

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/95242/>

Copyright and reuse:

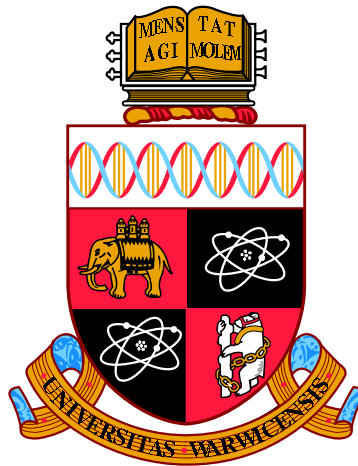
This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk



**Test of CP-invariance in Vector Boson Fusion
production of the Higgs boson using the Optimal
Observable method in the di-tau decay channel
with the ATLAS Detector**

by

Martin Spangenberg

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Physics

November 2016

THE UNIVERSITY OF
WARWICK

Contents

Acknowledgments	iv
Declarations	v
Abstract	vii
Chapter 1 Theory	1
1.1 The Standard Model of Particle Physics	1
1.1.1 Particles in the Standard Model	1
1.1.2 Quantum Field Theory	6
1.1.3 Electromagnetism	8
1.1.4 Yang-Mills Theory and the Strong Force	10
1.1.5 Electroweak Unification	14
1.1.6 Electroweak Symmetry Breaking	16
1.1.7 Masses of the Fermions	19
1.1.8 Unsolved Problems in the Standard Model	20
1.2 Higgs Phenomenology	24
1.2.1 Higgs Boson Production at the LHC	24
1.2.2 Higgs Boson Decays	27
1.3 CP Violation in VBF Higgs Production	29
1.3.1 Theoretical Model of CP Violation	30
1.3.2 Observables Sensitive to CP Violation	32
Chapter 2 The ATLAS Experiment	38
2.1 The Large Hadron Collider	38
2.2 Proton-Proton Collider Physics	41
2.2.1 Luminosity	41
2.2.2 Parton Distribution Functions	42
2.2.3 Partonic Cross Sections	44

2.3	The ATLAS Detector	46
2.3.1	Definition of Coordinates and Variables	47
2.3.2	The Inner Detector	48
2.3.3	Calorimeters	51
2.3.4	The Muon Spectrometer	54
2.3.5	The ATLAS trigger system	56
2.4	Data Taking with ATLAS in 2012	57
2.5	Muon Calorimeter Isolation	59
Chapter 3 Event Reconstruction		63
3.1	Tracking and Vertexing	64
3.2	Topological Clusters	65
3.3	Jets	67
3.3.1	Jet Reconstruction and Calibration	68
3.3.2	b -tagging	71
3.4	Electrons	72
3.4.1	Electron Trigger in $\tau\ell\tau_{\text{had}}$	73
3.4.2	Electron Reconstruction and Identification	74
3.5	Muons	77
3.5.1	Muon Trigger in $\tau\ell\tau_{\text{had}}$	77
3.5.2	Muon Reconstruction and Identification	78
3.6	Hadronically Decaying Taus	80
3.6.1	Reconstruction and Identification of Hadronic Tau Decays . .	82
3.6.2	Tau Energy Calibration	88
3.7	Missing Transverse Energy	90
Chapter 4 $H \rightarrow \tau\tau$ Analysis Method		93
4.1	Experimental Signatures of Signal and Background	94
4.1.1	Higgs Boson Processes	94
4.1.2	Background Processes	95
4.1.3	Reconstruction of the Ditau Mass	98
4.2	Monte Carlo Simulated Samples	102
4.2.1	Signal Samples	102
4.2.2	Background Samples	103
4.3	Signal Re-weighting	105
4.3.1	Validation of the Re-weighting Procedure	105
4.4	Data Samples and Triggers	108
4.4.1	Object Definitions	109

4.5	Data-driven background estimation	111
4.5.1	$Z \rightarrow \tau\tau$ Embedding Procedure	111
4.5.2	Fake Tau Background Estimation	113
4.6	Event Selection	116
4.6.1	Background Suppression	116
4.6.2	Signal and Control Regions	122
Chapter 5 Fit Model and Systematic Uncertainties		139
5.1	Description of the Log-likelihood Fit	140
5.1.1	Pruning and Smoothing of Nuisance Parameters	141
5.2	Experimental Uncertainties	143
5.3	Data-driven Background Modelling Uncertainties	148
5.4	Theoretical Uncertainties	149
Chapter 6 Results		155
6.1	Preliminary Studies	155
6.1.1	Gauge Curve from Neyman Construction	156
6.1.2	Construction of χ^2 measure	156
6.2	Expected sensitivity	158
6.3	Observed results	161
6.4	Validation of the Maximum-Likelihood Fit	167
6.5	Conclusion	174
Appendix A Electron Energy Scale		175
Appendix B List of Nuisance Parameters		177
Appendix C Effects of anomalous couplings in $H \rightarrow W^+W^-$ decays		179

Acknowledgments

I owe many people my thanks for their time, advice and friendship during the four years I have studied at the University of Warwick. First of all I want to thank the whole Warwick ATLAS group, both past and present members. Thank you Sinead Farrington and Paul Harrison for taking on the roles as my supervisors and guiding me whenever I needed it. Thank you Michel Janus for our close collaboration through large parts of my analysis work, for the stimulating discussions and for livening up the office. Also thank you to Elisabetta Pianori and Tim Martin for imparting your knowledge. Chris and Adomas, I hope I didn't frustrate you too much with my ramblings about digital currency! Joe, thanks for organising some enjoyable dinners in town. Everyone else not explicitly mentioned, know that I have enjoyed your company.

I would also like to thank the other members of the Higgs CP analysis group including Elias Coniavitis, Markus Schumacher, Vincent Croft and Christian Schillo for their professional and dedicated effort on the analysis, and especially Christian for our close collaboration between decay channels. Additionally I need to thank everyone who has worked on the $H \rightarrow \tau\tau$ couplings analysis as my work would not have been possible without the foundation they built.

Living in the UK (and France during my LTA) was quite a change for me after 25 years in Denmark, but the rock that is my parents held me steady. I could not wish for more supportive or caring people than you. To all my friends at Radio Halsnaes that I have mostly worked with remotely during these years, thank you for keeping me sane and maintaining my ties to North Zealand.

Declarations

This thesis is being submitted to the University of Warwick as part of my application for the degree of Doctor of Philosophy. It has been written by myself and has not been used in support of any previous degree application.

My contribution to the online running of ATLAS and the required qualification task to become an author of the collaboration was the development and validation of a muon trigger algorithm with an improved calorimeter isolation estimate, taking into account the energy density in the full detector as well as the muon self-energy. This work is described in section 2.5.

