hadronisation* and is prohibatively non-perturbative for matrix calculations for anything but the hardest partons. To account for this, parton shower models are applied that approximate the process by calculating the average number of radiation lengths for which the parton does not radiate additional particles. Parton shower models can be implemented with POWHEG [39] and SHERPA [40].

The last stage necessary to bring simulated events to resemble real events is a simulation of the detector. This must include detailed knowledge of the magnetic fields involved, and

an exact model of the geometry of the detector and the materials used. This is used to track the simulated particles flight path through the various detector elements and create the readout signals the real event would have generated. For the ATLAS detector, the simulation toolkit GEANT4 is used [41].

3.4. Object Reconstruction and Identification

In the ATLAS detector, measurements take the form of "hits" in the tracking systems and energy deposits in the clusters of the calorimeters. The hits are combined to reconstruct the tracks of charged particles and provide a measurement of the particle's momentum from the radius of the curvature under the influence of the magnetic field. These tracks are then matched to the energy deposits and the total information is used to reconstruct the physical objects that propagate through the detector. The particles with sufficient lifetimes to reach the detector subsystems are electrons, muons, photons and neutrinos, as well as bound hadronic states, such as protons, neutrons and pions that result in the physical entities known as jets. The identification of these objects is outlined here.
Electrons and photons are reconstructed from energy deposits in the electromagnetic calorimeter [42]. Clusters matched with well-reconstructed tracks in the inner detector resulting from the interaction region are classified as electrons. If the matched track originates from a separate vertex, the cluster is considered a converted photon, and if no track is matched, the cluster is classified as an unconverted photon. Photon and electron identification is based on shower shape variables in the EM calorimeter and the fraction of energy leaked into the hadronic calorimeter.
Muons are reconstructed from combining tracks in the inner detector with tracks in the outer muon system [43]. Muon identification prioritises the suppression of background contributions from pion and kaon decays. The main signature of these secondary muons are "kinks" in the flight path, recognised as non-matching inner detector and muon system tracks, as well as inconsistent momentum measurements by the two detector elements.
Jet reconstruction is based on topological clusters in the calorimeters with optional input from the inner detector [44]. Different jet clustering algorithms are used that establish the individual jets using the radius parameter R. The anti-K_t algorithm, for example, iteratively combines jet constituents using $R = 0.4$ and prioritising the hardest jet constituents first [45]. Jet vertex taggers suppress pile-up from additional collisions. b-jets originating from b-quarks are identified using a separate algorithm [46].
Overlapping reconstructed objects are typically removed based on the individual analysis requirement using ΔR criteria to ensure a minimum of separation.

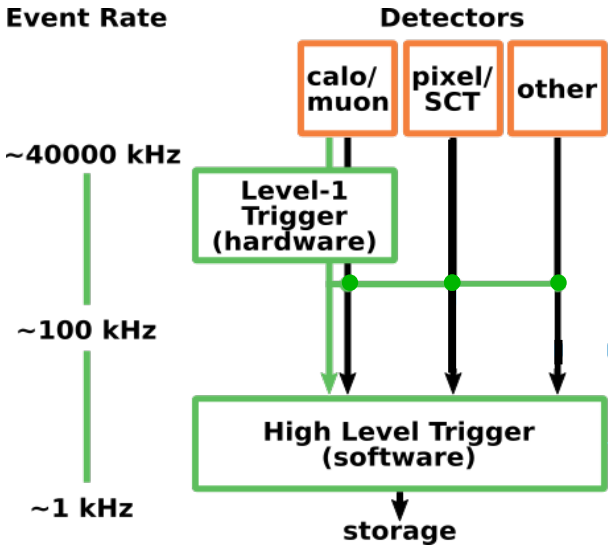


Figure 3.2.: A schematic of the trigger system used by ATLAS.

Neutrinos do not react with any material within the detector. Their presence is inferred from any missing transverse momenta $E_{\mathrm{T}}^{\mathrm{miss}}$. This is calculated from the negative of the magnitude of the sum of the transverse momenta of all other physical objects and tracks identified with a collision [47].

3.5. The ATLAS Trigger System

Collisions occur in the ATLAS detector at a rate of 40 MHz far exceeding modern storage capabilities. Since most of these are QCD multi-jet events and soft-scattering events of no use to physics analyses, ATLAS uses a two stage trigger system as of Run 2 to bring the data collection to a manageable rate while preserving the events of interest [48]. The trigger system is outlined in Figure 3.2. Calorimeter and muon system information is used in the first stage before this information is combined with that from the inner detector in the second stage. Events that survive are committed to storage.

3.5.1. The First Level Trigger

The first level trigger (L1) is hardware-based and has a maximum acceptance rate of 100 kHz. Multiple triggers run in parallel, designed to filter out events of interest. In the specific case of the tau trigger, tau candidates are built by considering a core region and a surrounding isolation region. These are defined in both calorimeters using trigger towers, which are blocks of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The transverse energy, E_T, of a $\tau_{\text{had-vis}}$ candidate

is the sum of the energy in the two most energetic neighbouring trigger towers within a 2×2 block in the EM calorimeter and the corresponding 2×2 block in the hadronic calorimeter. The EM isolation energy, E_T^{EMisol} is the transverse component of the energy deposited in the surrounding area up to 4×4 in the EM calorimeter. To suppress the background, the isolation energy is subjected to an upper threshold of $E_T^{EMisol}[\text{GeV}] \leq (E_T[\text{GeV}]/10 + 2)$ for $\tau_{\text{had-vis}}$ candidates up to 60 GeV. Above this threshold, the isolation requirement is no longer applied. This criteria is the result of fine-tuning to yield a selection efficiency of 98%.
The energy resolution at L1 level is poor compared to the offline reconstruction. No clustering algorithms are used when combining cells and no energy calibration specific to $\tau_{\text{had-vis}}$ candidates is applied. This, combined with the coarse granularity of the cells, is the main reason for the large loss in signal efficiency for low-p_{T}. The resolution of the L1 trigger for tau objects is limited to increments of 0.5 GeV.

3.5.2. The High-Level Trigger

The software-based high-level trigger (HLT) reduces the acceptance rate to an average of 1 kHz. For the tau trigger it is divided into *calo-only preselection*, *track preselection* and *offline-like selection* which are performed sequentially in an order aimed at minimising CPU usage.
In calo-only preselection the calorimeter cells in the region of interest (RoI) identified at L1 have the topo-clustering algorithm applied to them and are then calibrated using the local hadron calibration (LC) [49]. $\tau_{\text{had-vis}}$ candidate reconstruction is accomplished using the vectorial sum of these clusters as a 'jet seed' with the energy calculated from the LC clusters within $\Delta R < 0.2$ of the jet seed. Finally a dedicated $\tau_{\text{had-vis}}$ energy calibration (TES) is utilised. This p_{T} and η dependent calibration improves the energy measurement precision and includes pile-up corrections. Only $\tau_{\text{had-vis}}$ candidates that exceed a trigger-specific p_{T} value are passed on to the next stage.
The goal of track preselection is to add the track information to the reconstructed $\tau_{\text{had-vis}}$ candidate. This is done in two steps. The first step requires a leading track with $p_{\text{T}} > 1\,\text{GeV}$ within $\Delta R < 0.1$ of the RoI and with $|z| < 225\,\text{mm}$ along the beamline. If a track is found the second stage requires additional tracks within $\Delta R < 0.4$ of the RoI and within $|z| < 10\,\text{mm}$ of the leading track. This two-stage yields the highest efficiency within the allotted CPU time. To proceed to the final stage, a $\tau_{\text{had-vis}}$ candidate must fulfill $1 \leq N_{\text{core}}^{\text{trk}} \leq 3$ and $N_{\text{isol}}^{\text{trk}} \leq 1$ for the number of core and isolation tracks with $\Delta R < 0.2$ and $0.2 < \Delta R < 0.4$ for the number of core and isolation tracks $\leq N_{\text{core}}^{\text{trk}}$ and $N_{\text{isol}}^{\text{trk}}$ respectively.

For this thesis the offline-like selection is of particular importance. It identifies $\tau_{\text{had-vis}}$ candidates using a Boosted Decision Tree (BDT) algorithm with input variables calculated from calorimeter as well as track information [50]. Separate sets of input variables are used for the 1-prong and 3-prong algorithms. The BDT is trained on a signal of simulated $Z \to \tau\tau$ events and a QCD-jet enriched background taken from 2015 pp collision data. The identification score it provides is used to classify the *loose*, *medium* and *tight* working points for $\tau_{\text{had-vis}}$ candidates, where the medium working point is chosen to result in an efficiency of 96% and 82% for 1-prong and 3-prong tau candidates respectively. The performance of the BDT algorithm is illustrated in Figure 3.3 at HLT. The signal efficiency is defined as the ratio of signal events that pass the filter over the total signal events while the rejection is the inverse of the background efficiency defined the same as for the signal. Higher efficiencies result in reduced background rejection. 3-prong decays exhibit higher background rejection at a fixed efficiency, however, the 3-prong working points have been chosen to correspond to lower efficiencies compared to 1-prong to compensate for the decreased branching ratio. The BDT input variables used are as summarised in Table 3.2. Two of the twelve variables are used in the 1-prong classification only, while three are solely used in the 3-prong classification. The following detailed description is taken from [51]:

- **Central energy fraction (f_{cent}):** Fraction of the calorimeter transverse energy deposited in the region $\Delta R < 0.1$ with respect to all energy deposited in the region $\Delta R < 0.2$ around the $\tau_{\text{had-vis}}$ candidate. It is calculated by summing the energy deposited in all cells belonging to the three-dimensional clusters of cells (TopoClusters) with a barycenter in these regions, calibrated at the EM energy scale.

- **Leading track momentum fraction ($f^{-1}_{\text{leadtrack}}$):** The transverse energy sum, calibrated at the EM energy scale, deposited in all cells belonging to TopoClusters in the core region of the $\tau_{\text{had-vis}}$ candidate, divided by the transverse momentum of the highest-p_{T} charged particle in the core region.

- **Track radius (R_{track}):** p_{T}-weighted ΔR distance of the associated tracks to the $\tau_{\text{had-vis}}$ direction, using only tracks in the core region.

- **Leading track IP significance ($|S_{\text{leadtrack}}|$):** Absolute value of transverse impact parameter of the highest-p_{T} track in the core region, calculated with respect to the tau vertex, divided by its estimated uncertainty. The tau vertex is chosen from the list of primary vertex candidates to which the greatest fraction of p_{T} from the $\tau_{\text{had-vis}}$ candidate tracks is matched.

- **Fraction of tracks $p_{\mathbf{T}}$ in the isolation region ($f_{\mathbf{iso}}^{\mathbf{track}}$):** Scalar sum of the p_{T} of tracks associated with the $\tau_{\text{had-vis}}$ candidate in the region $0.2 < \Delta R < 0.4$ divided by the sum of the p_{T} of all tracks associated with the $\tau_{\text{had-vis}}$ candidate.

- **Maximum ΔR($\Delta R_{\mathbf{Max}}$):**The maximum ΔR between a track associated with the $\tau_{\text{had-vis}}$ candidate and the $\tau_{\text{had-vis}}$ direction. Only tracks in the core region are considered.

- **Transverse flight path significance ($S_{\mathbf{T}}^{\mathbf{flight}}$):** The decay length of the secondary vertex (vertex reconstructed from the tracks associated with the core region of the $\tau_{\text{had-vis}}$) in the transverse plane, calculated with respect to the tau vertex, divided by its estimated uncertainty. It is defined only for multi-track $\tau_{\text{had-vis}}$ candidates.

- **Track mass ($m_{\mathbf{track}}$):** Invariant mass calculated from the sum of the four-momentum of all tracks in the core and isolation regions, assuming a pion mass for each track.

- **Fraction of EM energy from charged pions ($f_{\mathbf{EM}}^{\mathbf{track\text{-}HAD}}$):** Fraction of the electromagnetic energy of tracks associated with the $\tau_{\text{had-vis}}$ candidate in the core region. The numerator is defined as the difference between the sum of the momentum of tracks in the core region and the sum of the cluster energy deposited in the hadronic part of each TopoCluster (including the third layer of the EM calorimeter) associated with the $\tau_{\text{had-vis}}$ candidate. The denominator is the sum of cluster energy deposited in the electromagnetic part of each TopoCluster (presampler and first two layers of the EM calorimeter) associated with the $\tau_{\text{had-vis}}$ candidate. All clusters are calibrated at the LC energy scale.

- **Ratio of EM energy to track momentum ($f_{\mathbf{track}}^{\mathbf{EM}}$):** Ratio of the sum of cluster energy deposited in the electromagnetic part of each TopoCluster associated with the $\tau_{\text{had-vis}}$ candidate to the sum of the momentum of tracks in the core region. All clusters are calibrated at the LC energy scale.

- **Track-plus-EM-system mass ($m_{\mathbf{EM+track}}$):** Invariant mass of the system composed of the tracks and up to two most energetic EM clusters in the core region, where EM cluster energy is the part of TopoCluster energy deposited in the presampler and first two layers of the EM calorimeter, and the four-momentum of an EM cluster is calculated assuming zero mass and using TopoCluster seed direction.

- **Ratio of track-plus-EM-system to$p_{\mathbf{T}}$($p_{\mathbf{T}}^{\mathbf{EM+track}}/p_{\mathbf{T}}$):** Ratio of the $\tau_{\text{had-vis}}$ p_{T}, estimated using the vector sum of track momenta and up to two most energetic EM clusters in the core region to the calorimeter-only measurement of $\tau_{\text{had-vis}}$ p_{T}.

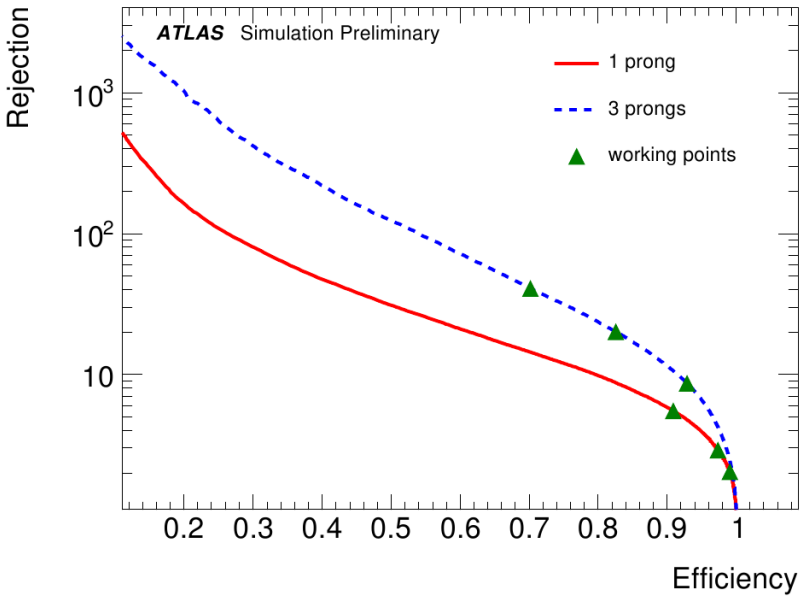


Figure 3.3.: Performance of the BDT algorithm at HLT in terms of efficiency versus background rejection for *loose*, *medium* and *tight* working points for 1-prong and 3-prong $\tau_{\text{had-vis}}$ candidates [52].

Variable	Description	1-prong	multi-prong
f_{cent}	Central energy fraction	•	•
$f_{\text{leadtrack}}^{-1}$	Leading track momentum fraction	•	•
R_{track}	Track radius	•	•
$\lvert S_{\text{leadtrack}} \rvert$	Leading track impact parameter significance	•	
$f_{\text{iso}}^{\text{track}}$	Fraction of p_{T} from tracks in the isolation region	•	
ΔR_{Max}	Maximum ΔR		•
$S_{\text{T}}^{\text{flight}}$	Transverse flight path significance		•
m_{track}	Track mass		•
$f_{\text{EM}}^{\text{track-HAD}}$	Fraction of EM energy from charged pions	•	•
$f_{\text{track}}^{\text{EM}}$	Ratio of EM energy to track momentum	•	•
$m_{\text{EM+track}}$	Track-plus-EM-system mass	•	•
$p_{\text{T}}^{\text{EM+track}}/p_{\text{T}}$	Ratio of track-plus-EM-system to p_{T}	•	•

Table 3.2.: List of track and calorimeter derived variables used in the tau-ID BDT algorithm [52].

4. The Tag-and-Probe Method for Tau Lepton Efficiency Measurements

This chapter discusses the Tag-and-Probe method which utilises semi-leptonic $Z \to \tau\tau$ decays to determine the efficiency of the tau trigger. The signal region is explained, along with the control regions used to model the background contributions. The first event yields are presented.

4.1. Event Selection and Signal Region

The selection criteria are optimised for a signal of $Z \to \tau\tau \to \mu\tau_{\text{had-vis}}3\nu$ events, where one tau lepton decays to a muon and the other decays hadronically as shown in Figure 4.1. The selected events are thus required to contain a muon, known as the *tag*, and an offline reconstructed $\tau_{\text{had-vis}}$ candidate, the *probe*. The selection criteria used are listed in Table 4.1. The muon requires a transverse momentum $p_T \geq 26\,\text{GeV}$, pseudorapidity $|\eta| < 2.5$ and must be trigger selected and identified with the medium muon selection [43], as well as

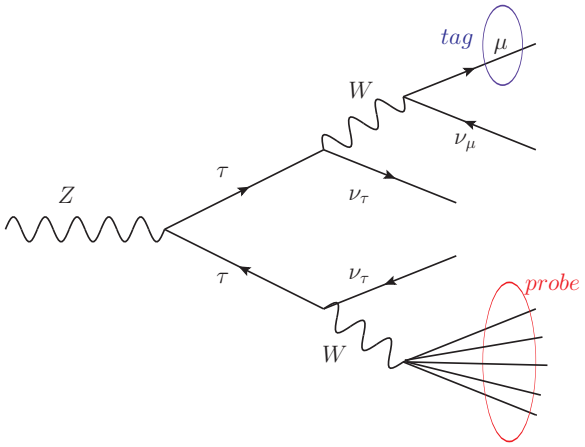


Figure 4.1.: The $Z \to \tau\tau \to \mu\tau_{\text{had-vis}}3\nu$ decay.

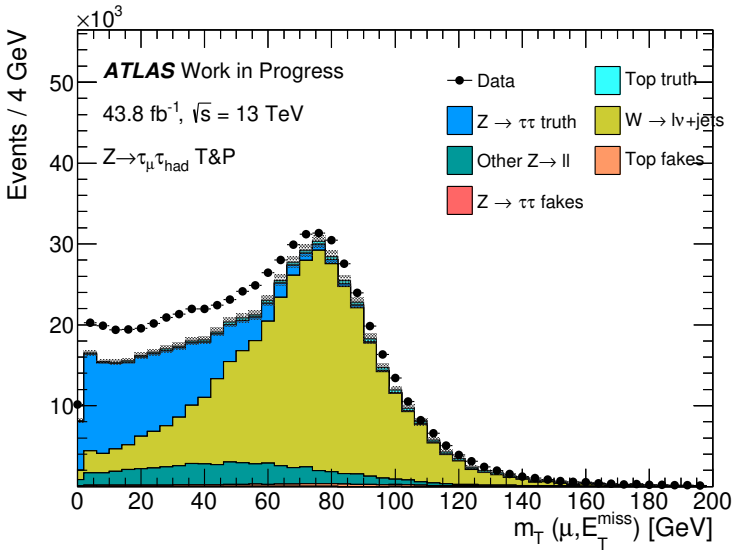


Figure 4.2.: The distribution of the transverse mass between muon and missing transverse energy, $m_T(\mu, E_T^{miss})$ with only preselection cuts applied.