The VBF $H \rightarrow \tau\tau$ CP analysis [1] has been a collaboration among several institutes directly involving more than ten people. In addition, it builds on top of the efforts of the many members of ATLAS who contributed to the $H \rightarrow \tau\tau$ evidence paper [2]. As the main analyser measuring the CP properties of the Higgs boson in the $H \rightarrow \tau_l\tau_{\text{had}}$ decay channel I

- developed the analysis framework in C++ and python using ROOT [3]
- defined and optimised signal and control regions (section 4.6.2)
- implemented the treatment of systematics and estimated additional systematic uncertainties arising from the use of a new observable and a different signal region compared to the VBF category in the $H \rightarrow \tau\tau$ evidence paper (chapter 5)
- produced preliminary sensitivity estimates using a simplified analysis using the mean value of the optimal observable (section 6.1)

- independently constructed the fit model and wrote the fitting code using HistFactory and RooStats (section 5.1)
- performed the standalone $\tau_\ell\tau_{\text{had}}$ fit as well as the combination with the input from $\tau_\ell\tau_\ell$ prepared by Christian Schillo [4] (sections 6.2 and 6.3)
- completed fit stability studies and investigated issues arising from low sample statistics and systematics (section 6.4)

Additionally, I produced the plots included in the CP measurement paper [1] using input from each channel as well as signal re-weighting validation work performed by others.

Abstract

A test of CP invariance in Higgs boson production via Vector Boson Fusion using the optimal observable method is presented. The analysis exploits the decay mode of the Higgs boson into a pair of tau leptons in the decay channels $H \rightarrow \tau_l \tau_l$ and $H \rightarrow \tau_l \tau_{\text{had}}$ and is based on 20.3 fb^{-1} of proton–proton collision data at $\sqrt{s} = 8 \text{ TeV}$ collected by the ATLAS experiment at the LHC. CP-violating interactions between the Higgs boson and electroweak gauge bosons are described in an effective field theory framework, where the strength of CP violation is governed by a single parameter \tilde{d} . The mean values and distributions of CP-odd observables agree with the expectation in the Standard Model and show no sign of CP violation. The CP-mixing parameter \tilde{d} is constrained to the interval $[-0.11, 0.05]$ at the 68% confidence level, consistent with the Standard Model expectation of $\tilde{d} = 0$.

Introduction

Elementary particle physics studies the smallest known building blocks of matter and the interactions between them in the search for a comprehensive theory of their structure and dynamics. All particles known today interact through four distinct forces: electromagnetism, the weak and strong nuclear forces and gravity. The Standard Model (SM) represents the modern description of particle physics including the former three interactions while unifying the electromagnetic and weak forces as established by Glashow, Salam and Weinberg. With the exception of gravity, which has not yet successfully been described by a quantum theory, the SM provides a framework that accurately describes almost all observed processes. It has led to an impressive list of predictions that subsequently resulted in the experimental discovery of new particles, the most recent being the 2012 discovery of the Higgs boson that enters the formulation of electroweak unification.

Despite the great success of the SM over many decades it is still known to be an incomplete theory. One very palpable piece of evidence is the fact that we live in a universe dominated by matter, even though the early universe should have contained equal amounts of matter and antimatter. It would be impossible for such an asymmetry to evolve over time if the laws of physics were symmetric under the exchange of matter and antimatter, as they would then annihilate to produce a universe filled with light but no matter. The SM does actually allow a small amount of matter-antimatter asymmetry of one part in 10^{-17} , but this is much lower than the observed asymmetry of 10^{-10} . One of the requirements for the existence of matter-antimatter asymmetry is CP violation, which means that the physical process in question is not invariant under the simultaneous inversion of charge and spatial coordinates. Searches for as yet unknown CP-violating contributions to SM processes could therefore hint new physics beyond the SM.

This thesis analyses the particular case of couplings of the Higgs boson to weak gauge bosons by investigating its Vector Boson Fusion production mode with subsequent decays to tau leptons. An effective field theory is used to model the CP-

violating contributions since the exact theory that might produce such couplings is not known. This allows a measurement of the strength of the contribution from such couplings, parametrised through the variable \tilde{d} . Discrimination between CP-even and CP-mixing states is achieved using the Optimal Observable method, which employs a multidimensional discriminating variable incorporating the full matrix element of the process. The analysed data is collected by the ATLAS experiment at the Large Hadron Collider (LHC) in 2012, comprising 20.3 fb^{-1} of proton-proton collisions. Chapter 1 presents the particle contents and theoretical framework of the SM, the dominant Higgs boson production modes at a proton-proton collider as well as its decay channels, followed by a brief introduction to CP violation and the discriminating variables used in the analysis. Chapter 2 describes the experimental facilities used in the data collection, namely the CERN accelerator complex including the LHC and the ATLAS detector. Chapter 3 discusses the reconstruction of collision events in the detector in terms of the objects from which they are built and how these are calibrated. After defining the relevant signal and background processes, chapter 4 then uses the kinematics of these objects to define the regions of interest and methods of background suppression. Chapter 5 introduces the fit procedure used to extract limits on the CP-odd coupling strength \tilde{d} followed by the treatment of systematic uncertainties and how they enter the fit. Finally, chapter 6 presents the results of the analysis and provides a confidence interval on \tilde{d} .

Chapter 1

Theory

1.1 The Standard Model of Particle Physics

The Standard Model (SM) is the theoretical framework that describes three of the four currently known fundamental interactions of the universe and its constituents. It has been experimentally verified to high precision during many generations of experiments. The first parts were developed in the 1960's with the formulation of the Electroweak Theory, a unified description of the electromagnetic and weak interactions, resulting in the 1979 Nobel Prize in Physics awarded to Sheldon Glashow, Abdus Salam and Steven Weinberg. The formulation of the electroweak interaction invokes another important concept called spontaneous symmetry breaking. Through this mechanism the Higgs field can be introduced to explain why some of the known particles have mass. This piece of the theoretical puzzle was already in place in 1964, but the experimental confirmation was to remain elusive until 2012, when the ATLAS and CMS experiments jointly announced the discovery of the Higgs boson, leading to the 2013 Nobel Prize being awarded to Peter Higgs and François Englert. The parton model and subsequent experimental evidence showing that hadrons are not fundamental particles but instead consist of quarks have led to our current understanding of the strong interaction through the theory of Quantum Chromodynamics. Another physics Nobel Prize was awarded to David Politzer, Frank Wilczek and David Gross in 2004 for their work on one aspect of the strong interaction called asymptotic freedom.

1.1.1 Particles in the Standard Model

Table 1.1 serves as a complete list of the known particles in the SM. It comprises gauge bosons, leptons and quarks in addition to the Higgs boson. Their properties include mass and spin as well as the charges they carry. The particles can be categorised in many ways according to their properties, giving rise to terms such as ‘antiparticles’ and ‘generations’ that simplify their presentation. The particles

Particle	Symbol	Spin	EM charge [e]	Weak charge (Isospin, I^3)	Colour charge	Mass [MeV]
electron	e	1/2	-1	-1/2	0	0.511
electron neutrino	ν_e	1/2	0	+1/2	0	$< 2 \cdot 10^{-6}$
muon	μ	1/2	-1	-1/2	0	105.7
muon neutrino	ν_μ	1/2	0	+1/2	0	< 0.19
tau	τ	1/2	-1	-1/2	0	1777
tau neutrino	ν_τ	1/2	0	+1/2	0	< 18.2
up	u	1/2	+2/3	+1/2	R, G, B	~ 2.3
down	d	1/2	-1/3	-1/2	R, G, B	~ 4.8
charm	c	1/2	+2/3	+1/2	R, G, B	$\sim 1.275 \cdot 10^3$
strange	s	1/2	-1/3	-1/2	R, G, B	~ 95
top	t	1/2	+2/3	+1/2	R, G, B	$\sim 173 \cdot 10^3$
bottom	b	1/2	-1/3	-1/2	R, G, B	$\sim 4.2 \cdot 10^3$
photon	γ	1	0	0	0	0
Z boson	Z^0	1	0	0	0	$91.188 \cdot 10^3$
W boson	W^\pm	1	± 1	± 1	0	$80.4 \cdot 10^3$
gluon	g	1	0	0	Colour pair	0
Higgs boson	H^0	0	0	1/2	0	$125.1 \cdot 10^3$

Table 1.1: *The Standard Model particles and their properties. Data has been taken from the Particle Data Group [5]. The particles are divided into classes by double lines with leptons at the top, followed by quarks and finally the gauge bosons and the Higgs boson at the bottom. Both quarks and leptons are divided by single lines into their first, second and third generations.*

can be divided into two groups, fermions and bosons, which will be described in the following. According to the spin-statistics theorem they are distinguished by their spin, which is defined as intrinsic quantised angular momentum with both integer and half-integer values allowed. Interactions between the particles are summarised in figure 1.1.