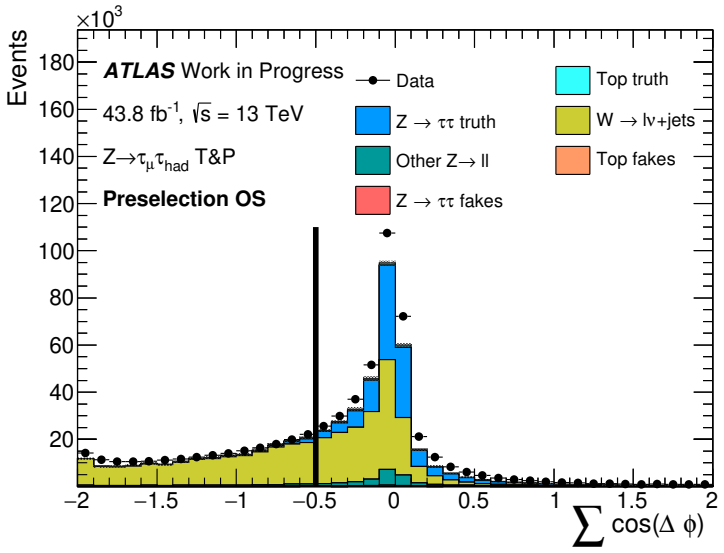


Figure 4.3.: The distribution of $\Sigma \cos \Delta\phi$ with only preselection cuts applied. The signal region is to the right of the vertical line.

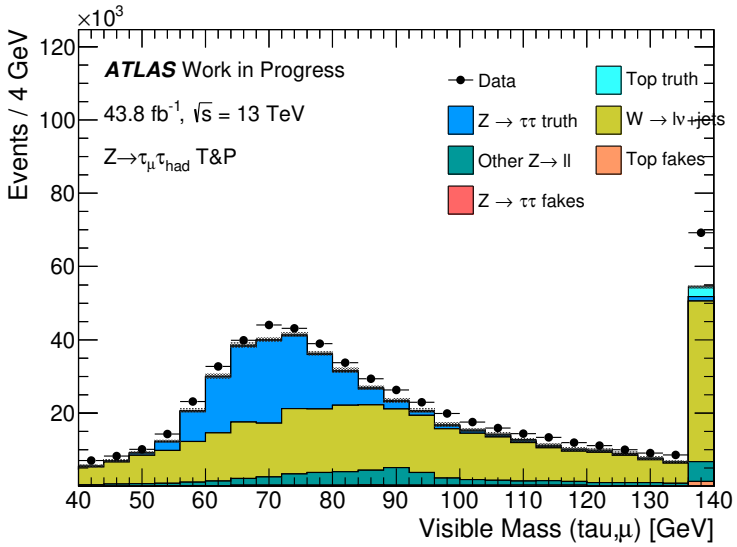


Figure 4.4.: The distribution of the invariant mass of the decay products, $m_{vis}(\tau_{\text{had-vis}}, \mu)$ with only preselection cuts applied.

passing calorimeter and track isolation requirements [52]. The $\tau_{\text{had-vis}}$ probe must satisfy, $p_T \geq 25\,\text{GeV}$, $|\eta| < 2.47$, excluding the transition zone between central region and barrel at $1.37 < |\eta| < 1.52$, must have either one or three associated charged tracks, and must pass at least the loose BDT identification discriminant against jets [51]. No electrons or b-tagged jets are allowed and the muon tag and the $\tau_{\text{had-vis}}$ candidate probe are required to have charge of opposite sign (OS). No tau trigger identification is required at this stage. For the signal region, the following cuts have been found to reduce the background contribution from W + jets events: the transverse mass between muon and missing transverse energy

$$m_T(\mu, E_T^{miss}) = \sqrt{2p_T^{\mu} E_T^{miss}(1 - \cos\Delta\phi(\mu, E_T^{miss}))} < 50\,\text{GeV},$$

the sum of the azimuthal angles of the muon and the $\tau_{\text{had-vis}}$ with the missing energy

$$\Sigma\cos\Delta\phi = \cos\Delta\phi(\tau_{\text{had-vis}}, E_T^{miss}) + \cos\Delta\phi(\mu, E_T^{miss}) > -0.5,$$

and the visible invariant mass

$$45\,\text{GeV} < m_{vis}(\tau_{\text{had-vis}}, \mu) < 80\,\text{GeV}.$$

The motivation for these cuts can be seen in the distributions of these variables at preselection in Figures 4.2, 4.3 and 4.4. The lack of simulated events is because the QCD

Event selection	
Muon tag:	$\tau_{\textbf{had-vis}}$ **probe:**
Medium quality	jet BDT medium
Trigger matched	Muon veto, no overlapping electron
$p_T > 26\,\text{GeV}$	$p_T > 25\,\text{GeV}$, $\lvert q\rvert = 1$
$\lvert\eta\rvert < 2.5$	$\lvert\eta\rvert < 1.37$, $1.52 < \lvert\eta\rvert < 2.47$
track+calo isolation	1 or 3 core tracks
$Z \to \tau\tau \to \mu\tau_{\textbf{had-vis}}3\nu$ **signal region:**	
$m_T(\mu, E_T^{miss}) < 50\,\text{GeV}$	
$\Sigma \cos\Delta\phi > -0.5$	
$45\,\text{GeV} < m_{vis}(\tau_{\text{had-vis}}, \mu) < 80\,\text{GeV}$	

Table 4.1.: Summary of the $Z \to \tau\tau \to \mu\tau_{\text{had-vis}}3\nu$ tag and probe event selections.

Process	MC generator	σ [pb]
$Z \to \tau\tau$	POWHEG-PYTHIA	1950
W + jets	POWHEG-PYTHIA	60200
$Z \to \ell\ell$ $(\ell = e, \nu)$	POWHEG-PYTHIA	3900
Single top and $t\bar{t}$	POWHEG-PYTHIA	526

Table 4.2.: Summary of the MC samples used, along with their event generators and cross sections σ.

background has not yet been modelled at this stage.

4.2. Monte Carlo Samples and Background Modelling

The main background contributions to the $Z \to \tau\tau$ signal are QCD multi-jet events and $W(\to \mu\nu)$ + jets with the jet passing for a $\tau_{\text{had-vis}}$ candidate. Additionally, smaller contributions originate from top and Z + jets events.

The MC samples used for this measurement are listed in Table 4.2. The $Z \to \tau\tau$, $Z \to \ell\ell$ $(\ell = e, \nu)$, single top and $t\bar{t}$ events can be estimated directly from Monte Carlo. For QCD multi-jet and W + jets events the process is more involved and the Background methods used will now be explained.

4.2.1. QCD Estimate

To estimate the QCD contribution, a data-driven approach is used. Since in a true Z boson decay, tag and probe will have have opposite charge, any same sign (SS) events are expected to be predominantly background. The number of events in the signal region $N(SR)$ can be written as follows:

$$N(SR) = N_{signal} + N_{bkg,sym} + N_{bkg,asym}, \tag{4.1}$$

where $N_{bkg,sym}$ is the contribution to the signal region yield from processes where tag and probe have no charge correlation and $N_{bkg,asym}$ have full correlation. The OS contribution to the QCD background can then be estimated from the SS contribution. The OS/SS symmetry is assumed to not be perfect, as the dearth of anti-particles in the initial protons introduces a bias to positively charged events and a correction factor r_{QCD} is thus applied. r_{QCD} is estimated as the ratio between OS and SS events in a control region designed to be enriched in QCD multi-jet events by inverting the muon isolation requirement since any non-isolated muons are expected to be produced in QCD jets. The symmetric background is thus

$$N_{bkg,sym} = r_{QCD} \cdot (N_{SS}(SR) - N_{SS}^{Signal}(SR)), \quad r_{QCD} = \frac{N_{OS}(CR)}{N_{SS}(CR)}, \tag{4.2}$$

where the signal subtraction has been performed. r_{QCD} is calculated separately above and below $p_{\mathrm{T}}(\tau_0) = 50\,\mathrm{GeV}$ with the choice of only two bins due to limited statistical power. The SS p_T distribution of the $\tau_{\text{had-vis}}$ candidate in the control region can be seen in Figure 4.5. The lack of agreement between the data and the MC samples is due to the large number of QCD multi-jet events that are not yet modelled. The uncertainty on r_{QCD} is obtained from varying the muon isolation criteria used to define the control region.

4.2.2. Fake rate correction factor k_W

The k_W correction factors account for a possible mismodelling of the $j \to \tau$ fake rate in simulation. They are calculated separately for OS and SS events in a control region designed to be enriched in W + jets events using the cuts $m_{\mathrm{T}}(\mu, E_{\mathrm{T}}^{miss}) > 60\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{miss} > 30\,\mathrm{GeV}$. The k_W factor is then the ratio of the observed to expected events in this region:

$$k_W^{OS/SS} = \frac{N_{Data}^{OS/SS}(CR)}{N_{MC}^{OS/SS}(CR)}. \tag{4.3}$$

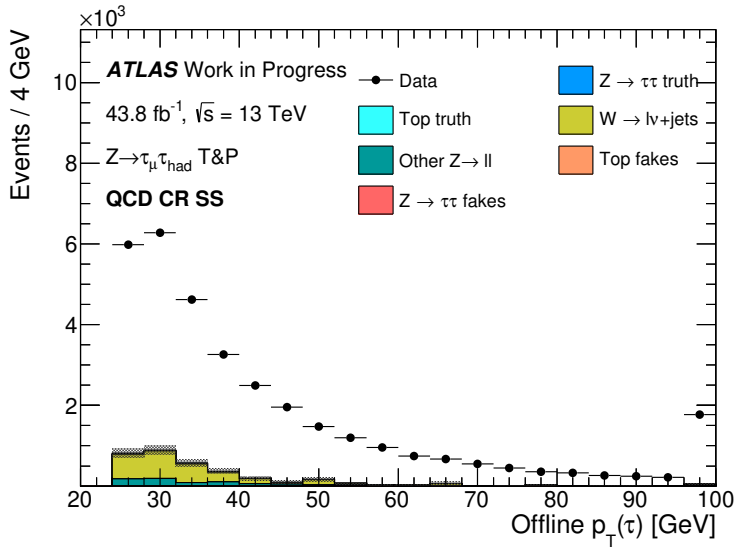


Figure 4.5.: The SS p_T distribution of the $\tau_{\text{had-vis}}$ candidate in the control region. The discrepancy between data and MC is due to QCD multi-jet events.

k_W is used in the estimation of the W + jets background detailed below. Its uncertainty is obtained from varying the cut on $m_{\text{T}}(\mu, E_{\text{T}}^{miss})$ defining the control region.

4.2.3. W + jets Estimate

Due to the small size of the W + jets samples, another data-driven approach is used. The shape of the W + jets channel is taken from data in a different control region, this time defined by changing the $\Sigma \cos \Delta\phi$ cut to $\Sigma \cos \Delta\phi < 0.6$. The transfer factor f_W is defined as

$$f_W = \frac{k_W^{OS} N^{OS}(SR) - k_W^{SS} r_{QCD} N^{SS}(SR)}{k_W^{OS} N^{OS}(CR) - k_W^{SS} r_{QCD} N^{SS}(CR)}, \tag{4.4}$$

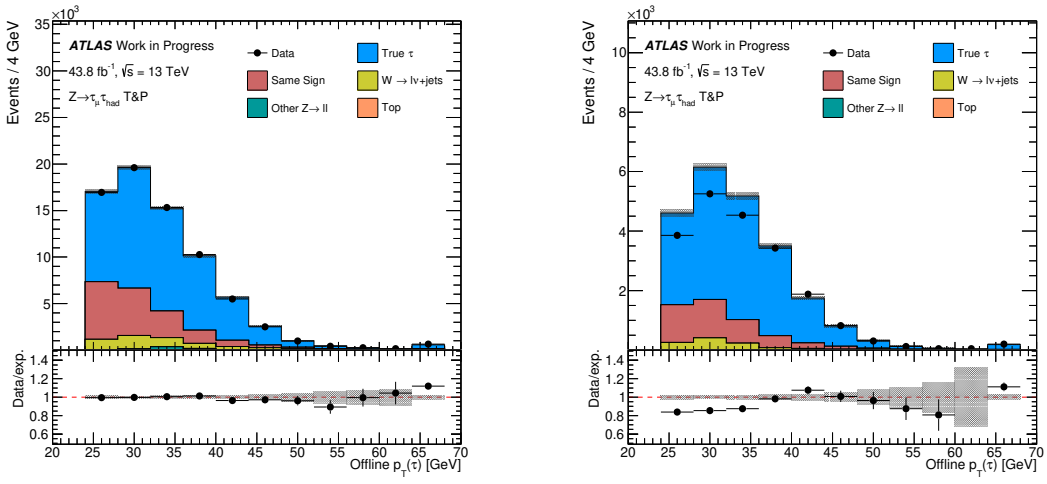
where the yields are taken from the MC samples. f_W is calculated separately below and above $p_{\text{T}}(\tau_0) = 35\,\text{GeV}$ and is applied as a scale factor on the data in the control region to obtain the W + jets estimate. The distribution of $\Sigma \cos \Delta\phi$ after preselection was shown in Figure 4.3. The control region is slightly to the left of the vertical line and the $\Sigma \cos \Delta\phi < -0.6$ cut used for the control region is clearly motivated by the almost total exclusion of the signal in this region.

Contribution	Yield 1-prong	Yield 3-prong
symm. bkg	16794 ± 152	4088 ± 78
τ fakes	5920 ± 230	1160 ± 98
true $\tau_{had,vis}$	50118 ± 297	17487 ± 183
Total Expected	72832 ± 405	22735 ± 222
Data	72477	20522
Data/Exp	0.995	0.903
Purity	68.8%	76.9%

Table 4.3.: Event yields for 1-prong and 3-prong taus. The errors are statistical only.

4.3. Event Yields

With the relevant backgrounds defined the expected yields are now compared to the data taken during 2017 with a total integrated luminosity of $43.8\,\text{fb}^{-1}$. The event yields are shown in Table 4.3 and the p_{T} distribution of the $\tau_{\text{had-vis}}$ candidate before application of a trigger is shown in Figure 4.6 separately for 1- and 3-prong $\tau_{\text{had-vis}}$ decays with the medium tau-ID working point. The symmetric background is obtained using the approach summarised in Equation 4.2 and the misidentified $\tau_{\text{had-vis}}$ candidates (τ fakes) are mostly W + jets scaled via Equation 4.4 with a negligible contribution from Z + jets and top events. In the 1-prong case, 72477 events were measured while approximately 72800 were expected, corresponding to a discrepancy of just 0.5% well within the statistical uncertainty. The signal purity achieved with the aforementioned cuts was 68.8%.
The expected number of 3-prong events is about 22700 of which 76.9% are thought to originate from the $Z \to \tau\tau$ channel. However, only 20522 events were measured, resulting in a 10% excess in simulated events. This mismodelling is subject to investigation in Chapter 5.2 where it is implied that the signal is the cause. The actual purity in the 3-prong channel is consequently assumed to be significantly lower.



(a) 1-Prong

(b) 3-Prong

Figure 4.6.: 1- and 3-prong p_T distributions in the signal region with the medium tau-ID working point applied before application of any trigger.

5. Efficiency and Scale Factor Measurements of the Tau Trigger

5.1. 2017 Efficiencies and Scale Factors

This chapter presents the results of the tau trigger efficiency measurements. Discrepancies between the tau trigger efficiencies obtained with data and Monte Carlo are accounted for using correction factors that can then be provided along with the efficiencies themselves to analysis teams using tau leptons. Later the investigation into issues and attempts to improve the performance are discussed. The data used was recorded in 2017 with a total luminosity of $43.8\,\mathrm{fb}^{-1}$.
The $Z \to \tau\tau \to \mu\tau_{\text{had-vis}}3\nu$ trigger efficiencies measured in this report are defined as the number of events that pass both the trigger and the tau-ID over the number of events to pass the tau-ID:

$$\epsilon(\text{Data}) = \frac{\#\text{ Data-Bkg Evts with tau-ID \& trigger}}{\#\text{ Data-Bkg Evts with tau-ID}} \tag{5.1}$$

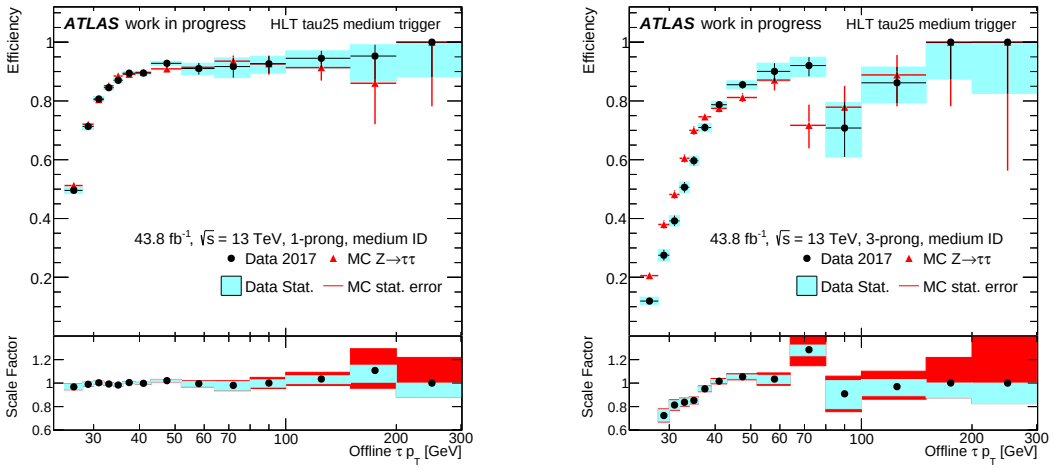
$$\epsilon(\text{MC}) = \frac{\#\text{ Signal Evts with tau-ID \& trigger}}{\#\text{ Signal Evts with tau-ID}}. \tag{5.2}$$

These efficiencies are measured both in background subtracted data and in simulated events and the ratios of both quantities are given as scale factors,

$$\text{SF} = \frac{\epsilon(\text{Data})}{\epsilon(\text{MC})}. \tag{5.3}$$

These scale factors allow analysis teams using τ leptons to account for the discrepancy between observed and simulated trigger efficiencies. Here, the efficiencies are calculated separately for 1- and 3-prong events with respect to offline $\tau_{\text{had-vis}}$ momentum p_T and pseudorapidity η, as well as with respect to the average number of interactions within the detector $\langle\mu\rangle$ known as *pile-up*. The trigger used is the *HLT tau25 medium* trigger which requires an isolated $E_T > 12\,\text{GeV}$ candidate at L1 and a $\tau_{\text{had-vis}}$ with $p_T > 25\,\text{GeV}$ at the HLT. At this stage the uncertainties provided are statistical only.

Figure 5.1 shows the efficiencies binned versus p_T. Both the 1-prong and the 3-prong



(a) 1-Prong (b) 3-Prong

Figure 5.1.: Efficiencies in 1- and 3-prong versus p_T with the medium tau-ID working point applied.

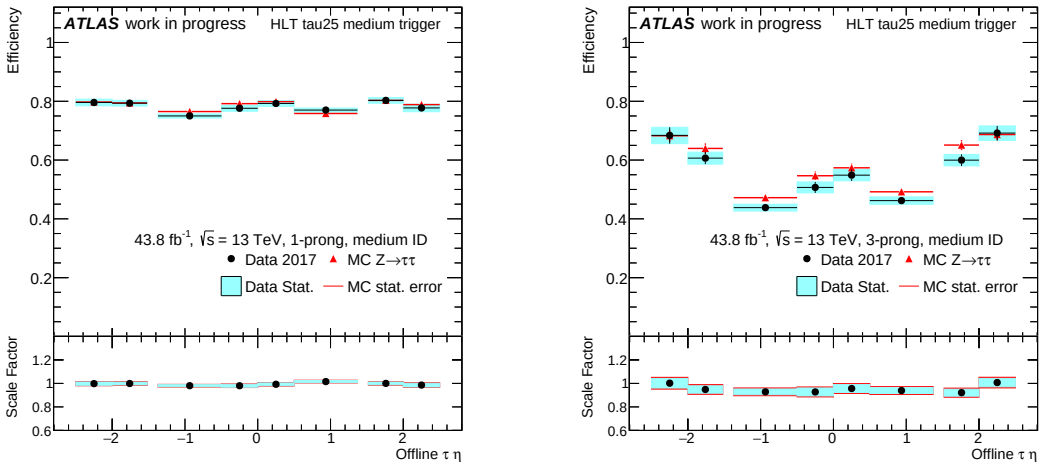
channel show an approximately linear increase in efficiency up to 40 GeV, above which they remain constant at around 90%. The scale factors are stable in the 1-prong case. For 3-prong the difference in performance of data and simulation lead to scale factors significantly less than 1. This is discussed later in the chapter.

The dependency of the efficiency versus pseudorapidity η and pile-up can be seen in Figures 5.2 and 5.3 respectively. 1-prong shows good modelling and constant efficiencies across the entire detector region, while 3-prong suffers from reduced performance towards the ends of the central barrel. In both channels a decrease in efficiency can be seen as pile-up increases, likely a consequence of the increased activity in the isolation regions used in the definition of the trigger.

5.2. Investigation of 3-Prong Mismodelling

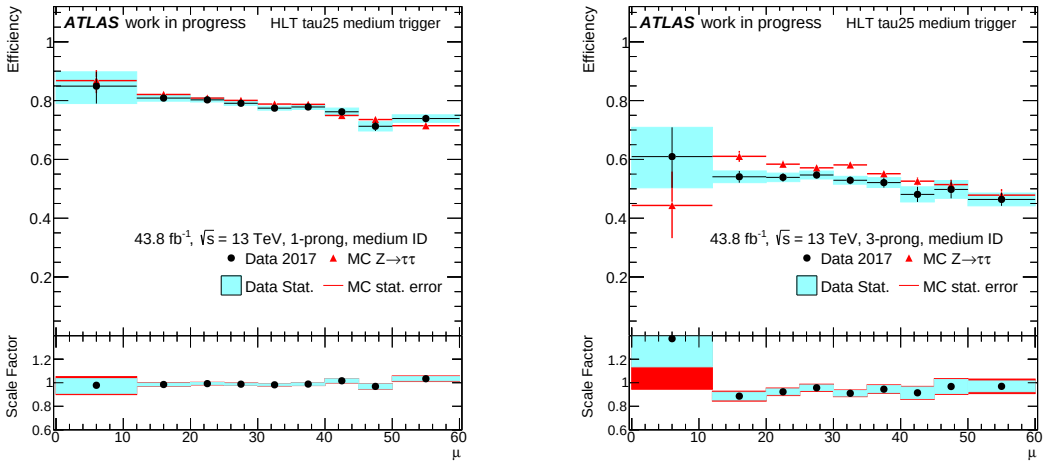
The measurements detailed previously in Figure 5.1 show a discrepancy between the efficiencies predicted by simulation and those from data. Furthermore, there is an excess of simulated events in the signal region compared to data that is present even before the HLT tau25 medium trigger is applied as shown in Figure 5.4.

As a first step in the investigation, the efficiency measurements were repeated using combined data from 2015 and 2016 with an integrated luminosity of 36.1 fb^{-1} (Figure 5.5).



(a) 1-Prong (b) 3-Prong

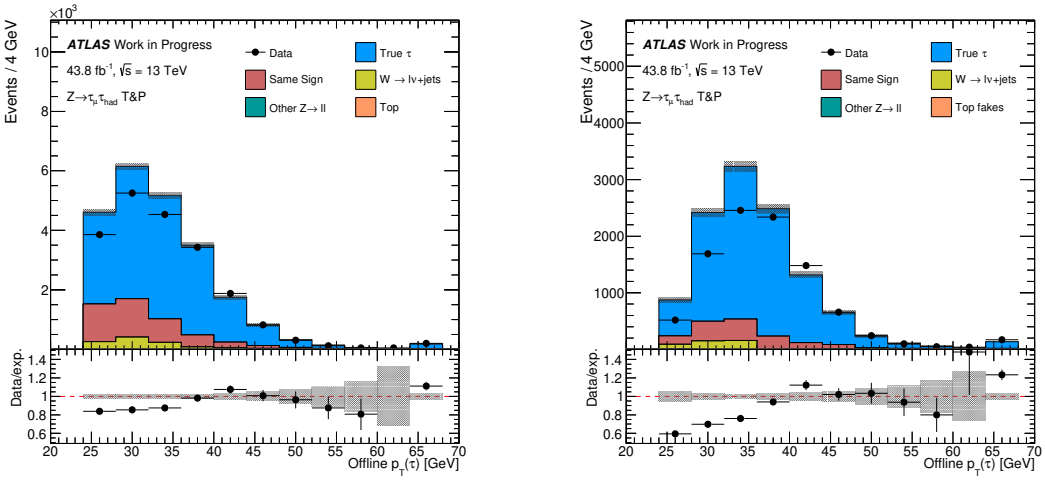
Figure 5.2.: Efficiencies in 1- and 3-prong versus η with the medium tau-ID working point applied.



(a) 1-Prong (b) 3-Prong

Figure 5.3.: Efficiencies in 1- and 3-prong versus pile-up with the medium tau-ID working point applied.

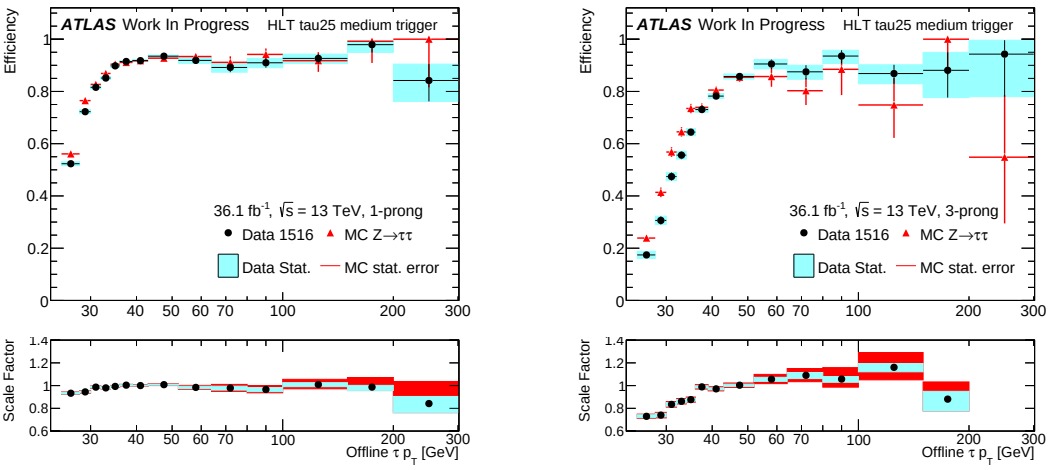
The mismodelling in the 3-prong channel was consequently found to not just be an artifact of the 2017 simulation. The following features were then investigated: A two-dimensional binning of the efficiencies in p_{T} and η, the W + jets background modelling and the input variables used in the BDT for the tau identification.



(a) Before trigger

(b) After trigger

Figure 5.4.: p_T distribution in the 3-prong signal region before and after applying the HLT tau25 medium trigger



(a) 1-prong

(b) 3-prong

Figure 5.5.: Efficiencies in 1- and 3-prong versus p_T with the medium tau-ID working point applied for combined 2015-2016 data.

5.2.1. p_{T} versus η Efficiencies

As the next step, the efficiency measurements were repeated but with a two-dimensional binning in $\tau_{\text{had-vis}}$ transverse momentum p_T and the absolute of the pseudorapidity $|\eta|$. This was done in order to see if the mismodelling occurred for a specific combination of detector region and momentum. A coarser binning than before was used to ensure

that statistical fluctuations are minimised. Once the low statistics in the high-p_T, high-η regions have been accounted for, no local mismodelling of any significance is observed, requiring further investigation.

While the two dimensional measurements ultimately did not resolve the 3-prong mismodelling issue, they can still be used to investigate the trigger modelling and to possibly provide more detailed scale factors to future analyses. For this reason Figures 5.6 and 5.7 show the $|\eta|$ dependence of the efficiency measurements in slices of p_T for the 1- and 3-prong channels respectively. The slices used are $p_T = [25, 32]\,\text{GeV}$, $p_T = [32, 40]\,\text{GeV}$, $p_T = [40, 54]\,\text{GeV}$ and $p_T = [54, 300]\,\text{GeV}$.

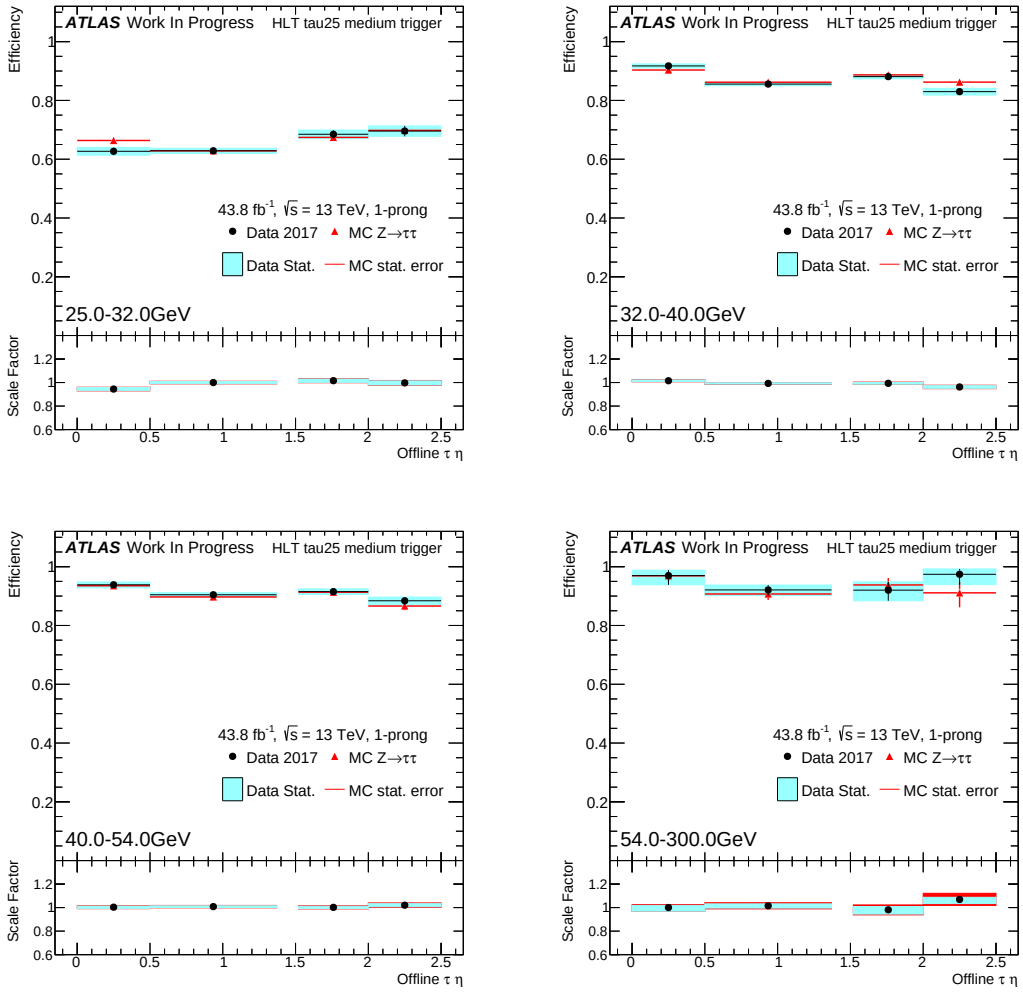


Figure 5.6.: $|\eta|$ dependence of 1-prong efficiencies in slices of p_T. From top left to bottom right: $p_T = [25, 32]]\,\text{GeV}$, $p_T = [32, 40]\,\text{GeV}$, $p_T = [40, 54]\,\text{GeV}$, $p_T = [54, 300]\,\text{GeV}$

Two-dimensional scale factor plots are displayed in the appendix B.

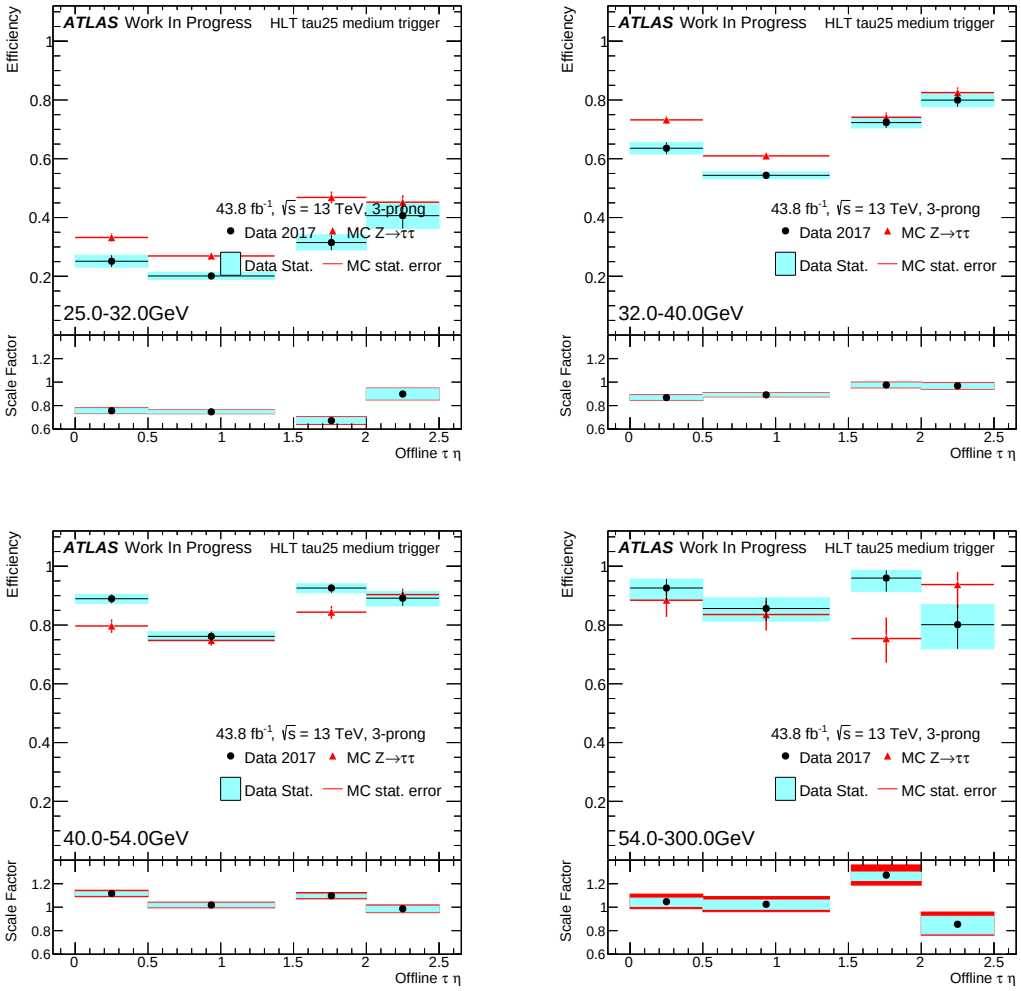


Figure 5.7.: $|\eta|$ dependence of 3-prong efficiencies in slices of p_T. From top left to bottom right: $p_T = [25, 32]]\,\text{GeV}$, $p_T = [32, 40]\,\text{GeV}$, $p_T = [40, 54]\,\text{GeV}$, $p_T = [54, 300]\,\text{GeV}$

5.2.2. W + jets Modelling

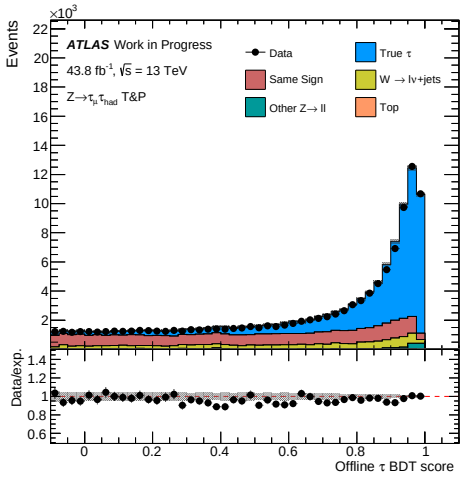
Another attempt to address the issue of 3-prong mismodelling was to improve the modelling of the W + jets background. The results are here relegated to the Appendix A.

5.2.3. BDT Input Variable Distributions

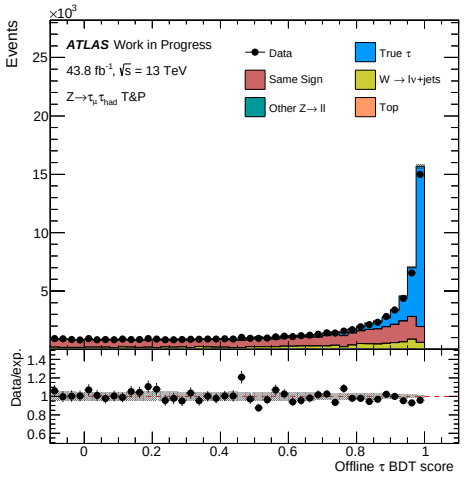
The final attempt to identify the source of the 3-prong mismodelling was to consider the input variables used in the tau-ID BDT algorithm described in Section 3.5.2. This was motivated by comparing the $\tau_{\text{had-vis}}$ p_T distributions of the signal region with the not-loose and the loose tau-ID working point (Figure 5.8) to the previous medium tau-ID working

point seen in Figure 4.6. The distribution of the BDT score itself with no tau-ID applied is also shown. The excess in simulated events is only apparent after a working point is chosen, implying that the BDT algorithm may have been trained using variables that do not properly model 3-prong $\tau_{\text{had-vis}}$ decays. The variables described previously in Section 3.5.2 were consequently evaluated for discrepancies in the shape of the data and simulated distributions.

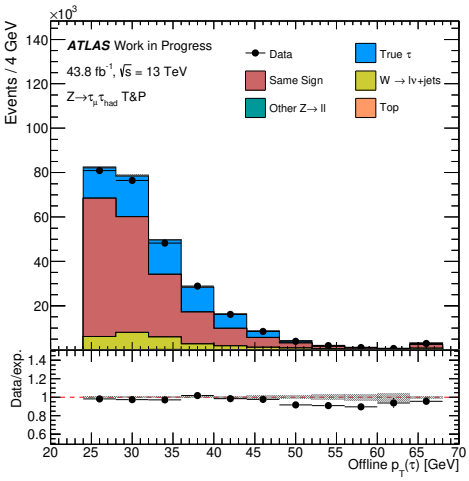
Figures 5.9 to 5.12 show the four variables with the most noticeable shape difference. These are the leading track momentum fraction $f^{-1}_{\text{lead track}}$, the central energy fraction f_{cent}, the ratio of EM energy to track momentum $f^{\text{EM}}_{\text{track}}$, and the ratio of track-plus-EM-system to p_T $P_T^{\text{EM+track}}/P_T$. The remaining BDT input variables are shown in the Appendix C. In general, the four variables with the worst performance show shape mismodelling where the peak in the simulated events is shifted relative to the data. These variables are also comprised of ratios of quantities including the energy deposited in the electromagnetic calorimeter. Since 1-prong decays are not affected it seems possible that energy measurements on the close proximity deposits from multiple tracks are not simulated correctly. The training of the τ classifier for 3-prong decays is hindered by the lack of events relative to 1-prong decays, as well as by resolution issues from the close proximity tracks. This issue should be investigated in the future to see if it requires additional attention. The development of hadronic τ lepton identification using a recurrent neural network (RNN) [56] instead of the BDT is also noteworthy. In the meantime the 3-prong mismodelling illustrates the necessity of providing scale factors to analysis teams using τ leptons in the final state.