Fermions

All particles with half-integer spin are called fermions and respect the Pauli exclusion principle [7]. Consequently their energies are described by Fermi-Dirac statistics [8, 9]. These properties give rise to the structure seen in the periodic table of the elements, and fermions are suitably called matter particles although only a few of them form the building blocks of atoms. A distinction is made between elementary fermions subject to the strong force, called quarks, and those that do not feel the strong force, called leptons. The known elementary fermions all have a spin of $\frac{1}{2}$, but the term applies to any particle, composite or elementary, whose total spin is a half-integer number.

Leptons are able to interact through the weak force and a subset of them

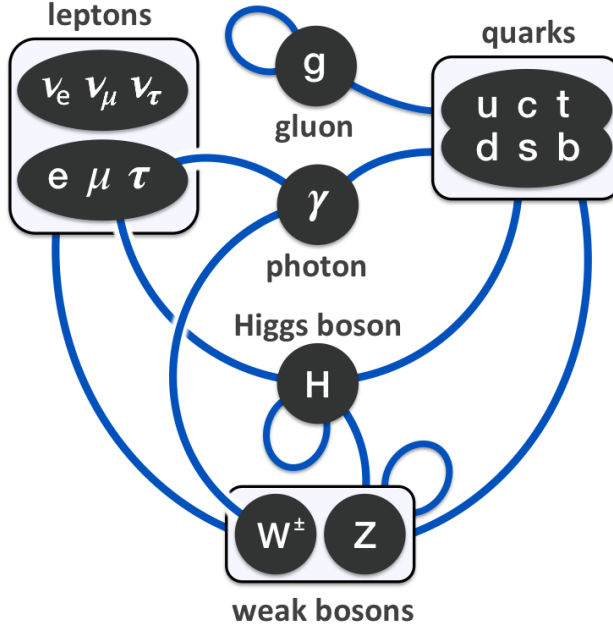


Figure 1.1: All possible interactions among particles in the Standard Model [6]. Particles are represented by ovals, while interactions are given by lines drawn between the particles participating in the interaction. Particles of multiple generations are joined into one oval. All particles in a box are subject to an interaction if the line ends at the edge of the box. Lines starting and ending on the same oval/box represent self-interactions.

through the electromagnetic force. This is determined by the charges they carry. The electron is an example of a lepton interacting through both of these forces. It has unit electric charge $-e$, which is the lowest nonzero EM charge of any known observable particle subject to electromagnetism. Closely related to the electron are the muon (μ) and the tau (τ) lepton. These can be thought of as heavier copies of the electron. Each of these has an associated neutrino denoted ν_e , ν_μ and ν_τ respectively. These interact only through the weak force and can be inferred in collider experiments only through the observation of missing momentum (see section 2.3.1). Neutrino physics has developed rapidly in recent years with the confirmation of neutrino oscillations in 1998 [10] indicating that neutrinos have nonzero and different masses. At present only squared mass differences and upper mass limits are known.

Quarks on the other hand can interact through all three Standard Model forces. All known quarks carry a fractional electric charge of $-\frac{1}{3}e$ or $+\frac{2}{3}e$, but all observed bound states have their quark contents arranged in such a way that the sum of charges is a multiple of the unit charge e . Only bound states of pairs

(mesons) or triplets (baryons) have been observed, not individual quarks. This concept is called colour confinement (see section 1.1.4). It is theoretically possible to have bound states containing more than three quarks, and a discovery of bound states containing five quarks was recently made by the LHCb experiment in the $\Lambda_b^0 \rightarrow J/\psi K^- p$ channel [11]. These states have been dubbed pentaquarks. The two lightest quarks are the up (u) and down (d) quarks that make up hadrons such as protons and neutrons. Combined with the electron they form all the chemical elements, but only a small fraction of the total mass of an atom. Protons and neutrons have masses far exceeding the sum of the masses of their three constituent quarks (2 u , 1 d for protons and 1 u , 2 d for neutrons). For example, the proton mass of 938.272046 MeV [5] mostly comes from the binding energy in the gluon field between the quarks.

All currently known fermions can be arranged in three generations, where each generation contains two quarks and two leptons. Evidence from experiments at LEP (Large Electron-Positron Collider) suggests that there cannot be more than three generations of particles in the universe. The number of light neutrino generations has been determined to be 2.9840 ± 0.0082 [12] in agreement with the known three generations. However, this result does not rule out a neutrino massive enough to avoid being detected at the LEP centre-of-mass energy of 209 GeV. Searches for fourth-generation particles have recently been performed with negative results, see e.g. [13], [14]. It is expected that the cross sections of several Higgs production and decay channels are increased in the presence of a fourth generation of fermions [15]. However, Higgs production and decay rates from LHC measurements so far do not show any evidence for this.

Bosons

The gauge bosons give rise to all three fundamental forces described by the SM, i.e. the electromagnetic, weak and strong forces. The force carriers are all spin-1 bosons and each of the fundamental forces is mediated by at least one such gauge boson.

Electromagnetism is the only force to have just one associated gauge boson. The photon, γ , is however quite versatile in its observable phenomena ranging from electricity and magnetism to the spectrum of electromagnetic radiation. Although the photon is responsible for binding matter together at the scale of atoms and molecules it also has infinite range due to its having no mass. The prerequisite for interaction with the photon is electric charge, carried by the W bosons as well as all fermions except the neutrinos.

The weak interaction is often called the weak nuclear force as a reference to

its key role in nuclear β -decays. It is mediated by three gauge bosons called the W^+ , W^- and Z^0 particles. They are themselves massive and thus have an extremely short range of approximately 10^{-17} m. Weak isospin determines whether a matter particle can interact through the weak force. Only particles with nonzero weak isospin undergo weak interactions, which is the case only for left-handed fermions. As the W bosons have electric charge they couple to the photon, and the weak force carriers also have self-couplings. The SM allows triple gauge boson couplings between a W^+W^- pair and either a photon or a Z boson, while the allowed quartic couplings are between four W bosons ($W^+W^-W^+W^-$), a pair of W bosons and either two Z bosons or two photons ($W^+W^-Z^0Z^0$ or $W^+W^-\gamma\gamma$) as well as two W bosons coupling to a Z boson and a photon ($W^+W^-Z^0\gamma$).