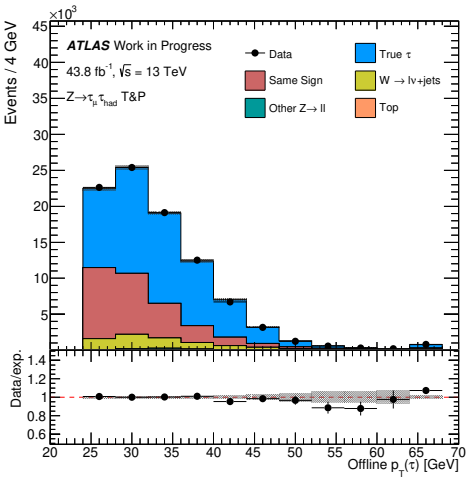
(a) BDT score, 1-prong



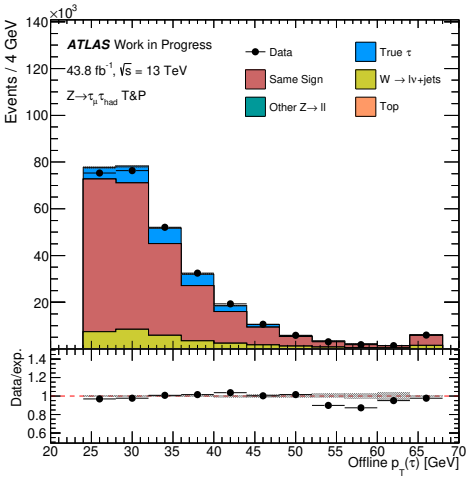
(b) BDTscore, 3-prong



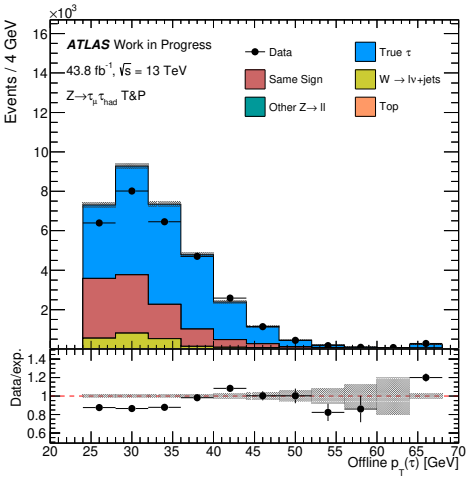
(c) 1-prong, not-loose-ID



(d) 1-prong, loose-ID

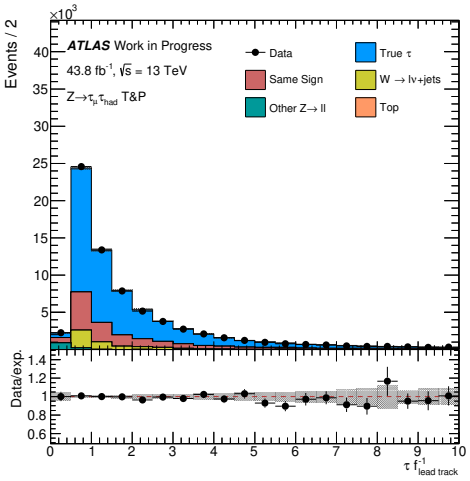


(e) 3-prong, not-loose-ID

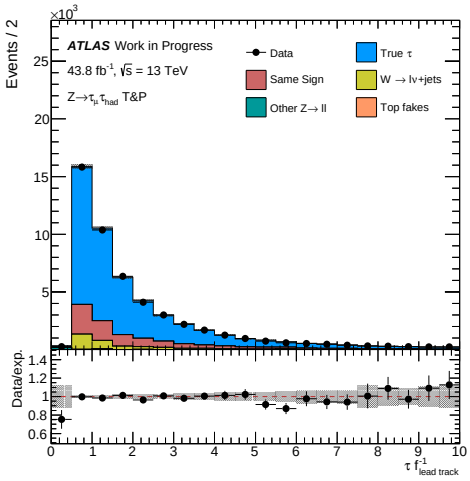


(f) 3-prong, loose-ID

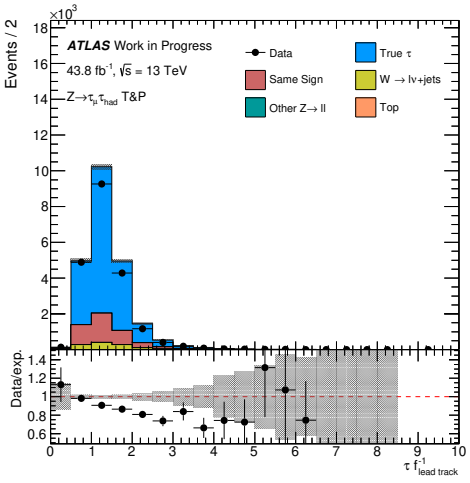
Figure 5.8.: BDT distribution with no tau-ID working point and $\tau_{\text{had-vis}}$ p_T distributions with not-loose and loose tau-ID working points.



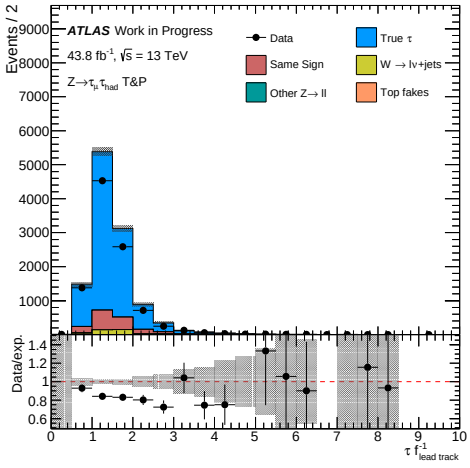
(a) $f^{-1}_{\text{lead track}}$, 1-prong, before trigger



(b) $f^{-1}_{\text{lead track}}$, 1-prong, after trigger

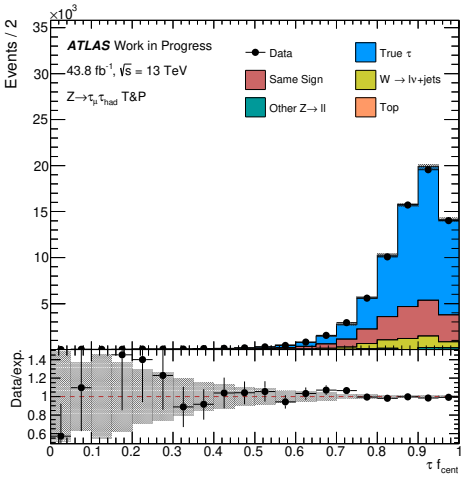


(c) f_{cent}, 3-prong, before trigger

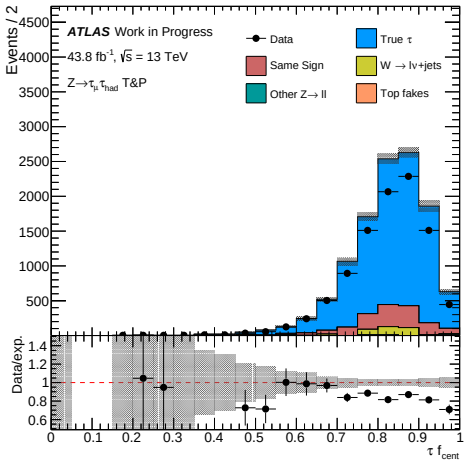


(d) f_{cent}, 3-prong, after trigger

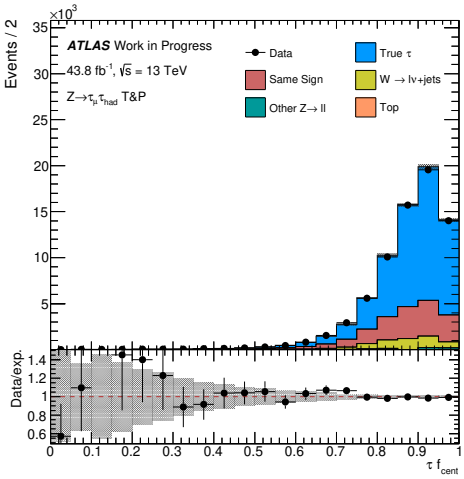
Figure 5.9.: The distributions of the leading track momentum fraction $f^{-1}_{\text{lead track}}$ used in the tau-ID BDT discriminant, shown here with the medium working point for both 1-prong and 3-prong decays before and after application of the *HLT tau25 medium* trigger.



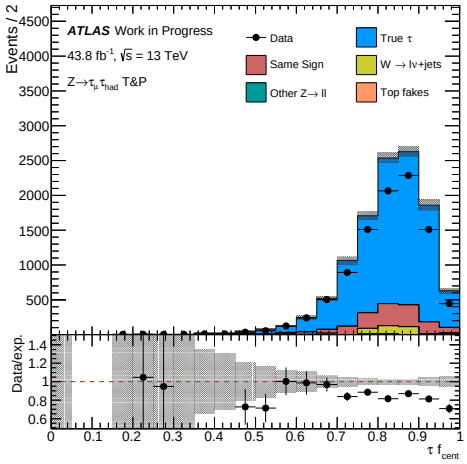
(a) $f^{-1}_{\text{lead track}}$, 1-prong, before trigger



(b) $f^{-1}_{\text{lead track}}$, 1-prong, after trigger

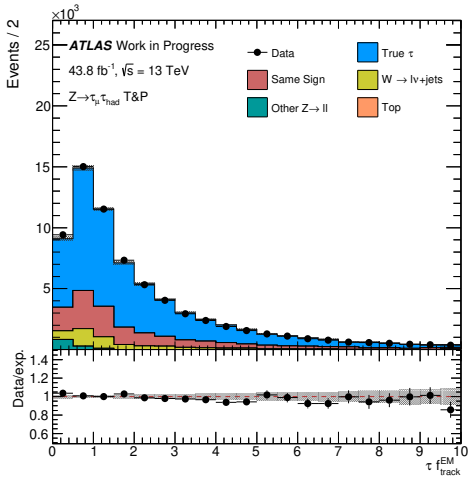


(c) f_{cent}, 3-prong, before trigger

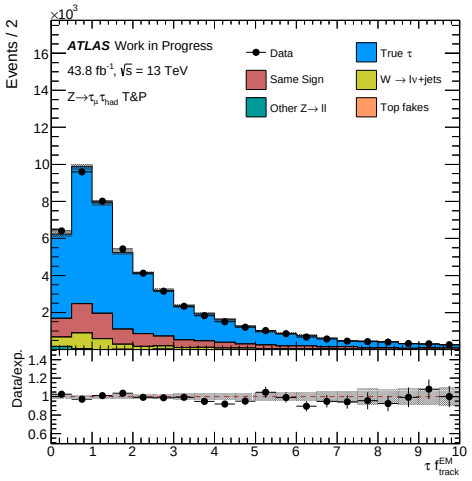


(d) f_{cent}, 3-prong, after trigger

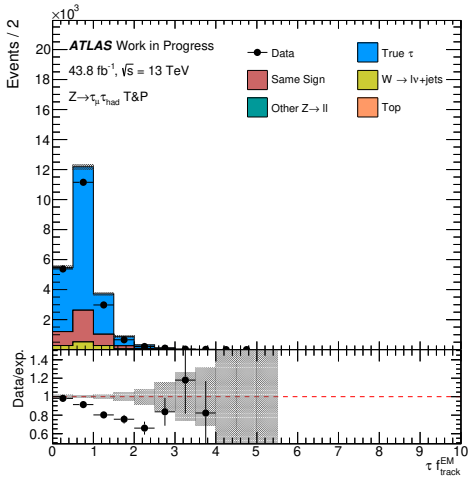
Figure 5.10.: The distributions of the central energy fraction f_{cent} used in the tau-ID BDT discriminant, shown here with the medium working point for both 1-prong and 3-prong decays before and after application of the *HLT tau25 medium* trigger.



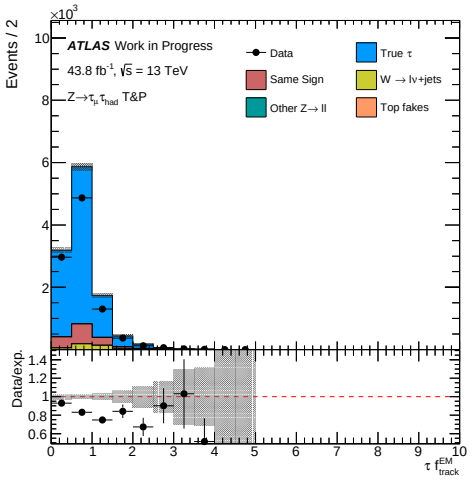
(a) $f^{-1}_{\text{lead track}}$, 1-prong, before trigger



(b) $f^{-1}_{\text{lead track}}$, 1-prong, after trigger

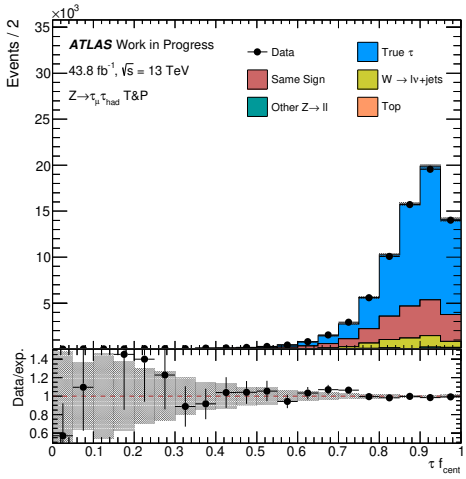


(c) f_{cent}, 3-prong, before trigger

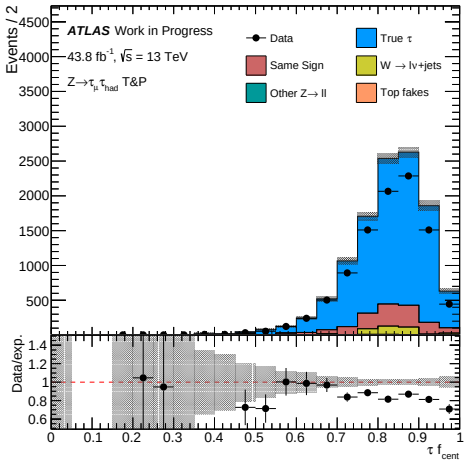


(d) f_{cent}, 3-prong, after trigger

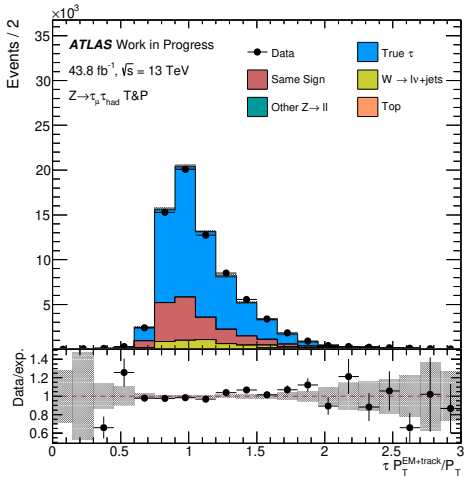
Figure 5.11.: The distributions of the ratio of EM energy to track momentum $f^{\text{EM}}_{\text{track}}$ used in the tau-ID BDT discriminant, shown here with the medium working point for both 1-prong and 3-prong decays before and after application of the *HLT tau25 medium* trigger.



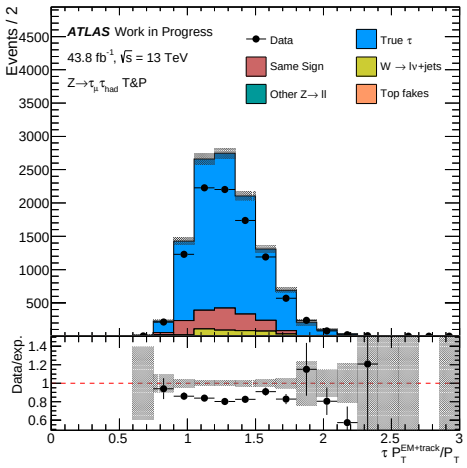
(a) $f^{-1}_{\text{lead track}}$, 1-prong, before trigger



(b) $f^{-1}_{\text{lead track}}$, 1-prong, after trigger



(c) f_{cent}, 3-prong, before trigger



(d) f_{cent}, 3-prong, after trigger

Figure 5.12.: The distributions of the ratio of track-plus-EM-system to p_T $P_T^{\text{EM+track}}/P_T$ used in the tau-ID BDT discriminant, shown here with the medium working point for both 1-prong and 3-prong decays before and after application of the *HLT tau25 medium* trigger.

6. The CP Measurement of the Higgs Boson in VBF Production for the $H \to \tau\tau$ Final State

The next two chapters of this thesis describe the CP measurement of the Higgs boson in VBF production for the $H \to \tau\tau$ fully hadronic final state. This chapter details the methods employed in the analysis with the 2015-2016 dataset at $\sqrt{s} = 13\,\text{TeV}$ with $\mathcal{L} = 36.1\,\text{fb}^{-1}$ [4], while the next chapter presents some early results obtained with the full Run-2 dataset with $140\,\text{fb}^{-1}$. The analysis uses the observation of the mean of a CP-odd Optimal Observable to perform a direct test of CP invariance in Higgs boson VBF production.

6.1. MC Samples and Event Preselection

This section details the Monte Carlo simulation samples used, as well as the cuts applied at the preselection level in the case of the fully hadronic sub-channel. At this stage, all $H \to \tau\tau$ production mechanisms are regarded as the signal. Consequently this includes production via gluon-gluon-fusion (ggF) as well as vector-boson-fusion (VBF) which is later selected for. Smaller contributions originate from the production via Higgs boson radiation (ZH, WH). Production of a Higgs boson in association with two top quarks (ttH) is not included as the cross section renders this mechanism negligible.
The background consists mainly of $Z \to \tau\tau$ events and QCD multi-jet events, the former due to the similar decay of a boson to two τ leptons, and the latter for its high cross section and ability of jets to mimic hadronically decaying τ leptons. Smaller contributions to the background consist of W + jets and $Z \to \ell\ell$ ($\ell = e, \mu$) events, single top and $t\bar{t}$ events and diboson VV ($V = W, Z$) events.
A summary of the MC samples, the generators used and their cross sections can be seen in Table 6.1. The cuts applied at the fully hadronic preselection stage are listed in Table 6.2. Events are required to have at least one primary vertex and pass the event cleaning for

Process	MC generator	σ [pb]
ggF, $H \to \tau\tau$	POWHEG-PYTHIA	12.1
VBF, $H \to \tau\tau$	POWHEG-PYTHIA	0.236
WH	POWHEG-PYTHIA	0.0857
ZH	POWHEG-PYTHIA	0.0552
$Z \to \tau\tau$	SHERPA	13000
W + jets	SHERPA	185000
$Z \to \ell\ell$ ($\ell = e, \mu$)	SHERPA	26900
Single top and $t\bar{t}$	POWHEG-PYTHIA	233
VV ($V = W, Z$)	SHERPA	129

Table 6.1.: Summary of the MC samples used, along with their event generators and cross sections σ. $Z \to \tau\tau$ and $Z \to \ell\ell$ ($\ell = e, \mu$) events include both QCD and electroweak production mechanisms. ttH simulation samples were not yet available.

Preselection Cuts
At least one primary vertex, loose event cleaning
No muons or electrons
$E_{\mathrm{T}}^{\mathrm{miss}} > 20\,\mathrm{GeV}$ and $0.8 < \Delta R_{\tau\tau} < 2.5$
$p_{T,\tau_{lead}} > 40\,\mathrm{GeV}$ and $p_{T,\tau_{sub-lead}} > 30\,\mathrm{GeV}$
$p_{T,j_{lead}} > 70\,\mathrm{GeV}$ and $\lvert\eta_{j0}\rvert < 3.2$
$p_{T,j_{sub-lead}} > 15\,\mathrm{GeV}$
$0.1 <$ ditau_coll_approx < 1.4
$\lvert q\rvert = 1$ and $q \times q = -1$
τ decay 1 or 3 prongs
Successful solution to MMC Fit
Leading jet matched at L1 level
Both tau leptons trigger matched

Table 6.2.: List of cuts used at the fully hadronic preselection stage.

jets [57]. A leading and sub-leading τ lepton with the loose tau ID criteria [51] is required with $p_{T,\tau_{lead}} > 40\,\mathrm{GeV}$ and $p_{T,\tau_{sub-lead}} > 30\,\mathrm{GeV}$ respectively. These τ leptons must have opposite charge, have 1 or 3 prongs, and be separated by $0.8 < \Delta R_{\tau\tau} < 2.5$, consistent with the decay from a Higgs boson. A jet with $p_{T,j_{lead}} > 70\,\mathrm{GeV}$ and $|\eta_{j_{lead}}| < 3.2$ must be found and matched at L1. A second jet with $p_{T,j_{sub-lead}} > 15\,\mathrm{GeV}$ is also required. Both τ leptons must pass the HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo trigger. Finally, the Missing Mass Calculator (MMC) fit which uses the visible τ lepton remnants to estimate the invariant mass of the ditau system must yield a successful solution [58]. The distribution of the MMC variable after preselection for the fully hadronic sub-channel can be seen in Figure 6.1. The data has been blinded within the range $100\,\mathrm{GeV} \leq M_{MMC}(\tau\tau) \leq 150\,\mathrm{GeV}$ around the Higgs boson mass to avoid influencing the

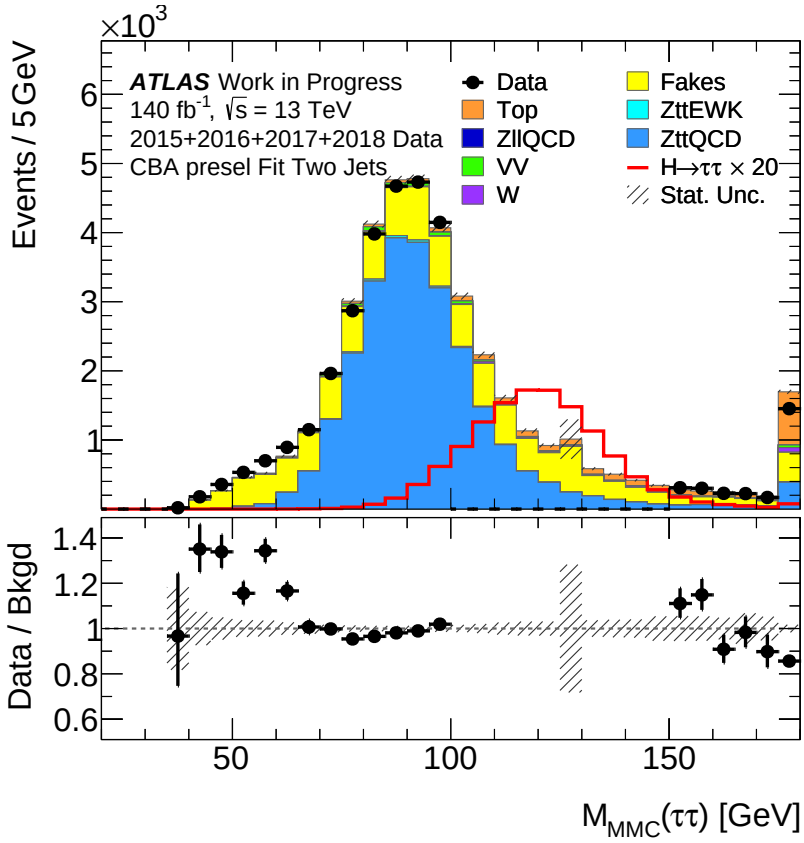


Figure 6.1.: The distribution of the $M_{MMC}(\tau\tau)$ variable in the fully hadronic preselection region. The MMC fit attempts to reconstruct the invariant mass of the ditau system from the visible τ lepton remnants. Note that the overlaid $H \to \tau\tau$ signal has been scaled by a factor of 20 to aid with visibility.

later analysis based on the distribution of the data. The distribution shows the large background from $Z \to \tau\tau$ decays centered around the Z boson mass.

6.2. Optimal Observable Generation and $\tilde{d}$ Reweighting

A test of CP invariance can be done almost model independently by performing a measurement on the mean of a CP-odd observable $\langle \mathcal{O}_{\mathrm{CP}} \rangle$. If the mean vanishes, then CP invariance holds. A result incompatible with $\langle \mathcal{O}_{\mathrm{CP}} \rangle = 0$ would be a clear sign of CP violation. The *optimal observable* $\mathcal{OO}$ chosen here combines seven phase-space variables into a single observable, and has been shown to have the highest sensitivity to small values of $\tilde{d}$ [3, 59, 60]. These seven phase-space variables result from the twelve given by the four-momenta of the Higgs boson and outgoing-jets after introduction of the five constraints

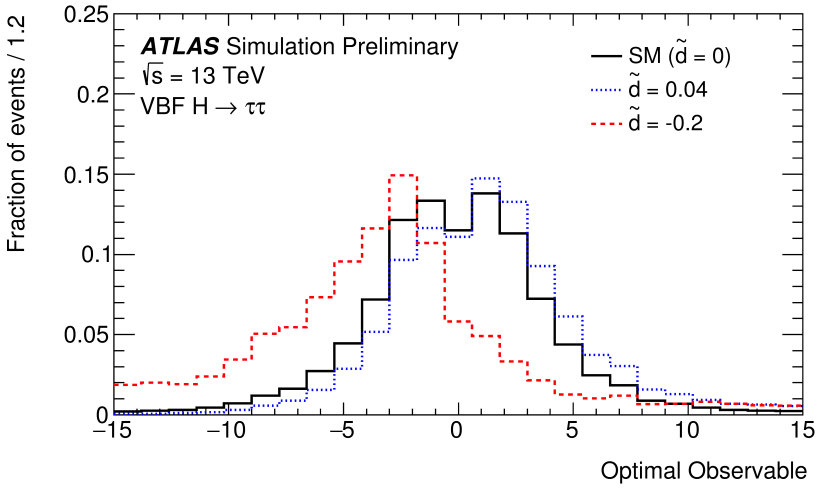


Figure 6.2.: The normalised distribution of the *Optimal Observable* for three different $\tilde{d}$ values [4].

given by the fixed Higgs boson mass, the negligible jet masses and the conservation of momentum in the transverse plane. Here, only the first order *Optimal Observable* $\mathcal{OO}_1$ proportional to $\tilde{d}$ is utilised, and the second order *Optimal Observable* $\mathcal{OO}_2$ which is proportional to $\tilde{d}^2$ is neglected. The former can be written as the ratio of the second term in the matrix element from Section 2.5 to the Standard Model contribution:

$$\mathcal{OO}_1 = \frac{2\mathrm{Re}(\mathcal{M}^*_{\mathrm{SM}}\mathcal{M}_{\text{CP-odd}})}{|\mathcal{M}_{\mathrm{SM}}|^2}$$

and the latter as the ratio of the third term in the equation to the SM contribution:

$$\mathcal{OO}_2 = \frac{|\mathcal{M}_{\text{CP-odd}}|^2}{|\mathcal{M}_{\mathrm{SM}}|^2}.$$

A distribution of this first-order *Optimal Observable* for a VBF $H \to \tau\tau$ sample at $\sqrt{s} = 13\,\mathrm{TeV}$ can be seen in Figure 6.2 for three different $\tilde{d}$ values. In the Standard Model case of $\tilde{d} = 0$, the distribution of $\mathcal{OO}$ is completely symmetric with a mean of 0. A non-zero $\tilde{d}$ leads to an asymmetric distribution and a mean that is shifted in the direction corresponding to the sign of $\tilde{d}$.

The calculation of the *Optimal Observable* is done on an event by event basis using the leading-order matrix elements computed by HAWK [61]. This requires the four-momenta of the Higgs boson and of the two accompanying jets as well as the momentum fractions of the partons from the initial jets $x_{1/2}$. The latter can be reconstructed via energy-

momentum conservation in the Hjj system [28] as:

$$x_{1/2}^{\text{reco}} = \frac{m_{Hjj}}{\sqrt{s}} e^{\pm y_{Hjj}}$$

Here, m_{Hjj} is the invariant mass of the sum of the Higgs boson and outgoing jet four-momenta, and y_{Hjj} is the rapidity of the same system. The final matrix elements are obtained by summing over all possible flavour combinations $ij \to klH$ of the initial and final state partons weighted using the relevant leading-order parton distribution functions (PDFs) $f(x)$ [62]:

$$2\text{Re}(\mathcal{M}_{\text{SM}}^{*}\mathcal{M}_{\text{CP-odd}}) = \sum_{i,j,k,l} f_i(x_1)f_j(x_2) \times 2\text{Re}((\mathcal{M}_{\text{SM}}^{ij\to klH})^{*}\mathcal{M}_{\text{CP-odd}}^{ij\to klH}),$$
$$|\mathcal{M}_{\text{SM}}|^2 = \sum_{i,j,k,l} f_i(x_1)f_j(x_2)|\mathcal{M}_{\text{SM}}^{ij\to klH}|^2.$$

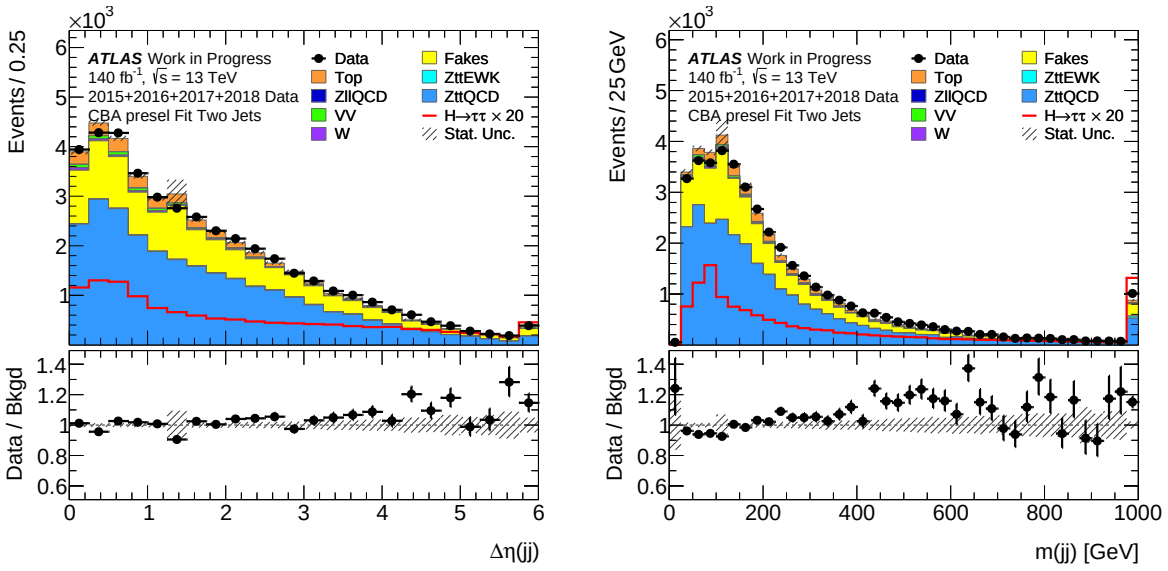
The CP violating VBF $H \to \tau\tau$ signal samples with non-zero $\tilde{d}$ values needed for this analysis are obtained from the Standard Model signal sample via a $\tilde{d}$-dependent reweighting of the events. The weight expression is

$$w(\tilde{d}) = \frac{|\mathcal{M}|^2}{|\mathcal{M}_{\text{SM}}|^2} = 1 + \tilde{d} \cdot \frac{2\text{Re}(\mathcal{M}_{\text{SM}}^{*}\mathcal{M}_{\text{CP-odd}})}{|\mathcal{M}_{\text{SM}}|^2} + \tilde{d}^2 \cdot \frac{|\mathcal{M}_{\text{CP-odd}}|^2}{|\mathcal{M}_{\text{SM}}|^2}.$$

The linear and quadratic coefficients in $\tilde{d}$ can thus be calculated at the same level as the *Optimal Observable* on an event-by-event basis, allowing for the easy access to the event weights needed to produce a signal with any level of CP violation.

6.3. Background Modelling

All but one of the background processes are modelled using Monte Carlo (MC) simulations with the samples listed in Table 6.1. However, the second largest contributor to the background, after $Z \to \tau\tau$ decays is from QCD multi-jet events faking τ leptons that cannot be usefully modelled using Monte Carlo. As in the Tag-and-Probe method described previously, a data-driven approach is needed.
Since the kinematic distributions of QCD events are not expected to depend on the charge of the reconstructed fake τ leptons, a control region (nOS CR) is defined from the preselection by inverting the opposite sign requirement on the τ lepton candidates, $q(\tau_0) \times q(\tau_1) \neq -1$, and expanding the cut on the number of tracks to allow for the selection of 2-prong τ lepton candidates. The resulting region is heavily dominated by



(a) Jet separation $\Delta\eta_{jj}$

(b) Invariant mass of the jet system M_{jj}

Figure 6.3.: The distribution of the $\Delta\eta_{jj}$ and M_{jj} variables used to define the fully hadronic VBF region.

QCD mult-jet events with a small contamination of true τ leptons from $Z \to \tau\tau$ events. The latter are subtracted from the data to create a fake template. The normalisation of this template is expected to be left free in the final analysis, however for comparisons in the preselection region, it is normalised by fitting to the $\Delta\eta_{\tau\tau}$ distribution.
Uncertainties on the background modelling are not yet included in this thesis but can be obtained for the extrapolation from the preselection to the nOS CR as well as for the subtraction of the true τ lepton contaminants in the nOS CR.

6.4. VBF BDT and Negative Log-Likelihood Fit

In this section, an overview of the likely analysis method for the full Run-2 dataset is detailed, based on the procedure for the 2015-2016 dataset at $\sqrt{s} = 13\,\text{TeV}$ with $\mathcal{L} = 36.1\,\text{fb}^{-1}$ [4].
The precise strategy differs depending on the decay mode of the two τ leptons. In the fully hadronic case, a VBF region is defined by requiring a jet separation of $\Delta\eta_{jj} > 3.0$ and an invariant mass of the jet system $M_{jj} > 300\,\text{GeV}$. The motivation for these choices can be seen in Figure 6.3. From here, boosted decision trees (BDT) separate the VBF signal from the background and the other Higgs boson production modes. The signal region is then defined by a channel-dependent cut on the BDT score [28] and the *Optimal Observable*

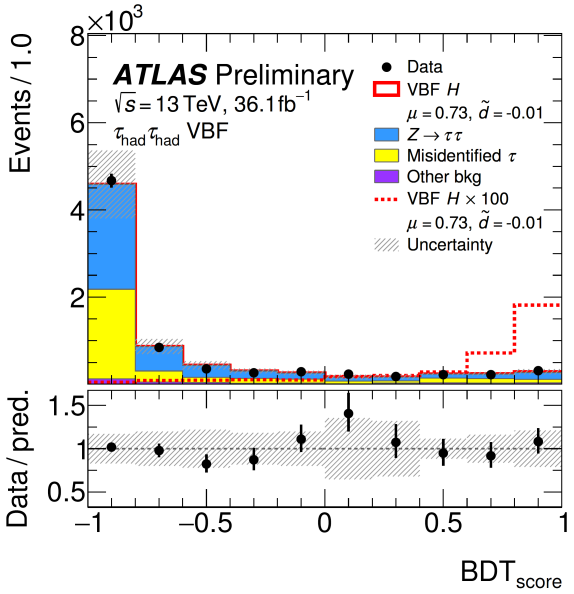


Figure 6.4.: The distribution of the BDT score in the fully hadronic VBF region using the 2015-2016 dataset with $\mathcal{L} = 36.1\,\text{fb}^{-1}$ [4]. For this previous analysis a cut of BDT > 0.87 was chosen to define the signal region.

can then be used to probe CP violation in this region. The background modelling of the *Optimal Observable* is validated using dedicated channel-specific control regions. Figure 6.4 shows the distribution of the BDT score in the fully hadronic VBF region from the previous analysis using the 2015-2016 dataset. A cut of BDT > 0.87 was chosen to define the signal region.

A maximum likelihood fit is performed on the *Optimal Observable* distribution for each decay mode in order to obtain the best estimate of the parameter $\tilde{d}$. This binned likelihood function, $\mathcal{L}(\mathbf{x}; \mu, \theta)$ is a function of the data $\mathbf{x}$, the signal strength μ, which is the ratio of the measured cross section to the SM prediction, and additional nuisance parameters. It is given as the product of the Poisson probability for each bin in the *Optimal Observable* distribution assuming an underlying model of signal plus background. This likelihood is evaluated multiple times with the different signal templates described in Section 6.2, each modelling a different $\tilde{d}$ hypothesis, and then adapted to construct a negative log-likelihood (NLL) curve. The best estimate for the value of $\tilde{d}$ can be read from the minimum of this curve NLL_{min}, and the associated uncertainty at 68% confidence level (CL) can be read from the points at which $\Delta\text{NLL} = \text{NLL} - \text{NLL}_{\text{min}} = 0.5$. The observed and expected ΔNLL curve for the Run-1 analysis can be seen in Figure 6.5 for the combination of all decay channels. The expected values are obtained from an Asimov dataset using the best-fit signal strength parameter $\mu = 1.55$. Using the interpretation described above, it

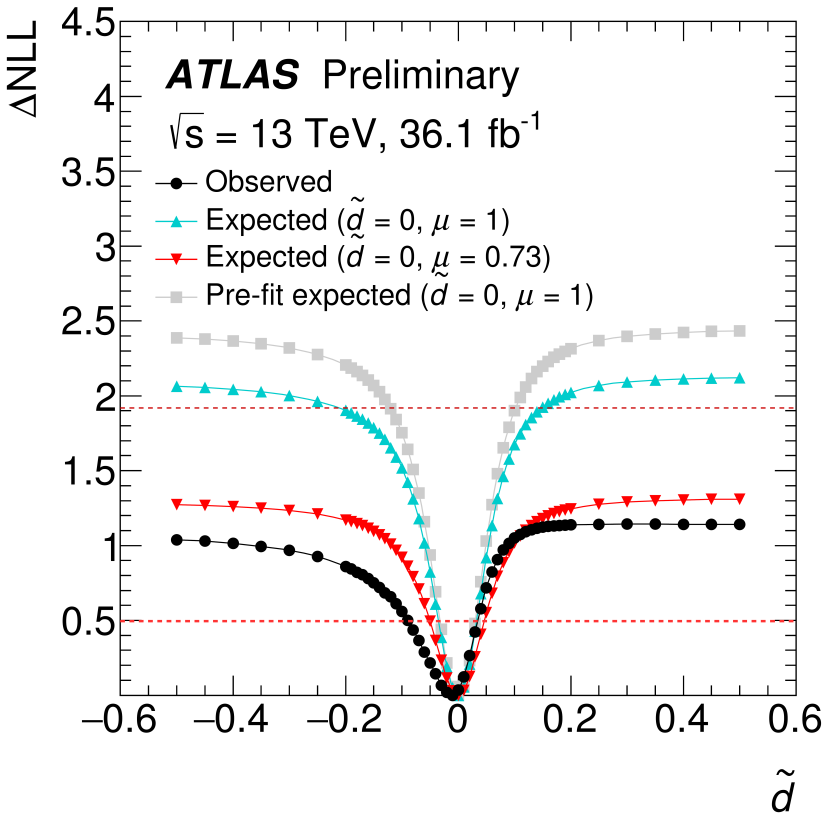


Figure 6.5.: The ΔNLL curve for the previous analysis with the 2015-2016 dataset. Shown are the expected values as well as the observed values for the combination of decay channels. [4].

can be seen that the final reported value for $\tilde{d}$ in the publication was $-0.090 < \tilde{d} < 0.035$ at 68% CL [28].

7. Results of the CP Measurement of the Higgs Boson in VBF Production for the $H \to \tau\tau$ Final State

This chapter presents some preliminary results of the $H \to \tau\tau$ CP investigation in the fully hadronic channel. It discusses the event yields at preselection level, analyses the modelling of important kinematic variables and examines the effect of various $\tilde{d}$ values on the signal sample. The full Run-2 dataset is used, which was recorded from 2015 to 2018 with the ATLAS detector. The integrated luminosity is $140\,\text{fb}^{-1}$.

7.1. Event Yields

The event yields from the fully hadronic preselection region described in Section 6.1 are listed in Table 7.1. Monte Carlo simulation predicts a total of 40756.5 events from background processes at this stage. Of these approximately 64% are predicted to be from the main background contributor of $Z \to \tau\tau$ events and a further 29% from QCD mult-jet events. Single top and $t\bar{t}$ events contribute around 5% to the total background, with W + jets, diboson and $Z \to \ell\ell$ events making smaller contributions.
In the preselection region the signal, here seen as the total number of $H \to \tau\tau$ events, contributes only 1.5% to the total simulation. Its composition is expected to be 28% vector-boson- fusion (VBF), which will later be selected for via BDT, 62% gluon-gluon-fusion (ggF) and 10% Higgs boson radiation (WH,ZH). The agreement between data and simulation is $\frac{N_{\text{data}}}{N_{\text{MC}}} = 0.9992$.