The strong interaction is responsible for binding together the quarks forming mesons and hadrons including the nucleons of atoms. The associated gauge bosons are the gluons, g , of which eight unique states exist by virtue of the underlying $SU(3)$ symmetry. The charge analogous to the electric charge of electromagnetism is called colour charge. Colour can take three values called red, green and blue with each gluon having a colour-anticolour pair or a superposition of these pairs associated with it. In contrast the quarks are only charged with one colour which is altered during the exchange of gluons. Only particles with colour charge interact via the strong force making it exclusive to quarks and the gluons themselves. Gluons are massless, but due to the nature of the strong force its effective range is approximately equal to the radius of a nucleon. The strength of the interaction decreases at smaller distances, or equivalently higher energies, such that quarks can be considered free particles at sufficiently high energies. This concept is called asymptotic freedom [16, 17]. At larger distances the interaction strength does not decrease unlike electromagnetism, and extrapolating this behaviour the gluon field reaches a point where it has enough energy to produce new particles. All observed mesons and hadrons are colour-neutral.

The weaker binding force that holds together the nucleons to form nuclei also has its origin in the strong force but in a residual form. The particles mediating this residual strong force are virtual mesons, and the force is felt by nucleons at distances of up to 1 – 3 fm.

Gravity is thought to be mediated by spin-2 bosons, but a consistent framework has not yet been developed to describe gravity in a renormalisable way in the language of Quantum Field Theory. Renormalisation [18] is a set of methods for treating infinities arising in calculated quantities from e.g. loop diagrams. However, effective theories exist [19] where the coupling constants at each order of the

expansion have to be experimentally determined.

The Higgs Boson

The description of the particles in the Standard Model given in the above sections is still incomplete since there is no way to introduce masses for the particles while preserving the gauge invariance of the electroweak force. This problem is remedied by introducing another field which has been hypothesised since the 1960's. The Higgs field has a nonzero vacuum expectation value which leads to spontaneous breaking of electroweak symmetry. By coupling to the Higgs field the particles can acquire mass. This topic is further discussed in section 1.1.6.

This formulation of spontaneous symmetry breaking leads to the existence of a quantum associated with the Higgs field, the scalar Higgs boson. Several generations of accelerator experiments have searched for the Higgs boson until the 4th of July 2012, when the ATLAS and CMS experiments at the LHC jointly announced that a particle with properties consistent with the Standard Model Higgs boson had been discovered (see [20] and [21]). The Higgs mass quoted in table 1.1 is the latest combined measurement from ATLAS and CMS [22].

1.1.2 Quantum Field Theory

The Standard Model of particle physics is formulated within the theoretical framework known as Quantum Field Theory (QFT). This section will give a brief introduction to the concepts and principles behind the current theories of nature as they are written in the language of QFT. It is however not an attempt to give a fully satisfactory mathematical description, which could fill many theses in itself.

Dynamics in a Field Theory

Taking a starting point in classical field theory, all the dynamical information of a system can be contained within a quantity called the action. Given a system with Lagrangian L , the action S of the system is given by

$$S = \int L dt = \int \mathcal{L}(\phi, \partial_\mu \phi) d^4x, \quad (1.1)$$

where \mathcal{L} is called the Lagrangian density, but is commonly simply referred to as the Lagrangian, which will also be the convention in the following. The Lagrangian is a function of the fields and their derivatives, which are again functions of both time and spatial coordinates. Requiring that the action is stationary, that is $\delta S = 0$, will lead to the equations of motion for that particular system. This is called the

principle of least action and leads to the Euler-Lagrange equations of motion for a classical system [23, p. 16]:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \quad (1.2)$$

In Quantum Field Theory the generalisation of this canonical formulation is the *path integral* originally developed by Richard Feynman. According to the path integral formulation the transition amplitude between two states $|\phi\rangle$ and $|\phi'\rangle$ can be written as

$$\langle \phi' | \phi \rangle = \int \mathcal{D}\phi e^{iS[\phi]}, \quad (1.3)$$

where $\mathcal{D}\phi$ denotes the sum over all possible ‘paths’ or configurations, and S is again the action of the system. The path integral generalises the action principle of classical mechanics and does away with the notion of a single, unique path followed by the system. Instead the transition amplitude is a functional integral over a potentially infinite number of paths. Through the use of perturbation theory we can then recover the Feynman rules and diagrams for the desired process. An interesting property of the path integral is that in many cases the contributions from the various paths cancel in such a way that at lowest order only the classical analog remains. Higher order perturbations then give the quantum fluctuations/corrections. In quantum calculations this is referred to as leading order (LO), next-to-leading order (NLO) etc.

Symmetries

One of the most fundamental concepts in modern physics is that of symmetries. Here a symmetry should be understood in the sense that if a system is symmetric under a given transformation of variables, then the observable physical properties of the system will be identical before and after the transformation. Specifically in the case of a classical field theory the equations of motion will be invariant. The utility of studying the symmetries of a theory comes from Noether’s theorem [23, p. 17] which links a given symmetry to a conserved current. In this manner it is possible to discover conserved quantities that might otherwise not be immediately obvious.

Standard Model Symmetries

In the Standard Model one finds a plethora of symmetries with various characteristics. One such group of symmetries is the Poincaré group in Minkowski spacetime.

The Lagrangian of a system must be invariant under these transformations describing coordinate translations, rotations and so-called boosts transforming the system from one relativistic frame of reference to another. The corresponding conserved quantities include linear momentum, angular momentum and energy. The Poincaré symmetries are examples of global symmetries that are independent of the coordinates. They are also continuous since the possible values of the variables are not discretised in any way.

The Standard Model also contains a discrete symmetry constructed from the concepts of time reversal ($t \rightarrow -t$), charge conjugation (charge changes sign, $c \rightarrow -c$, or equivalently particle \rightarrow antiparticle) and parity inversion (all spatial coordinates are flipped, $x_i \rightarrow -x_i$). None of these transformations are symmetries by themselves, and experimental evidence has shown that the weak interaction is not CP-invariant [24]. CP, short for Charge Parity, is a symmetry stating that the laws of physics should be invariant under the simultaneous inversion of charge and spatial coordinates. Thus only the joint operation CPT including a time reversal is considered a true symmetry in all parts of the Standard Model.

The kind of symmetries furthest removed from classical intuition are local gauge symmetries that are dependent on the position in space-time. Such symmetries give rise to all the gauge bosons responsible for the fundamental forces. The full set of gauge symmetries in the Standard Model can be formulated in group theory and can be written as $SU(3) \times SU(2) \times U(1)$, with $SU(3)$ describing the symmetry associated with the strong force and the gauge group $SU(2) \times U(1)$ being the foundation of the electroweak interaction.

1.1.3 Electromagnetism

Electromagnetism was the first force to be described in the language of Quantum Field Theory, and its formulation is a natural place to start. The goal is to extend the global $U(1)$ symmetry to also be valid locally by introducing a new term in the Dirac Lagrangian for a fermion in free space

$$\begin{aligned}\mathcal{L} &= \bar{\psi}(i\cancel{\partial} - m)\psi \\ &= i\bar{\psi}\cancel{\partial}\psi - m\bar{\psi}\psi,\end{aligned}\tag{1.4}$$

where $\psi(x)$ ($\bar{\psi}(x)$) is the wave function of a fermion (anti-fermion) and $\cancel{\partial}$ is a contraction of the sum $\gamma^\mu \partial_\mu$ with μ running over both time and space. γ^μ are the

gamma matrices defined through the anticommutation relations

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu} \mathbb{1}_4, \quad (1.5)$$

where $\eta^{\mu\nu}$ is the Minkowski metric and $\mathbb{1}_4$ is the 4×4 unit matrix. In the Dirac representation they can be written in terms of the Pauli spin matrices σ^i as

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1}_2 \\ \mathbb{1}_2 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}. \quad (1.6)$$

Defining q as the coupling constant and $\lambda(x)$ as a scalar function, it can easily be verified that the local (coordinate-dependent) $U(1)$ transformation corresponding to a simple phase rotation of the wave function

$$\psi \rightarrow e^{-iq\lambda}\psi, \quad \bar{\psi} \rightarrow \bar{\psi}e^{iq\lambda} \quad (1.7)$$

leaves the mass term $m\bar{\psi}\psi$ invariant, while the derivative term $i\bar{\psi}\not{\partial}\psi$ changes such that

$$\mathcal{L} \rightarrow \mathcal{L} + q\bar{\psi}\gamma^\mu\psi\partial_\mu\lambda. \quad (1.8)$$

This can be counteracted by introducing an additional term in the Lagrangian containing a new field $A_\mu(x)$ that transforms as $A_\mu \rightarrow A_\mu + \partial_\mu\lambda$ resulting in the modified Lagrangian

$$\mathcal{L} = i\bar{\psi}\not{\partial}\psi - m\bar{\psi}\psi - q\bar{\psi}\gamma^\mu\psi A_\mu. \quad (1.9)$$

The field is usually introduced into the Lagrangian by defining a covariant derivative given as $\mathcal{D}_\mu = \partial_\mu + iqA_\mu$. The further addition of a propagation term $F_{\mu\nu}F^{\mu\nu}$ with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, which is itself invariant under the condition that the vector field is massless, then gives the final Lagrangian of Quantum Electro Dynamics:

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\not{\mathcal{D}} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (1.10)$$

The newly introduced field A_μ is identified as the photon, and the propagation term $F_{\mu\nu}F^{\mu\nu}$ is equivalent to the Maxwell equations in free space. To conclude, it is possible to derive the QED Lagrangian from the Dirac Lagrangian when adding the simple demand that it be invariant under $U(1)$ gauge transformations. The same method of identifying symmetries is used as a tool in the formulation of both the strong and electroweak theories.

1.1.4 Yang-Mills Theory and the Strong Force

The above procedure of modifying a global symmetry to also hold locally can be extended to apply to cases with more than one field. This results in new transformations based on matrices instead of scalars.

Local Symmetry with Two Fermion Fields

As an example we will follow in the footsteps of Yang and Mills [25] and study the case of two spin- $\frac{1}{2}$ fields ψ_1 and ψ_2 . In this case the full Lagrangian in free space will simply be the sum of two Dirac Lagrangians, one for each field. A more economical way of writing this is in vector form with

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \quad \bar{\psi} = \begin{pmatrix} \bar{\psi}_1 & \bar{\psi}_2 \end{pmatrix}, \quad (1.11)$$

giving the Lagrangian

$$\mathcal{L} = \bar{\psi}(i\not{\partial} - M)\psi, \quad (1.12)$$

with M being a 2×2 diagonal matrix containing the masses of the two fields. In the case where $m_1 = m_2 = m$ it can be interpreted as a real number. We now have a Lagrangian identical in form to equation 1.4 but with ψ being a vector. The global invariance of the system now generalises to a transformation involving a unitary 2×2 matrix U that can be written as $\psi \rightarrow U\psi = e^{i\theta} e^{i\sigma_a \phi_a} \psi$, where we have exploited the unitarity of the matrix ($U^\dagger U = 1$) to rewrite it in exponential form. σ_a is a three-vector containing the Pauli matrices while ϕ_a is a vector containing three real numbers. Disregarding the phase factor of $e^{i\theta}$ that has already been explored in the $U(1)$ transformation leading to electromagnetism, we instead study the $SU(2)$ (unitary with determinant 1) transformation

$$\psi \rightarrow e^{i\sigma_a \phi_a} \psi = e^{-iq\sigma_a \lambda_a} \psi, \quad (1.13)$$

where in the equality we have extracted a coupling constant $-q$ from ϕ_a . This is a global symmetry of the Lagrangian in 1.12, and the Lagrangian must now be changed to make the symmetry hold locally. Making λ_a depend on the coordinates x , the symmetry can be upheld by introducing a new covariant derivative

$$\mathcal{D}_\mu = \partial_\mu + iq\sigma_a A_\mu^a \quad (1.14)$$

that includes three new gauge fields via A_μ^a . In the limit of small $|\lambda_a|$, the transfor-

mation rule for A_μ^a that keeps the Lagrangian invariant can be shown to be

$$A_\mu^a \rightarrow A_\mu^a + \partial_\mu \lambda^a + 2q\epsilon^{abc}\lambda^b A_\mu^c, \quad (1.15)$$

where ϵ^{abc} is the Levi-Civita symbol. As in the $U(1)$ case, the last step is to add a propagation term for the three new vector fields. In addition to the requirement that the fields be massless, the expression needs a new cross-product term in order to uphold the local invariance of the Lagrangian:

$$F_a^{\mu\nu} = \partial^\mu A_a^\nu - \partial^\nu A_a^\mu - 2q(\epsilon_{abc}A_b^\mu A_c^\nu). \quad (1.16)$$

The final Yang-Mills Lagrangian can then be written as

$$\begin{aligned} \mathcal{L}_{\text{YM}} &= i\bar{\psi}\not{D}\psi - m\bar{\psi}\psi - (q\bar{\psi}\gamma^\mu\sigma^a\psi)A_\mu^a - \frac{1}{4}F^{a,\mu\nu}F_{\mu\nu}^a \\ &= \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}F^{a,\mu\nu}F_{\mu\nu}^a. \end{aligned} \quad (1.17)$$

We end up with a Lagrangian in free space describing two spin- $\frac{1}{2}$ fields with identical masses coupling to three massless vector fields. This particular model is not known to describe any real-world interactions, but the derivation is useful as a template when developing related models. It will also play a role as part of the Electroweak Theory that will be discussed in section 1.1.5.