CBA Preselection	
$Z \to \tau\tau$ QCD	25856.8 ± 137.3
$Z \to \tau\tau$ EWK	214.2 ± 3.5
Fakes	11641.6 ± 301.5
W + jets	344.4 ± 21.9
Top	2198.5 ± 21.8
VV	477.6 ± 5.5
$Z \to \ell\ell$ ($\ell = e, \mu$)	23.4 ± 11.3
$\sum$ Bkgd	40756.5 ± 303.0
Data	41347 ± 203.3
ggF $H \to \tau\tau$	389.2 ± 2.9
VBF $H \to \tau\tau$	172.4 ± 0.7
ZH $H \to \tau\tau$	25.5 ± 0.4
WH $H \to \tau\tau$	37.5 ± 0.6
$\sum$ Signal	624.6 ± 3.1
$\sum$ All	41381.1 ± 303.0

Table 7.1.: The event yields in the preselection region. Note that ttH events are included as they are expected to be negligible. Uncertainties are statistical only

7.2. Kinematic Distributions

Figures 7.1 and 7.2 show the distributions of multiple kinematic variables investigated in the fully hadronic preselection region. The $M_{\text{MMC}}(\tau\tau)$ variable was shown previously in Figure 6.1. The signal is overlaid with the background and amplified by a factor of 20 for visibility. Here all $H \to \tau\tau$ events are regarded as signal regardless of the production mechanism. Only statistical uncertainties are included here. They are shown as the grey bands on the MC bins and in the ratio graph in the lower third of each plot.

The distribution of the two τ lepton momenta seen in 7.1(a) and 7.1(b) is well-modelled across the range. Due to its lower mass, decays from a Z boson are shifted to lower p_{T} values than the signal. The same can be said for the multi-jet events since QCD events dominate at lower energies and for top decays, because of the two undetected neutrinos. The cuts of 40 GeV and 30 GeV for the leading and sub-leading τ lepton are well motivated here as they leave the signal mostly untouched.

The p_{T} distributions of the two jets are displayed in 7.1(c) and 7.1(d). The low p_{T} region of both jets is well-modelled, however, simulation predicts a lack of events towards the tail-ends of the p_{T} spectra, and this discrepancy increases the harder the jet is. This could be due to issues with the MC simulation of the L1 jet matching not occurring for high-p_{T}

events, and is potentially worth investigating since high energy jet systems are selected for in the later VBF region, and the jet four-momenta are required for the generation of the *Optimal Observable*. The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ is modelled in Figure 7.1(e) where it can be seen that the cut imposed removes a significant portion of the QCD background. The separation between the two τ leptons $\Delta R(\tau\tau)$ (Figure 7.1(f)) also has a cut centered on the signal. The lower bound of $0.8 < \Delta R(\tau\tau)$ is specific to the fully hadronic decay channel and is to guarantee that the two trigger matched τ leptons with RoIs of $\Delta R(\tau\tau) = 0.4$ do not overlap. This distribution has a slight shift in the simulation towards higher values where $t\bar{t}$ events make up a percentage contribution.

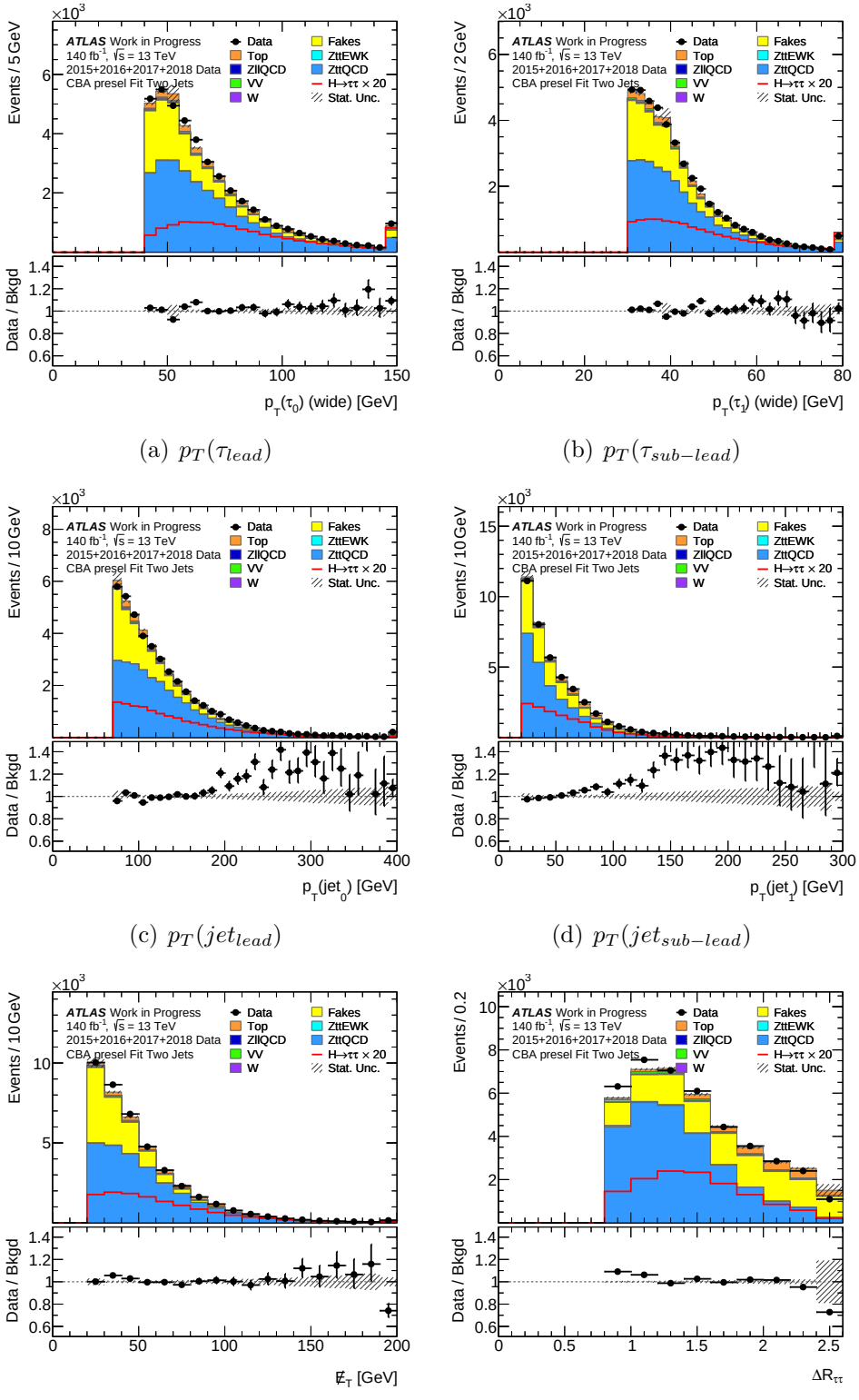
The distribution of the corrected average pileup $< \mu >_{corrected} = \mathcal{L}_{\mathrm{bunch}} \times \sigma_{\mathrm{inel}} / f_r$, where $\mathcal{L}_{\mathrm{bunch}}$ is the instantaneous luminosity per bunch, σ_{inel} is the inelastic cross section and f_r is the LHC revolution frequency, is shown in Figure 7.2(a). The number of reconstructed vertices N_{Vertex} can be seen in 7.2(b). Both exhibit noticeable shape mismodelling, although this is less likely to have an effect on the upcoming analysis as these variables are not utilised. By contrast the simulation of the number of jets and the Higgs boson transverse momentum N_{jets} and p_T^H, the latter of which is calculated from the transverse momentum of the ditau system and the missing transverse energy, in Figures 7.2(c) and 7.2(d) are mostly well-modelled.

An evaluation of the performance *Optimal Observable* at first order at this stage of the analysis is critical, since it is the CP-odd variable used to test for CP violation later. Since the Standard Model background is assumed to be CP invariant, the distribution is expected to have a vanishing mean and be symmetrically distributed. Figure 7.2(e) shows that this is indeed the case and that the *Optimal Observable* is well constructed. The SM signal sample also demonstrates these properties. The distribution of the second order component of the *Optimal Observable* is also plotted in Figure 7.2(f).

7.3. Signal Sample $\tilde{d}$ Reweighting Results

The results of the reweighting of the signal samples as described in Section 6.2 are shown here. Figures 7.3(a) to 7.3(d) show the distribution of the first order *Optimal Observable* reweighted for values of $\tilde{d} = 0$, 0.02, -0.1, 0.5. Here only the VBF $H \to \tau\tau$ events are included in the signal, since only these events are considered to exhibit CP violation. The results show a clear non-zero mean value and an asymmetric distribution of the signal even in the case of $\tilde{d} = 0.02$. The direction of the shift indicates the sign of the parameter and becomes more noticeable for higher values. The reweighted signals have not been normalised to the Standard Model cross section and consequently the signal strength

increases considerately for $\tilde{d} = 0.5$. It should be noted that the analysis method which relies only on the first order term of the *Optimal Observable* is no longer sensitive for values beginning around this range as the second order terms begin to dominate. To enable a comparison with only the 2015 and 2016 dataset at $36.1\,\mathrm{fb}^{-1}$ the same corresponding distributions are displayed in the Appendix.

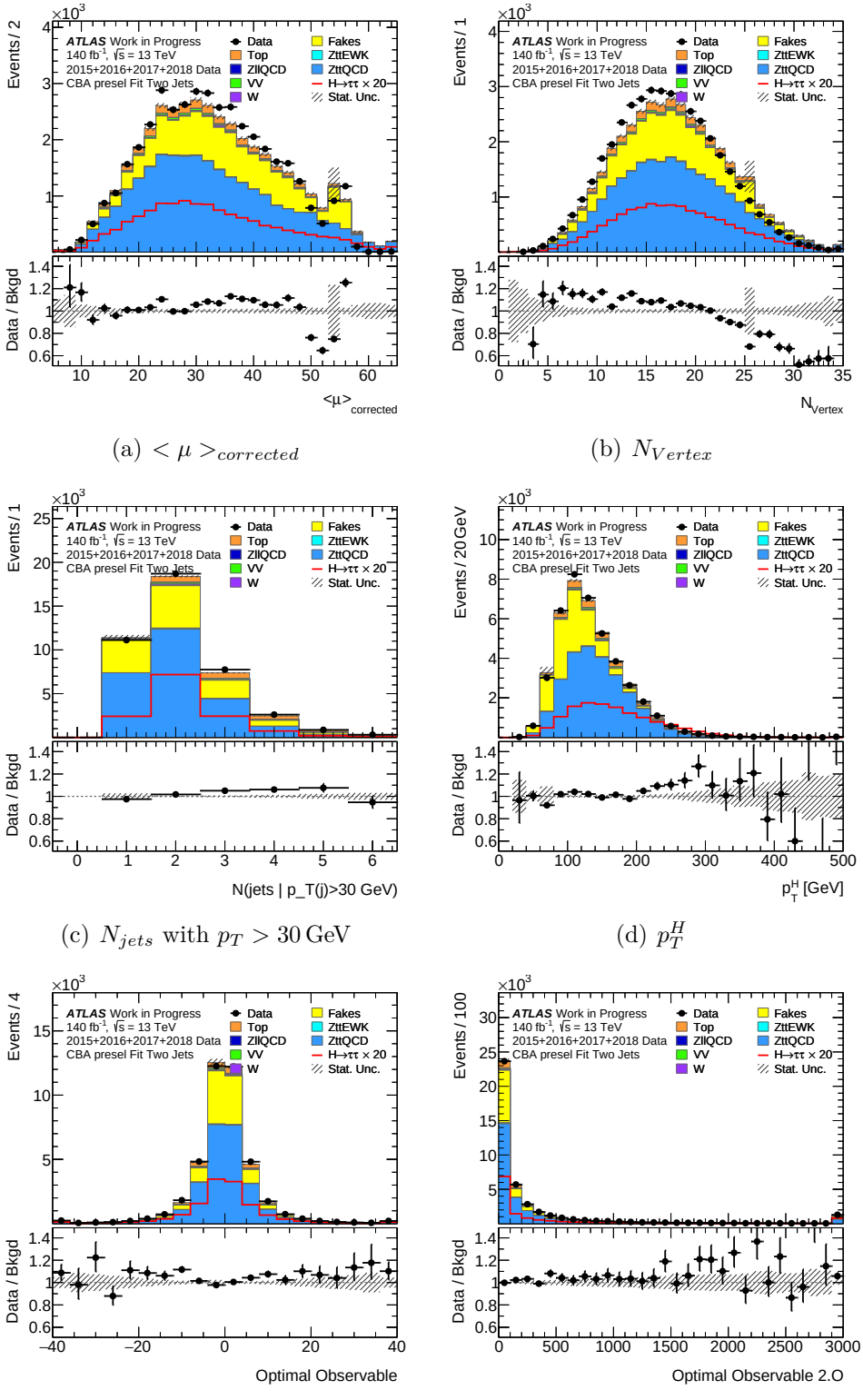


(a) $p_T(\tau_{lead})$ (b) $p_T(\tau_{sub-lead})$

(c) $p_T(jet_{lead})$ (d) $p_T(jet_{sub-lead})$

(e) $E_{\mathrm{T}}^{\mathrm{miss}}$ (f) $\Delta R(\tau\tau)$

Figure 7.1.: Kinematic variables distributions in the preselection region. Only statistical uncertainties are shown here.



(a) $< \mu >_{corrected}$

(b) N_{Vertex}

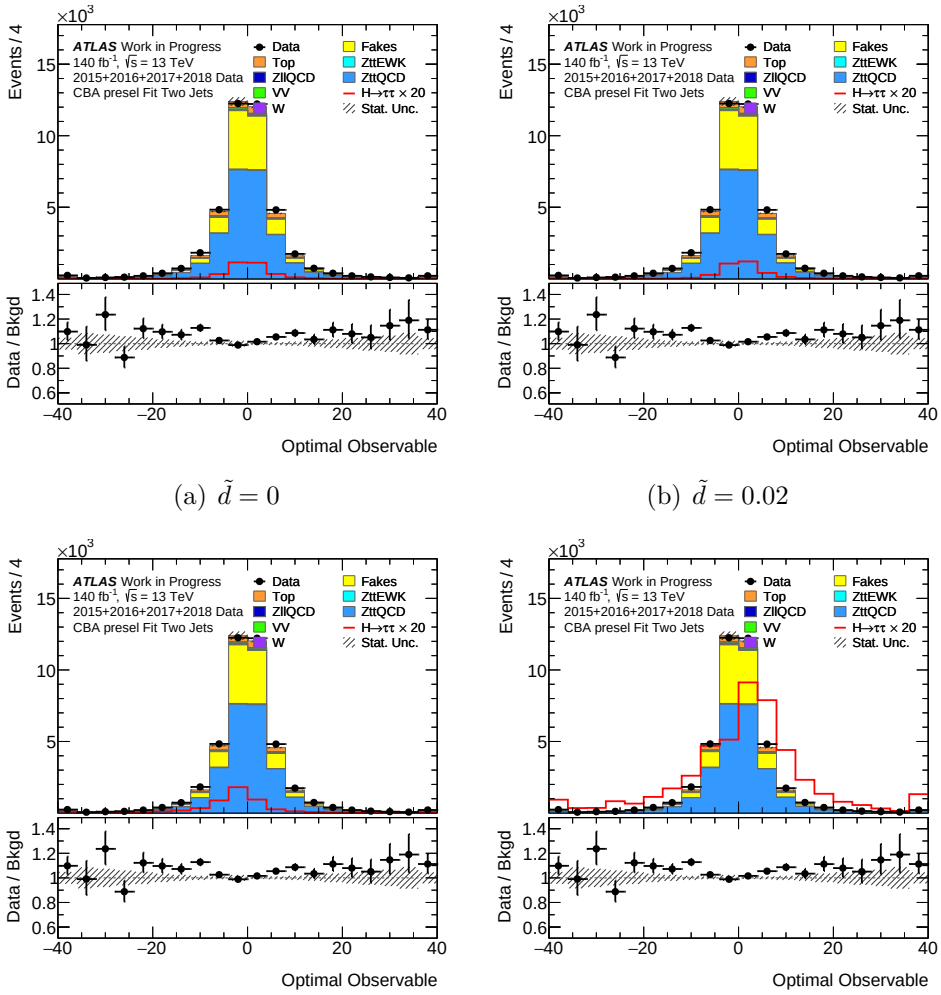
(c) N_{jets} with $p_T > 30\,\mathrm{GeV}$

(d) p_T^H

(e) Optimal Observable

(f) Optimal Observable 2nd Order

Figure 7.2.: Kinematic variables distributions in the preselection region. Only statistical uncertainties are shown here.



(a) $\tilde{d} = 0$ (b) $\tilde{d} = 0.02$

(c) $\tilde{d} = -0.1$ (d) $\tilde{d} = 0.5$

Figure 7.3.: Distributions of the *Optimal Observable* where the signal sample has been reweighted to various $\tilde{d}$ values. he signal here is solely comprised of VBF $H \to \tau\tau$ events.

8. Conclusion

8.1. The Tau Trigger Efficiency Measurements

The first half of this thesis detailed the $Z \to \tau\tau$ trigger efficiency measurements using the Tag-and-Probe method. A significant portion was dedicated to locating the cause of the mismodelling in the 3-prong decays. This was quickly shown to not only be related to the HLT tau25 medium trigger itself, while also being present in measurements from the previous two years.

It should be noted that the ideal performance of the 1-prong decay channel provided a perfect control during the measurements and the investigation as both channels were treated identically. Simulations of 1-prong $\tau_{\text{had-vis}}$ decays predict the same number of events in the signal region to within 0.5% and the trigger efficiencies are well-modelled with respect to tau p_T, η and pile-up. Contrasting this with the approximate 10% excess in simulated events in the 3-prong decay channel shows the extent of the problem.

The 3-prong issue could not be traced to a particular combination of p_T and detector region. The two-dimensional scale factors calculated were approximately constant everywhere, although statistics were limited in high-p_T, high-η regions.

Similarly the data-derived method of estimating the W + jets background did not lead to a resolution, although the possibility of this background being a contributing factor could not be excluded due to the poor performance of the replacement samples. Future samples tailored to the $Z \to \tau\tau$ trigger measurement could provide clarity.

The discovery of the shape mismatch of relevant input variables used to train the tau-ID BDT discriminant is promising but ultimately further investigation is outside the scope of this thesis that was originally intended to focus on the trigger itself. Energy measurements in the electromagnetic calorimeter are implicated in the BDT input variables that exhibit the worst modelling.

The investigation led to some other results. First, the two-dimensional scale factors could be of use to future analysis teams, requiring information on the pseudorapidity binning, while being willing to sacrifice precision on the momentum or vice versa. Second, the method implemented to determine the systematic uncertainty of the transform factor

shows promise, as no alternative method is currently used. This would of course depend on the existence of better samples.
These additions and the enhanced understanding of the 3-prong issue are also of relevance to future tau trigger efficiency measurements and should contribute to improving the performance in the years to come.
The trigger measurements, in particular the 3-prong decay issue encountered demonstrate the need for physics analysis teams to compensate with the scale factors to obtain accurate results where τ leptons are involved.

8.2. $H \to \tau\tau$ CP Analysis

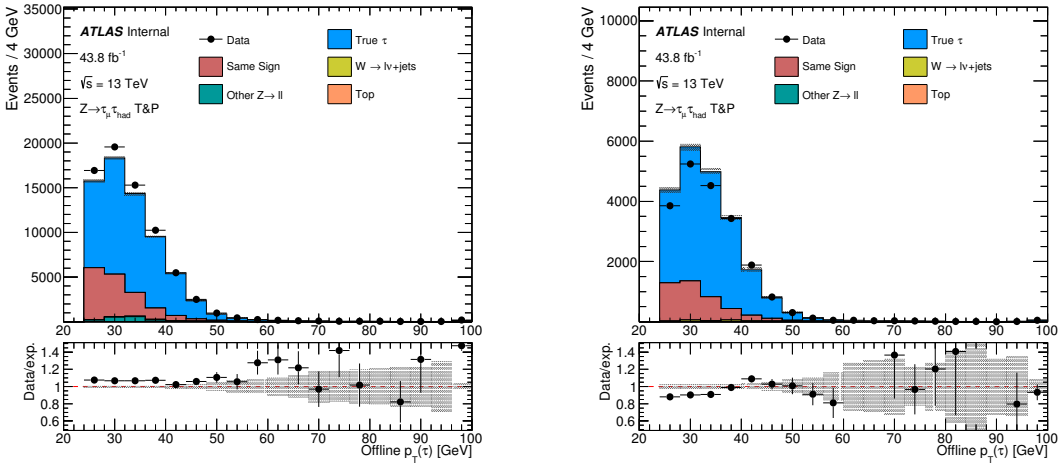
The final measurement of CP invariance in the VBF production mechanism of the Higgs boson will benefit from having access to nearly four times the luminosity of the previous analysis. This should allow for a tighter constraint on the value of $\tilde{d}$ and further narrow the window in which any BSM CP violation can occur.
The evaluation of the event yields in the fully hadronic decay channel at preselection emphasises the challenge of the analysis. At this stage $H \to \tau\tau$ decays make up only approximately 1.5% of the total events and the VBF production mechanism contributes only 0.4%. The variables used in the definition of the VBF region shown in Figure 6.3 only marginally improve these ratios at the cost of excluding some signal events. At preselection the agreement between simulation and data is exceptionally good with the ratio in number of events observed versus expected at 99.92%.
The first order *Optimal Observable* used to probe the VBF vertex is calculated from the four momenta of the Higgs boson and the outgoing jets, and is consequently dependent on the behaviour of these quantities being well understood. This thesis provides evidence that this is indeed the case, with the possible exception of the jet transverse momenta at high values.
This thesis shows that the distribution *Optimal Observable* itself is consistently symmetrical with a vanishing mean for the background processes as well as for the SM signal sample. At the same time it is shown that the shift in the mean is sufficiently sensitive to events generated from matrix elements that include a CP-odd contribution.
The results obtained here in the hadronic channel provide a good foundation for the later optimisation of the VBF region and the BDT used for the signal region. This should allow for the most rigorous test yet of the CP nature of the VBF vertex.

A. Attempt at W + jets Modelling Improvement

The W + jets background method described in Section 4.2.3 is used because the reliability of the MC samples is unknown and a pure MC simulation-based approach currently cannot provide the necessary sample size for accurate modelling of the process. Although W + jets make up less than 10% of the events in the signal region, they could nevertheless play a role in the observed 3-prong discrepancies. Previous results used W + jets samples created using events generated with POWHEG [53] and a parton shower simulated using PYTHIA [54]. Redoing the analysis using samples created with SHERPA [55] provides the opportunity to rule out the background as the cause of the issue as well as the possibility of quantifying the uncertainty introduced by the method itself. This is achieved by varying the transform factor f_W by the ratio of the transfer factors calculated using either generator:

$$f_W = f_W * \left(1 \pm \left|1 - \frac{f_W(S\text{HERPA})}{f_W(PoPy)}\right|\right), \tag{A.1}$$

where $f_W(S\text{HERPA})$ and $f_W(PoPy)$ are the transfer factors calculated using SHERPA and POWHEG-PYTHIA for the W + jets samples respectively. The $\tau_{\text{had-vis}}$ p_T distributions from the SHERPA MC simulation are shown in Figure A.1. Whereas the simulation and data were previously in good agreement in 1-prong decays when using the POWHEG-PYTHIA samples, as can be seen by comparing with Figure 4.6, the switch to SHERPA generated W + jets events leads to a lack of simulated 1-prong events in the distribution while having next to no impact on the MC excess in the 3-prong case. The current SHERPA samples predict only a trace contribution by W bosons in the signal region, compared with the approximate $5 - 10\%$ seen previously. This is indicative of a much lower sample size and is quantified in Table A.1 where the number of W and Z events as simulated by both generators before and after application of the HLT tau25 medium trigger is listed along with the transform factors f_W used. f_W is computed separately for low- and high-p_T regions, with the boundary at $p_T = 35\,\text{GeV}$.



(a) Sherpa, 1-prong

(b) Sherpa, 3-prong

Figure A.1.: The $\tau_{\text{had-vis}}$ p_T background modelling in the 1- and 3-prong decay channel with the W + jets events generated via Sherpa.

The Sherpa-derived transfer factors are shown to typically adopt much lower values than for Powheg-Pythia, leading to the lack of events. At the same time, the highly varying values lead to the prediction of an increase in events after the trigger is applied. These results were traced to large uncertainties in the number of events used in Equation 4.4 to calculate f_W. To further explain the poor performance, the individual Sherpa samples were evaluated and the number of raw events in the signal region in each sample was compared with the yield after application of all weights. Any irregularities in the ratios of these quantities along with large errors would show that the samples are not optimised to provide the statistics for the particular analysis required. This is shown in Table A.2. The samples are sorted based on the invariant mass of the W boson m_W as well as the presence of c- and b-quark filters. From the results it is apparent that low mass events with $0 \leq m_W \leq 70\,\text{GeV}$ dominate in the signal region. This is expected since the W bosons must pass selection criteria designed for tau leptons which are significantly lighter. The provided samples, on the other hand mostly supply events in the $70\,\text{GeV} \leq m_W \leq 280\,\text{GeV}$ range. Similarly, while events with b-vetoes are required due to the preselection criteria, the samples provide more events with b-quarks.
This evaluation shows that the Sherpa samples do not provide the statistics needed in the relevant regions for tau scale factor measurements. They consequently do not improve the 3-prong modelling. Using these particular Sherpa samples to provide an uncertainty on

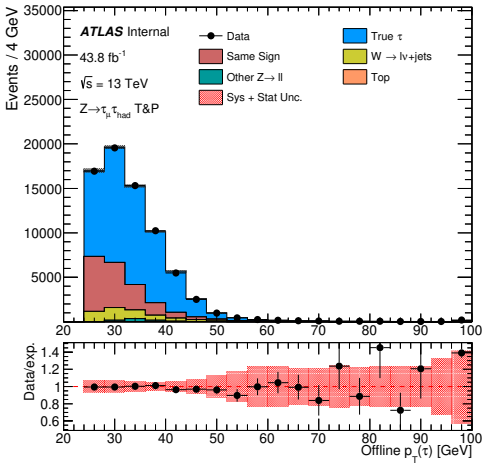
Process/fWs	1-prong		3-prong	
Generator	POWHEG-PYTHIA	SHERPA	POWHEG-PYTHIA	SHERPA
		Before Trigger		
$W \to l\nu$ + jets	5616 ± 77	363 ± 5	1302 ± 62	416 ± 17
$Z \to ll$	894 ± 206	1541 ± 209	-118 ± 73	-110 ± 56
fW lowPt	0.26	0.01	0.44	0.10
fW highPt	0.21	0.03	0.24	0.14
		After Trigger		
$W \to l\nu$ + jets	3715 ± 55	1485 ± 26	514 ± 47	-32 ± 7
$Z \to ll$	-7 ± 149	284 ± 136	-73 ± 57	-38 ± 31
fW lowPt	0.25	0.13	0.72	0.07
fW highPt	0.20	0.04	0.11	-0.07

Table A.1.: Fake tau event yields and transform factors f_W in the signal region, as predicted by the data-driven approach using POWHEG-PYTHIA and SHERPA as generators, before and after the HLT tau25 medium trigger.

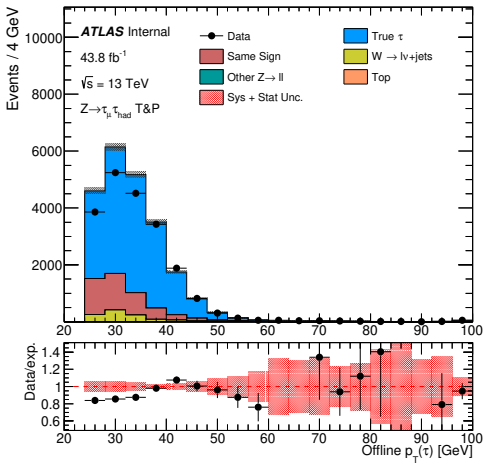
Sample	Signal Region 1-Prong			Signal Region 3-Prong		
SHERPA $W \to \mu\nu$ + jets	Raw	Weighted	Error	Raw	Weighted	Error
m_W/GeV = $[0, 70]$, c veto, b veto	137	1658	1787	24	-10.80	735.8
m_W/GeV = $[0, 70]$, c filter, b veto	177	971.3	309.4	65	478.2	154.5
m_W/GeV = $[0, 70]$, b filter	281	516.6	68.20	69	107.8	26.41
m_W/GeV = $[70, 140]$, c veto, b veto	459	803.2	142.2	96	233.6	64.16
m_W/GeV = $[70, 140]$, c filter, b veto	405	317.4	53.46	173	156.2	29.46
m_W/GeV = $[70, 140]$, b filter	768	86.01	10.88	205	22.31	5.508
m_W/GeV = $[140, 280]$, c veto, b veto	185	166.4	20.58	47	46.74	11.17
m_W/GeV = $[140, 280]$, c filter, b veto	159	106.4	13.62	65	36.54	6.165
m_W/GeV = $[140, 280]$, b filter	460	29.40	2.067	126	8.169	1.141
m_W/GeV = $[280, 500]$, c veto, b veto	51	14.10	3.015	19	6.099	1.597
m_W/GeV = $[280, 500]$, c filter, b veto	50	12.56	2.289	14	2.990	0.9254
m_W/GeV = $[280, 500]$, b filter	33	2.952	0.7170	10	0.7808	0.3219
m_W/GeV = $500, 1000]$	28	2.332	0.6002	10	0.5568	0.214
m_W/GeV > 1000	5	0.03883	0.01834	3	0.03047	0.01831

Table A.2.: Raw and weighted events by SHERPA samples which are sorted by invariant mass range m_W and presence of c- and b-quark filters. Negative yields can result as a consequence of a negative transfer factor as calculated via Equation 4.4.

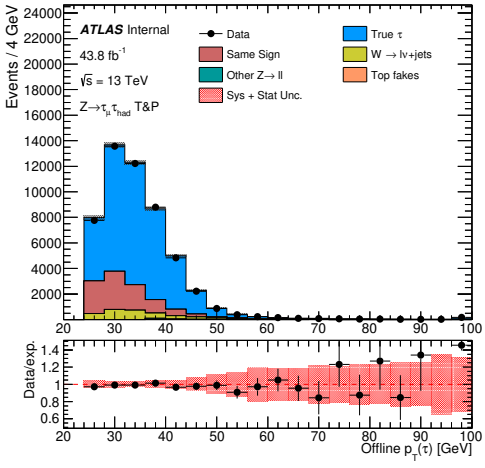
the f_W factors as provided by POWHEG-PYTHIA via Equation A.1 results in unrealistically large errors that are of little use. Here, they are shown in Figure A.2.



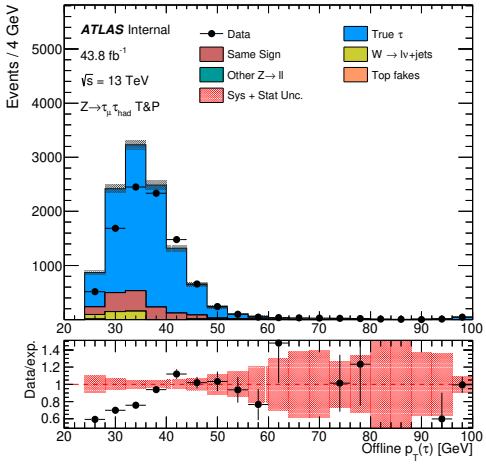
(a) 1-prong, before trigger



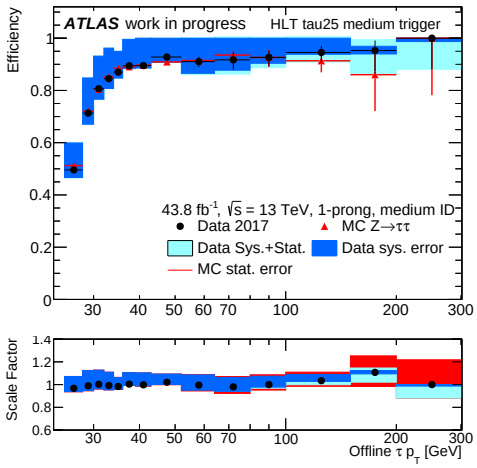
(b) 3-prong, before trigger



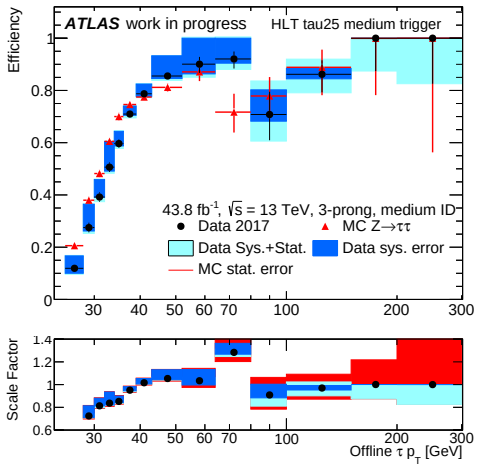
(c) 1-prong, after trigger



(d) 3-prong, after trigger



(e) 1-prong efficiency



(f) 3-prong efficiency

Figure A.2.: $\tau_{\text{had-vis}}$ p_T distributions and HLT tau25 medium trigger efficiencies with systematic uncertainties on f_W, calculated by varying the POWHEG-PYTHIA-derived values by the difference between POWHEG-PYTHIA and SHERPA.

B. Two-Dimensional Scale Factors

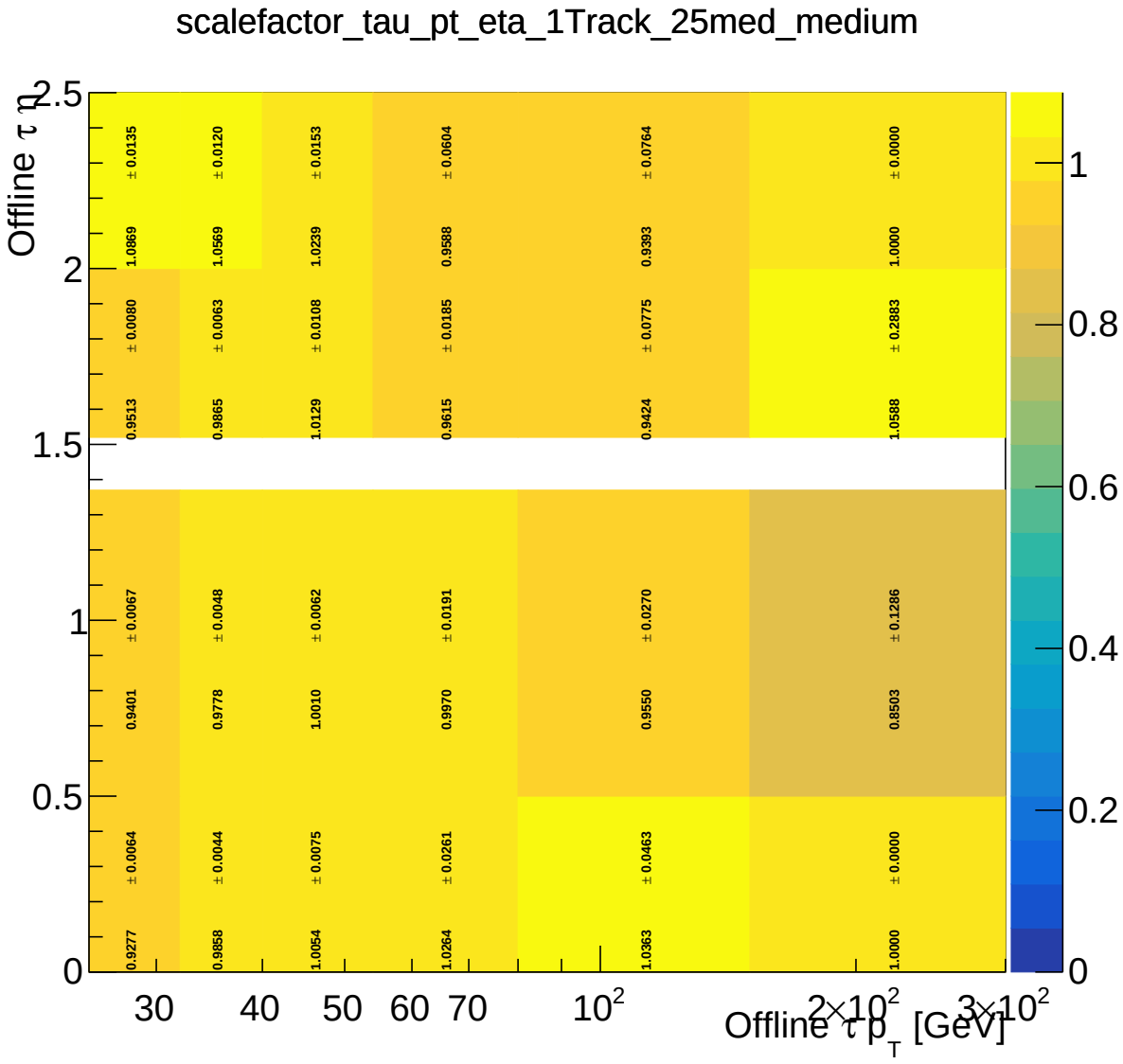


Figure B.1.: 1-prong scale factors binned in p_T and $|\eta|$ of the $\tau_{\text{had-vis}}$ candidate

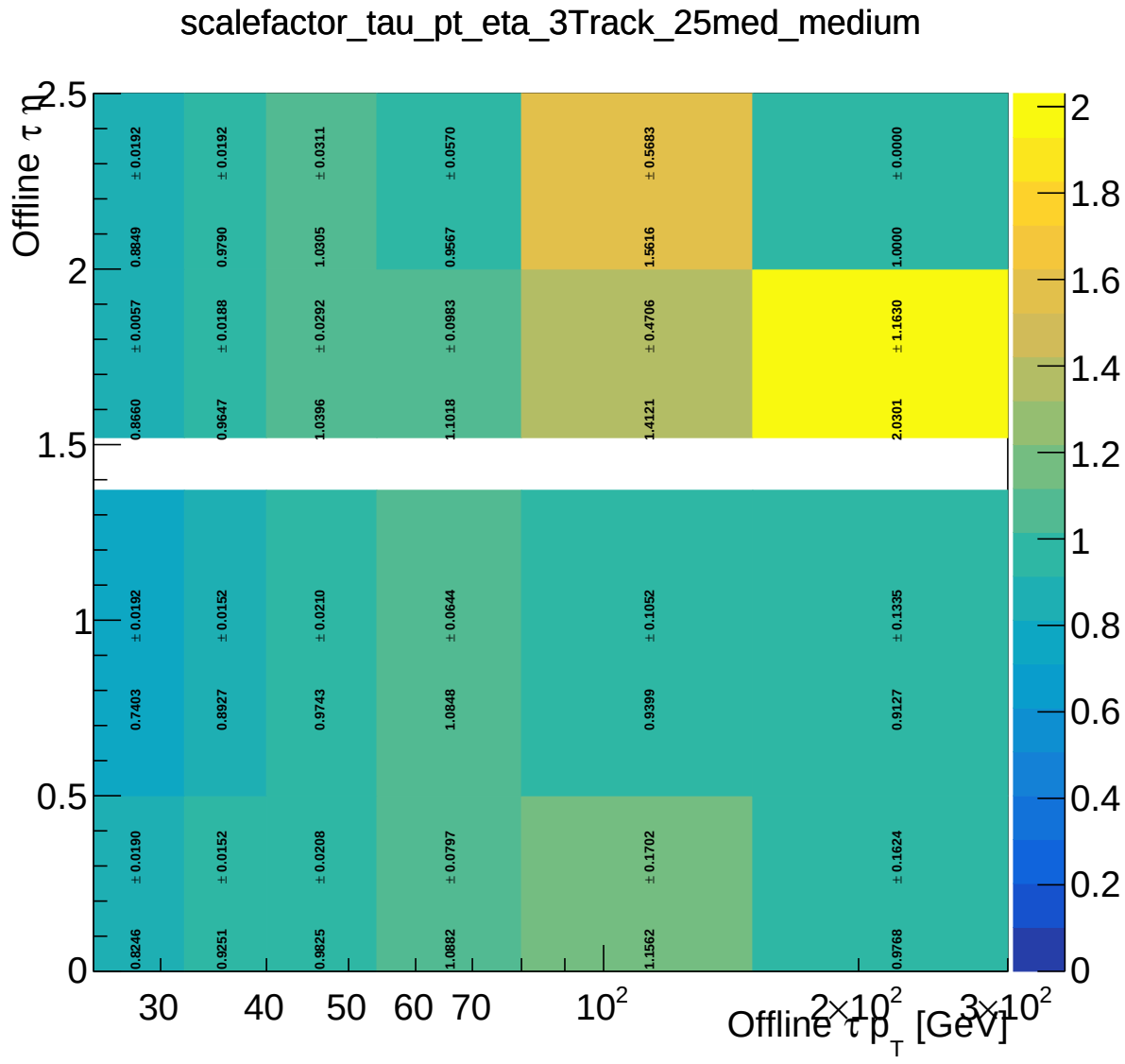
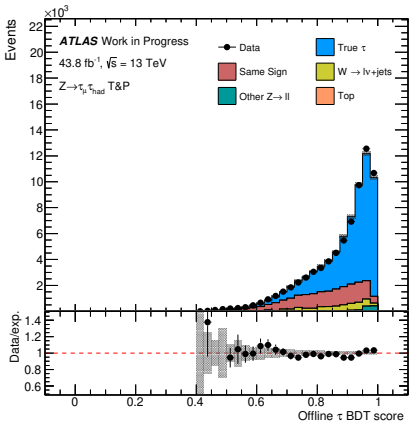
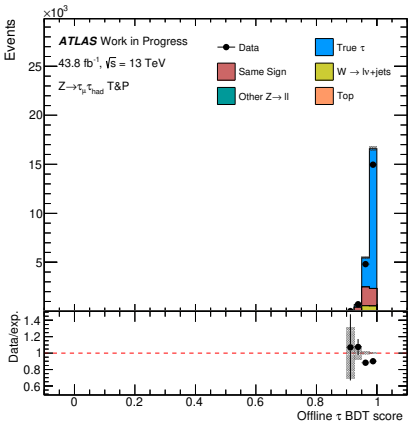