Quantum Chromodynamics

Through deep inelastic scattering experiments starting in the late 1960's it was deduced that the proton is not itself an elementary particle but is composed of so-called partons, today known to be spin- $\frac{1}{2}$ quarks and spin-1 gluons. Around the same time, the field of hadron spectroscopy was also compelled to introduce the concept of colour charge under the quark model in order to explain the properties of the many new observed particles. These efforts culminated in the formulation of Quantum Chromodynamics (QCD), the theory underlying the strong interaction. The derivation of QCD is analogous to the Yang-Mills procedure above. However, now the assumption is that each quark flavour has three colour states called red, green and blue. Attributing the same mass m to all three colour states again leads

to a Dirac Lagrangian of the form given in equation 1.4, but with

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}. \quad (1.18)$$

This leads to a $U(3)$ symmetry such that $\psi \rightarrow U\psi$ leaves the Lagrangian invariant, where U is any unitary 3×3 matrix. Writing the transformation in exponential form while ignoring the phase factor $e^{i\theta}$ we are left with the $SU(3)$ transformation

$$\psi \rightarrow e^{-iq\phi_a T_a} \psi, \quad (1.19)$$

where ϕ_a is a vector of eight real numbers and T_a is a vector containing the Gell-Mann matrices λ_a [26]. The elements are also identified as the generators of $SU(3)$ and can be written as

$$\begin{aligned} T_a &= \frac{1}{2} \lambda_a, \\ [T_a, T_b] &= if_{abc} T_c, \end{aligned} \quad (1.20)$$

containing the $SU(3)$ structure constants f_{abc} . The non-commutating behaviour of the generators means that $SU(3)$ is a non-abelian group, which leads to an extra term in the field strength tensor $G_{\mu\nu}^a$ compared to $F_{\mu\nu}$ in electromagnetism:

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - qf^{abc} G_\mu^b G_\nu^c. \quad (1.21)$$

defining a vector G_μ^a containing eight new gauge fields that can be identified as gluons. Local gauge invariance with $\phi_a = \phi_a(x)$ can be upheld analogously to the $SU(2)$ case by introducing the covariant derivative

$$\mathcal{D}_\mu = \partial_\mu + iq \frac{\lambda_a}{2} G_\mu^a, \quad (1.22)$$

resulting in the final Lagrangian

$$\mathcal{L}_{\text{QCD}} = i\bar{\psi}\not{D}\psi - m\bar{\psi}\psi - q(\bar{\psi}\gamma^\mu T_a \psi)G_\mu^a - \frac{1}{4}G_{\mu\nu}^a G^{a,\mu\nu}, \quad (1.23)$$

which is symmetric under any local $SU(3)$ gauge transformation.

By studying the QCD Lagrangian and comparing it to the QED counterpart in equation 1.10 it is clear that the non-abelian nature of $SU(3)$ leads to more complicated field dynamics through cubic and quartic terms in G that describe

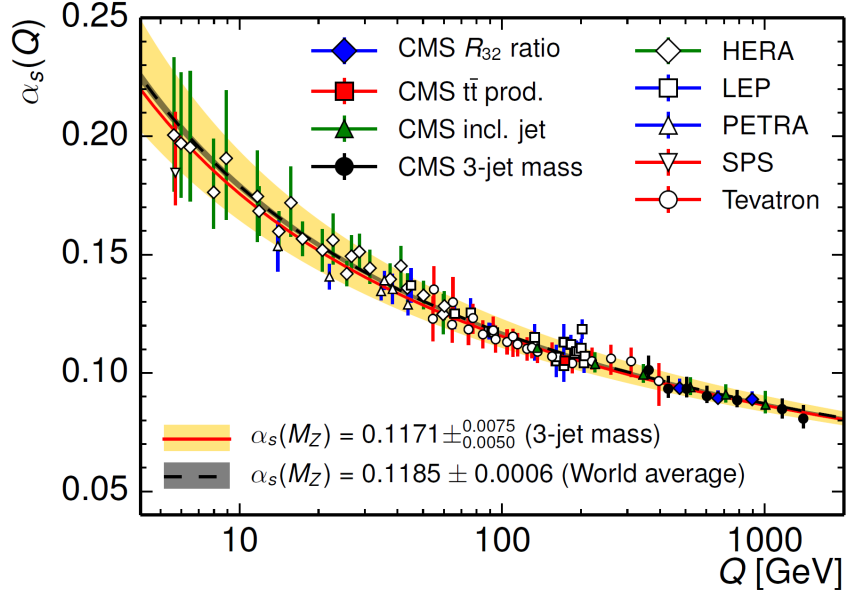


Figure 1.2: Experimental measurements of α_s at various energy scales Q plotted against the QCD scaling prediction [27].

gluon self-interactions (diagrams with three or four external legs respectively). To understand the physical consequences of this, we can investigate the relation between the coupling constant g and the energy scale of the interaction Q^2 , also called the renormalisation group equation:

$$\beta(g) = \frac{\partial g}{\partial \log(Q^2)} = Q^2 \frac{\partial \alpha_s(Q^2)}{\partial Q^2}. \quad (1.24)$$

The β function describes how the coupling strength evolves as the energy scale changes and can be determined using perturbation theory. In terms of the fine structure constant of QED, $\alpha = e^2/4\pi$, and the strong coupling, $\alpha_s = g^2/4\pi$, the β functions of the two theories to first order can be written as

$$\begin{aligned} \beta(\alpha) &= \frac{2\alpha^2}{3\pi}, \\ \beta(\alpha_s) &= -\left(11 - \frac{2n_f}{3}\right) \frac{\alpha_s^2}{2\pi}, \end{aligned} \quad (1.25)$$

where n_f is the number of quark flavours that can partake in the interaction at the energy scale Q^2 . $\beta(\alpha)$ shows that the α of electromagnetism decreases when the energy scale is lowered, describing the effect of charge shielding by polarisation of the vacuum. The opposite is true in QCD, where α_s instead *increases* as the energy

scale is lowered. This has important consequences for the applicability of perturbation theory at low energies as higher orders in α_s no longer converge, making QCD a non-perturbative theory at low energies (large distances). Figure 1.2 shows experimental measurements of α_s inside a wide range of energies. The strongly coupled behaviour leads to *colour confinement*, which means that coloured particles are never experimentally observed as free particles and will only be found as colour neutral bound states. Any partons or coloured hadron remnants produced in scattering experiments will hadronise through the creation of new coloured particles from the vacuum. In the high-energy limit (small distances) the coupling of QCD instead becomes weak, leading to *asymptotic freedom*, which enables the approximate treatment of partons as free particles in high-energy collider experiments.

1.1.5 Electroweak Unification

The weak interaction was first observed in radioactive β decay and was initially interpreted by Enrico Fermi as the four-fermion interaction $pe^- \rightarrow n\nu_e$. Later developments led to the observation of parity violation after a literature review by Lee and Yang [28] followed by an experiment investigating the properties of beta decay in cobalt-60 in collaboration with Wu [29]. It is now clear that parity is violated maximally with only left-handed particles (right-handed anti-particles) being subject to the interaction. Below the electroweak energy scale at the order of 100 GeV electromagnetism and the weak interaction appear as two distinct forces, while at higher energies they can be treated as components of the same force. This close relation was made clear by Glashow, Salam and Weinberg in their GSW model of the Electroweak Theory that unifies the two forces into a shared gauge group with the structure $SU(2)_L \times U(1)_Y$, where Y is the so-called weak hypercharge and the subscript L limits the interaction to only apply to left-handed particles. The weak hypercharge is described by the relation $Y = 2Q - 2I^3$, where Q is the electric charge and I^3 is the third component of the weak isospin. The left-handed fermions can be grouped into doublets of constant hypercharge, while right-handed fermions are singlets as illustrated in table 1.2. Right-handed neutrinos are not shown, as these are singlets under both $SU(2)_L$ and $U(1)_Y$ and have no couplings to other particles (and are not part of the SM).