Figure B.2.: 3-prong scale factors binned in p_T and $|\eta|$ of the $\tau_{\text{had-vis}}$ candidate

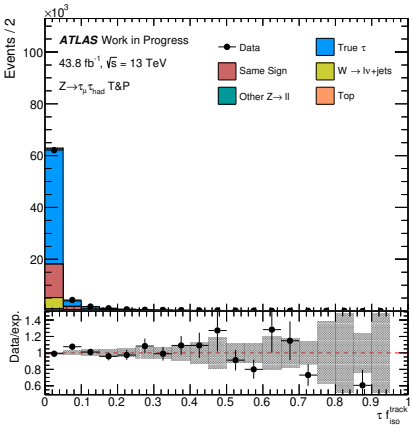
C. Additional HLT Variables



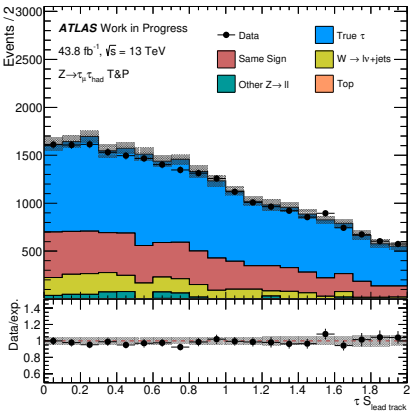
(a) Jet BDT score, 1-prong



(b) Jet BDT score, 3-prong

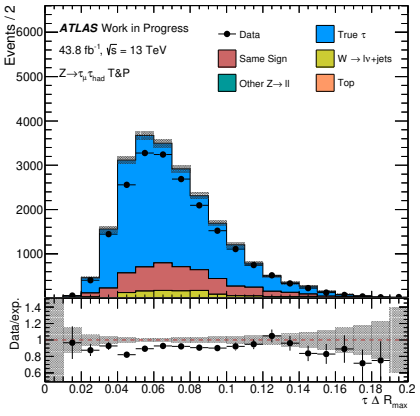


(c) $f_{\text{iso}}^{\text{track}}$, 1-prong

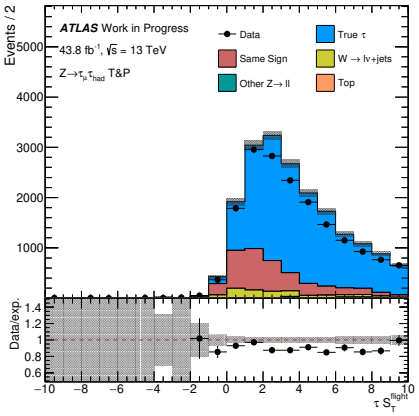


(d) $S_{\text{lead track}}$, 1-prong

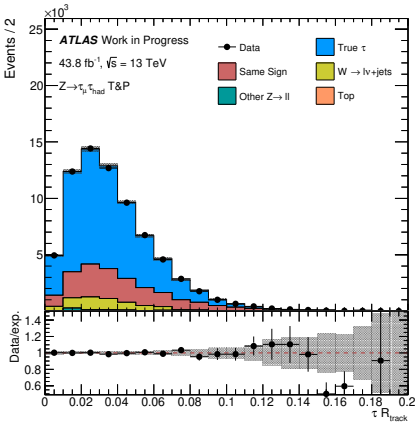
Figure C.1.: Jet BDT score, fraction of tracks p_T in the isolation region $f_{\text{iso}}^{\text{track}}$ (1-prong only) and leading track IP significance $S_{\text{lead track}}$ (1-prong only).



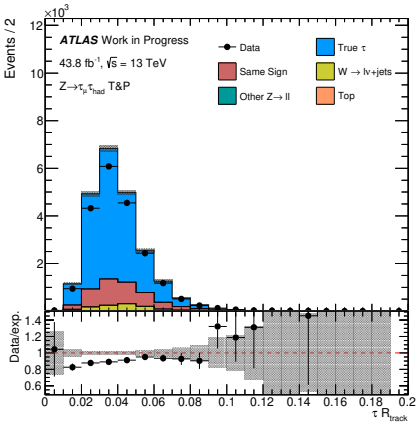
(a) ΔR_{Max}, 3-prong



(b) S_T^{flight}, 3-prong

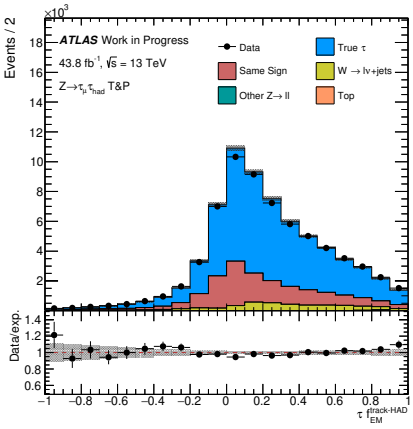


(c) R_{track}, 1-prong

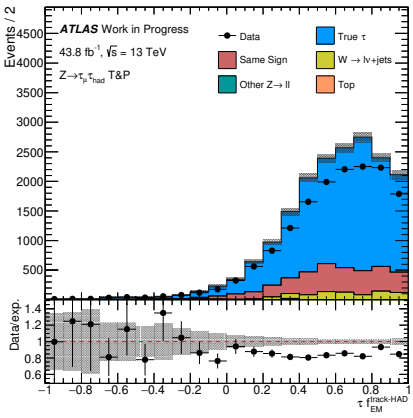


(d) R_{track}, 3-prong

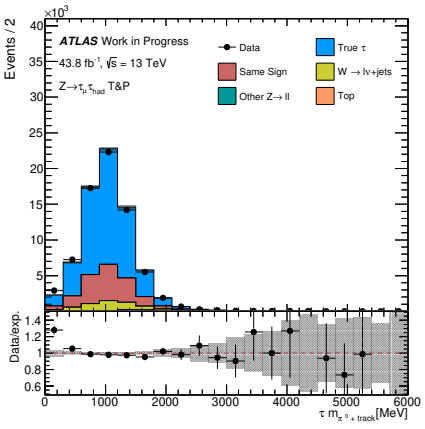
Figure C.2.: Maximum ΔR_{Max} (3-prong only), transverse flight path significance S_T^{flight} (3-prong only) and Track Radius R_{track}.



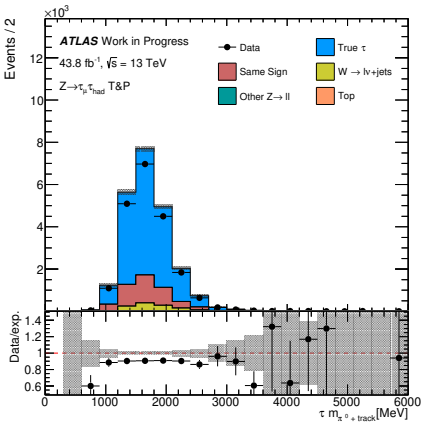
(a) $f_{\mathrm{EM}}^{\text{track-HAD}}$, 1-prong



(b) $f_{\mathrm{EM}}^{\text{track-HAD}}$, 3-prong



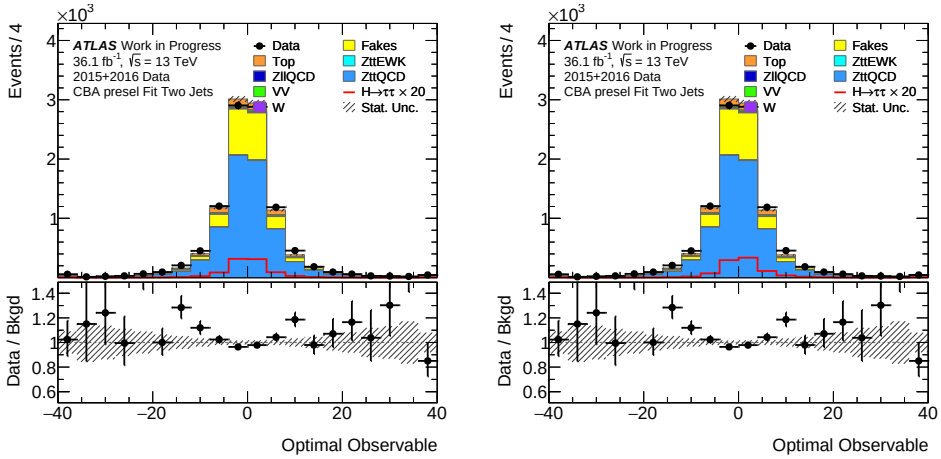
(c) $m_{\mathrm{EM+track}}$, 1-prong



(d) $m_{\mathrm{EM+track}}$, 3-prong

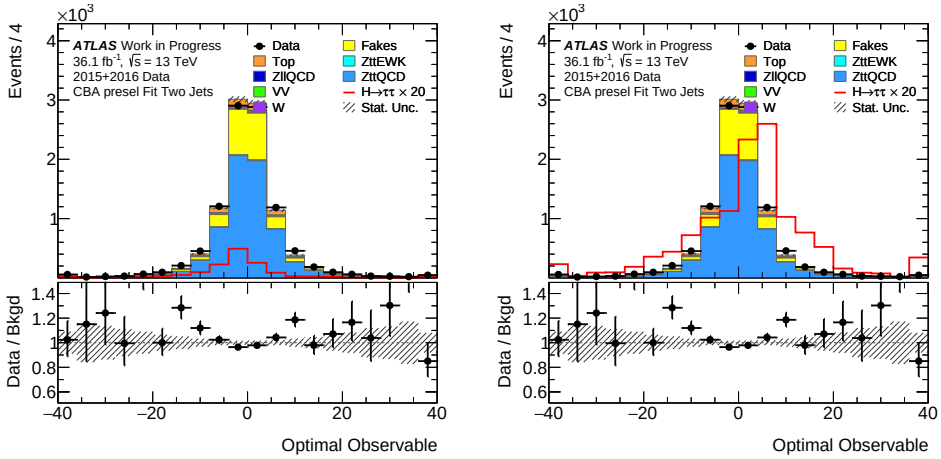
Figure C.3.: Fraction of EM energy from charged pions $f_{\mathrm{EM}}^{\text{track-HAD}}$ and Track-plus-EM-system mass $m_{\mathrm{EM+track}}$.

D. 2015/2016 Optimal Observable Distributions



(a) $\tilde{d} = 0$

(b) $\tilde{d} = 0.02$



(c) $\tilde{d} = -0.1$

(d) $\tilde{d} = 0.5$

Figure D.1.: Distributions of the *Optimal Observable* with reweighted VBF signal samples using only the 2015 and 2016 datasets.

Bibliography

[1] The ATLAS Collaboration, *Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. **D99**, 072001 (2019), 1811.08856

[2] The 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(19)**, 191801 (2018), [Erratum: Phys. Rev. Lett.122,no.8,089901(2019)], 1808.00336

[3] D. Atwood, A. Soni, *Analysis for magnetic moment and electric dipole moment form-factors of the top quark via e+ e- —> t anti-t*, Phys. Rev. **D45**, 2405 (1992)

[4] The ATLAS Collaboration, *Test of CP invariance in vector-boson fusion production of the Higgs boson in the $H \to \tau\tau$ channel in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Technical Report ATLAS-CONF-2019-050, CERN, Geneva (2019), URL https://cds.cern.ch/record/2693960

[5] The 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. **B716**, 1 (2012)

[6] The CMS collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716**, 30 (2012)

[7] M. Thomson, *Modern Particle Physics*, Cambridge university Press (2013)

[8] M. Kobayashi, T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49**, 652 (1973)

[9] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22**, 579 (1961)

[10] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519**, 367 (1968)

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

[12] T. P. D. Group, *Review of Particle Physics*, Phys. Rev. D **98**, 030001 (2018)

[13] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*, Technical Report arXiv:1307.1347. CERN-2013-004, Geneva (2013)

[14] The Atlas Collaboration, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*, Phys. Lett. **B786**, 59 (2018), 1808.08238

[15] S. Borowka, et al., *Higgs boson pair production in gluon fusion at NLO with full top-quark mass dependence*, Phys. Rev. Lett. **117(1)**, 012001 (2016)

[16] Particle Data Group, *Review of Particle Physics**, Phys. Rev. D **86**, 010001 (2012)

[17] The BESIII Collaboration (BESIII), *Precision measurement of the mass of the τ lepton*, Phys. Rev. **D90(1)**, 012001 (2014), 1405.1076

[18] The Atlas Collaboration, *Measurement of the Z to tau tau Cross Section with the ATLAS Detector*, Phys. Rev. **D84**, 112006 (2011), 1108.2016

[19] C. Ferro (ATLAS, CMS), *Top antitop quarks production in tau lepton + b-jets final states*, Nucl. Phys. Proc. Suppl. **253-255**, 184 (2014)

[20] The Cms Collaboration, *Search for neutral MSSM Higgs bosons decaying to a pair of tau leptons in pp collisions*, JHEP **10**, 160 (2014), 1408.3316

[21] H. E. Haber, G. L. Kane, *The search for supersymmetry: Probing physics beyond the standard model*, Phys. Rep. **117**, 75 (1985)

[22] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967)

[23] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105**, 1413 (1957)

[24] J. H. Christenson, J. W. Cronin, V. L. Fitch, R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13**, 138 (1964)

[25] V. A. Kostelecky, *The Status of CPT*, in *Physics beyond the standard model. Proceedings, 5th International WEIN Symposium, Santa Fe, USA, June 14-19, 1998*, pages 588–600 (1998), hep-ph/9810365

[26] J. E. Kim, *A Review on axions and the strong CP problem*, AIP Conf. Proc. **1200(1)**, 83 (2010), 0909.3908

[27] H. Nunokawa, S. J. Parke, J. W. F. Valle, *CP Violation and Neutrino Oscillations*, Prog. Part. Nucl. Phys. **60**, 338 (2008), 0710.0554

[28] The ATLAS Collaboration, *Test of CP Invariance in vector-boson fusion production of the Higgs boson using the Optimal Observable method in the ditau decay channel with the ATLAS detector*, Eur. Phys. J. **C76(12)**, 658 (2016), 1602.04516

[29] G. Aad, B. Abbott, J. Abdallah, O. Abdinov, R. Aben, M. Abolins, O. S. AbouZeid, H. Abramowicz, H. Abreu, et al., *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, The European Physical Journal C **75(10)** (2015), URL http://dx.doi.org/10.1140/epjc/s10052-015-3685-1

[30] L. Evans, P. Bryant, *LHC Machine*, JINST **3**, S08001 (2008)

[31] The ATLAS collaboration, *ATLAS: Detector and physics performance technical design report. Volume 1* (1999)

[32] E. Stanecka (ATLAS), *The ATLAS Inner Detector operation, data quality and tracking performance*, in *Proceedings, 32nd International Symposium on Physics in Collision (PIC 2012): Strbske Pleso, Slovakia, September 12-15, 2012*, pages 383–388 (2013), 1303.3630

[33] D. E. Boumediene (ATLAS), *ATLAS Calorimeter: Run 2 performance and Phase-II upgrades*, PoS **EPS-HEP2017**, 485 (2017)

[34] The ATLAS Collaboration, *ATLAS muon spectrometer: Technical design report* (1997)

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

[36] The ATLAS Collaboration (ATLAS), *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70**, 823 (2010), 1005.4568

[37] E. Bothmann, et al., *Event Generation with Sherpa 2.2*, SciPost Phys. **7**, 034 (2019), 1905.09127

[38] T. Sjostrand, S. Mrenna, P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05**, 026 (2006), hep-ph/0603175

[39] C. Oleari, *The POWHEG-BOX*, Nucl. Phys. Proc. Suppl. **205-206**, 36 (2010)

[40] A. e. a. Schaelicke, *Event generator for particle production in high-energy collisions*, Prog. Part. Nucl. Phys. **53**, 329 (2004), hep-ph/0311270

[41] The GEANT4 Collaboration, *GEANT4: A Simulation toolkit*, Nucl. Instrum. Methods Phys. Res., A **506**, 250 (2002)

[42] The ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. **C74(10)**, 3071 (2014), 1407.5063

[43] The ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$=13 TeV*, Eur. Phys. J. **C76(5)**, 292 (2016), 1603.05598

[44] The ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. **C77**, 490 (2017), 1603.02934

[45] M. Cacciari, G. P. Salam, G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04**, 063 (2008)

[46] The ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11(04)**, P04008 (2016), 1512.01094

[47] The ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. **C78(11)**, 903 (2018), 1802.08168

[48] The ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. **C77(5)**, 317 (2017), 1611.09661

[49] P. Giovannini, *Local hadron calibration with ATLAS*, Journal of Physics: Conference Series **293** (2010)

[50] The ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV*, Eur. Phys. J. **C75(7)**, 303 (2015), 1412.7086

[51] The ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, Technical Report ATL-PHYS-PUB-2015-045, CERN, Geneva (2015), URL https://cds.cern.ch/record/2064383

[52] The ATLAS collaboration, *The ATLAS Tau Trigger in Run 2*

[53] S. Alioli, P. Nason, C. Oleari, E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06**, 043 (2010), 1002.2581

[54] T. Sjostrand, et al., *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178**, 852 (2008)

[55] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, J. Winter, *Event generation with SHERPA 1.1*, JHEP **02**, 007 (2009), 0811.4622

[56] The Atlas Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, Technical Report ATL-PHYS-PUB-2019-033, CERN, Geneva (2019)

[57] *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of sqrts = 7 TeV*, Technical Report ATLAS-CONF-2010-038, CERN, Geneva (2010)

[58] A. Elagin, P. Murat, A. Pranko, A. Safonov, *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, Nucl. Instrum. Meth. **A654**, 481 (2011), 1012.4686

[59] M. Davier, L. Duflot, F. Le Diberder, A. Rouge, *The Optimal method for the measurement of tau polarization*, Phys. Lett. **B306**, 411 (1993)

[60] M. Diehl, O. Nachtmann, *Optimal observables for the measurement of three gauge boson couplings in e+ e- —> W+ W-*, Z. Phys. **C62**, 397 (1994)

[61] A. Denner, S. Dittmaier, S. Kallweit, A. Mueck, *HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders*, Comput. Phys. Commun. **195**, 161 (2015), 1412.5390

[62] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, C. P. Yuan, *New parton distributions for collider physics*, Phys. Rev. **D82**, 074024 (2010), 1007.2241

Acknowledgements

I would like to thank the members of the II. Institute of Physics for their help during my time there. In particular, I would like to acknowledge the following people:
First, Prof. Dr. Stan Lai for giving me the opportunity to be a part of his research team, and for his guidance and supervision over the course of the year.
Similarly, my thanks goes to Dr. Michel Janus, Dr. Jason Veatch and Dr. Tobias Bisanz for their advice, and for their comments and proof-reading.
Serhat Ördek introduced me to all the tools needed for my research and provided helpful advice and assistance throughout. I would like to thank him for this, and for his knowledge and patience in helping me get to the bottom of the many issues and bugs I encountered throughout.
Additional thanks goes to Timo Dreyer and Lino Gerlach who were always there to answer any questions I had.

Erklärung nach §17(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 30. Januar 2020

(Kieran Amos)