Constructing the electroweak Lagrangian can be done analogously to sections 1.1.3 and 1.1.4. The field dynamics are a simple sum of the terms already introduced for $U(1)$ and $SU(2)$:

$$\mathcal{L}_{\text{dyn}} = -\frac{1}{4}W_{\mu\nu}^i W^{i,\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \quad (1.26)$$

	Gen. 1	Gen. 2	Gen. 3	I^3	Y
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	1/2	1/3
	u_R	c_R	t_R	-1/2	1/3
	d_R	s_R	b_R	0	4/3
				0	-2/3
Leptons	$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$	1/2	-1
	e_R	μ_R	τ_R	-1/2	-1
				0	-2

Table 1.2: Grouping of quarks and leptons in the *Electroweak Theory*. Left-handed particles become doublets under hypercharge Y , while right-handed particles are singlets.

with the field tensors

$$\begin{aligned}
W_{\mu\nu}^i &= \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon^{ijk}W_\mu^jW_\nu^k \\
B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu
\end{aligned} \tag{1.27}$$

containing the structure constant ϵ^{ijk} of $SU(2)$. The structure of the fermion sector can be expressed in terms of left- and right-handed wave functions as

$$\mathcal{L}_f = i\bar{\Psi}_L \not{D} \Psi_L + i\bar{\Psi}_{l_R} \not{D} \Psi_{l_R} + i\bar{\Psi}_Q \not{D} \Psi_Q + i\bar{\Psi}_{u_R} \not{D} \Psi_{u_R} + i\bar{\Psi}_{d_R} \not{D} \Psi_{d_R}, \tag{1.28}$$

where Ψ_L contains the left-handed lepton doublets of $SU(2)$ and Ψ_{l_R} the right-handed lepton singlets of $SU(2)$, while Ψ_Q are the left-handed quark $SU(2)$ doublets and Ψ_{u_R} and Ψ_{d_R} are the right-handed $SU(2)$ singlets of up-type and down-type respectively. Construction of the covariant derivative is done by adding terms to uphold the local gauge invariance for both $SU(2)$ and $U(1)$:

$$\mathcal{D}_\mu = \partial_\mu + igI^iW_\mu^i + ig'\frac{Y}{2}B_\mu. \tag{1.29}$$

The electroweak sector thus contains four vector fields, namely W_μ^1 , W_μ^2 , W_μ^3 and B_μ , where W_μ^3 and B_μ are electrically neutral. Mixing occurs between these fields through the mechanism of electroweak symmetry breaking (see section 1.1.6) resulting in the final physically observable fields, which can be expressed in the following

way:

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \quad (1.30)$$

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2).$$

The mixing of the electrically neutral fields is parametrised by the Weinberg angle θ_W , leading to the fields associated with the photon A_μ and the weak neutral current Z_μ , while the weak charged currents become W_μ^\pm . This mixing structure also means that the electric charge becomes a combination of the couplings g and g' from the $U(1)_Y$ and $SU(2)_L$ groups respectively:

$$e = \frac{gg'}{\sqrt{g^2 + (g')^2}} = g' \cos\theta_W. \quad (1.31)$$

So far the theory has introduced the experimentally observed interactions between fermions and the electroweak bosons as well as self-coupling terms between the weak bosons through the additional term in the $SU(2)_L$ field tensor. However, the observed weak bosons are massive, while the above description cannot support mass terms since this would break local gauge invariance as mentioned in section 1.1.4. Fermion masses are also not allowed due to the parity violating features of the weak interaction resulting in different transformation properties of left- and right-handed particles, which violates local gauge invariance in the mass terms $\bar{\Psi}_L \Psi_R + \Psi_L \bar{\Psi}_R$ that will be discussed further in section 1.1.6. If such mass terms are introduced ‘by hand’, the resulting theory becomes unrenormalisable and thus meaningless due to the inability to treat infinities. The GSW model introduces the concept of spontaneous symmetry breaking in order to preserve the gauge invariance with mass terms included and to explain the observed mixing of the symmetry groups.

1.1.6 Electroweak Symmetry Breaking

Spontaneous symmetry breaking describes the concept of a symmetric system spontaneously transitioning into a ground state that no longer reflects the full symmetry. A simplified example in two dimensions is a flexible rod that has a constant force applied parallel to its length axis at both ends in opposite directions (always pushing inward). The rod starts out unbent. Let the two coordinates be the radial distance from the centre and the angle around the length axis of the central point on the rod. At the outset the rod is rotationally symmetric around this axis, but when the force

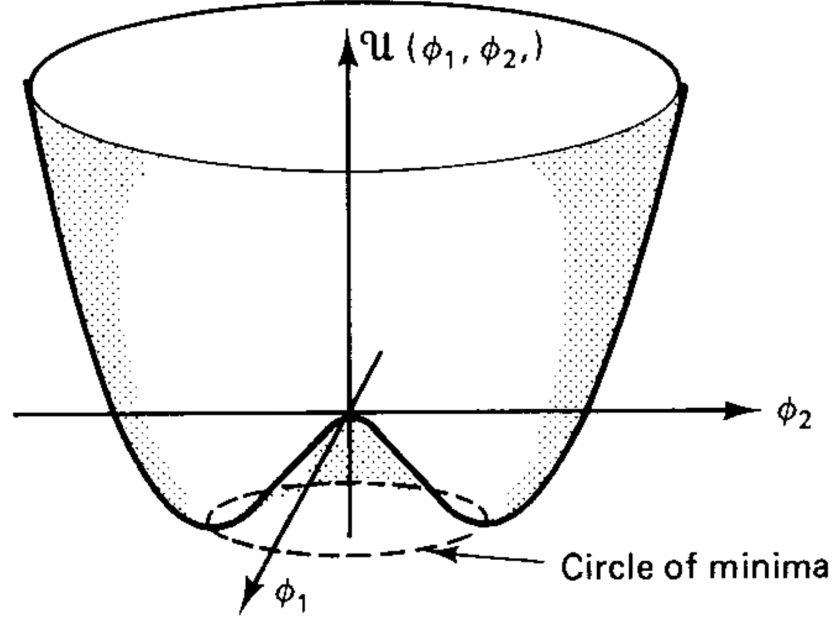


Figure 1.3: Two-dimensional Higgs potential, also colloquially called the Mexican hat potential, with $\lambda > 0$ and $\mu^2 < 0$. The stable minima are located in a circle with nonzero field values [30, p. 365].

is applied this configuration is no longer a ground state, only a stationary point in the energy of the system. The rod subsequently bends in a random direction to reach a new ground state breaking the rotational symmetry, which is now ‘hidden’.

This kind of symmetry breaking can be used in the context of the SM to break the symmetry of the subgroup related to the gauge bosons that are required to have nonzero masses, in this case the weak bosons of $SU(2)_L$. An additional complex scalar isospin doublet with hypercharge $Y = 1$ is introduced as [31, p. 334]:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.32)$$

with a Lagrangian that can be written in the form

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} &= (\mathcal{D}_\mu \phi)^\dagger (\mathcal{D}^\mu \phi) - V(\phi) \\ &= (\mathcal{D}_\mu \phi)^\dagger (\mathcal{D}^\mu \phi) + \mu^2 (\phi^\dagger \phi) - \frac{\lambda}{4} (\phi^\dagger \phi)^2, \end{aligned} \quad (1.33)$$

where \mathcal{D}_μ is defined as in equation 1.29. Choosing $\lambda > 0$ and $\mu^2 < 0$ the vacuum

expectation value (vev) of the field is nonzero (see figure 1.3 for an illustration) and is found to be

$$\phi_0^\dagger \phi_0 = \frac{v^2}{2}, \quad v = 2\sqrt{\frac{\mu^2}{\lambda}}, \quad (1.34)$$

whereby the full symmetry of the system is spontaneously broken. By expanding around the vev in a particular point corresponding to what is called the unitary gauge, the field can be written as

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \quad (1.35)$$

The unitary gauge leaves only one massive field $h(x)$ (the Higgs boson) while eliminating what would otherwise have been massless Goldstone bosons that arise due to ‘directions’ in the Lagrangian potential that are flat and hence have no resistance to excitations of the field. After substituting the vev ϕ_0 into 1.33 and performing some algebra the relevant terms that include the weak currents become

$$\begin{aligned} (\mathcal{D}'_\mu \phi)^\dagger (\mathcal{D}'^\mu \phi) &= \left(\frac{1}{2}vg\right)^2 W_\mu^+ W^{-,\mu} + \frac{1}{8}v^2 \begin{pmatrix} W_\mu^3 & B_\mu \end{pmatrix} \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} W^{3\mu} \\ B^\mu \end{pmatrix} \\ &= \left(\frac{1}{2}vg\right)^2 W_\mu^+ W^{-,\mu} + \frac{1}{8}v^2 \left[gW_\mu^3 - g'B_\mu \right]^2 + 0 \left[g'W_\mu^3 + gB_\mu \right]^2, \end{aligned} \quad (1.36)$$

where $\mathcal{D}'_\mu = igI^i W_\mu^i + ig'\frac{Y}{2}B_\mu$, removing the ∂_μ from the covariant derivative. Through the relation $W^\pm = (W^1 \mp iW^2)/\sqrt{2}$ we can identify the mass of the W boson through the expected mass term for a charged boson $M_W^2 W^+ W^-$ to be

$$M_W = \frac{1}{2}vg. \quad (1.37)$$

For the sake of clarity the second line of equation 1.36 has been written in terms of the eigenvalues of the matrix, one of which is zero. Since the fields Z_μ and A_μ diagonalise the mass matrix we can identify the remaining terms in equation 1.36

as $\frac{1}{2}M_Z^2 Z_\mu^2 + \frac{1}{2}M_A^2 A_\mu^2$ giving us

$$\begin{aligned} A_\mu &= \frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2 + g'^2}}, & M_A &= 0 \\ Z_\mu &= \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}}, & M_Z &= \frac{1}{2}v\sqrt{g^2 + g'^2}. \end{aligned} \quad (1.38)$$

So the above choice of isospin doublet has the appropriate effect of keeping the gauge boson associated with $U(1)_{em}$ massless. The generator of this symmetry is

$$Q = \frac{Y}{2} + I^3, \quad (1.39)$$

and the vacuum is invariant under the $U(1)$ transformation $\phi_0 \rightarrow e^{i\alpha(x)Q}\phi_0$ as required.

1.1.7 Masses of the Fermions

Another attractive feature of the SM is that it can simultaneously accommodate the fermion masses alongside the generation of the gauge boson masses. In order to preserve the gauge invariance of the Lagrangian this is done through Yukawa couplings between the Higgs $SU(2)$ doublet, the left-handed fermion $SU(2)$ doublets and the right-handed fermion singlets. Figure 1.4 shows such a coupling at tree-level. The Higgs doublet has exactly the required weak isospin and hypercharge to participate in such an interaction. In the unitary gauge after spontaneous symmetry breaking, the term for a lepton can be written

$$\mathcal{L}_{\text{lepton}} = -\frac{G_f}{\sqrt{2}}(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)(v + h), \quad (1.40)$$

giving the lepton mass as

$$m_{\text{lepton}} = \frac{G_f v}{\sqrt{2}}. \quad (1.41)$$

Hence, the theory can include fermion masses, but since G_f is an arbitrary coupling constant the mass value cannot be predicted.

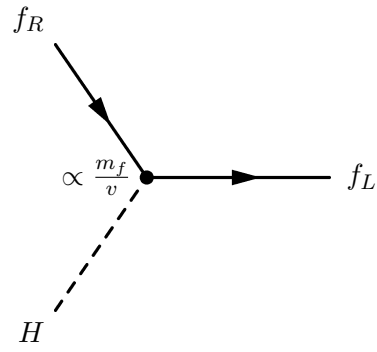


Figure 1.4: Tree-level Feynman diagram of the Higgs boson coupling to fermions. The coupling strength is proportional to the fermion mass m_f .

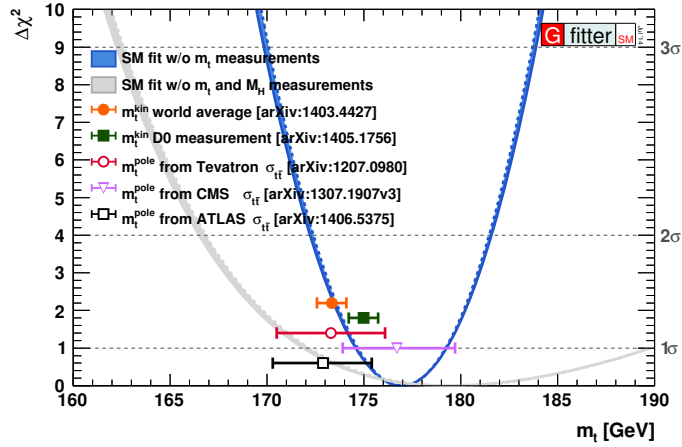


Figure 1.5: Direct measurements of the top mass from a range of experiments as well as estimates from SM fits including and excluding the direct measurements [41, 42].

1.1.8 Unsolved Problems in the Standard Model

The Standard Model is an extremely powerful predictive framework which is compatible with experimental data in all areas except in the case of neutrino masses in recent years. It is highly accurate in estimating the cross sections of its interactions, and it has successfully predicted the existence of the b -quark [32, 33], the W and Z bosons [34–37], the top quark [38, 39], the τ neutrino [40] and the Higgs boson [20, 21]. Many of its free parameters have been experimentally determined both directly and indirectly to the extent of being over-constrained with a high level of consistency between the estimates. As an example, figure 1.5 shows several direct measurements of the top mass from current and previous experiments as well as SM fits including and excluding these measurements. All measurements and fit results are consistent within the errors. The precision of theoretical predictions has also improved through new techniques that have made higher order cross section calculations feasible. However, there are theoretical considerations that make it improbable that the SM is the final answer in the search for the laws governing particle physics. This problem goes much deeper than just the glaring absence of gravity in the theory. Many aspects of the framework such as the number of particles and generations as well as the relative scale of the particle masses were not implemented based on rigorous theoretical insights, but in order to describe the observed data phenomenologically, which they do with impressive accuracy. Many fundamental problems are still left unsolved, some of which will be discussed briefly in the following.

Baryon Asymmetry in the Universe

The observable universe is heavily dominated by matter. For this matter-antimatter asymmetry to occur, the three Sakharov conditions [43] need to be fulfilled. Among these is the requirement that baryon number violation must occur, which does not happen in any SM processes. Additionally, processes that violate C and CP symmetries must exist, but the amount of CP violation present in the weak interaction is too small to explain the degree of baryon asymmetry that is observed. These considerations hint at physics beyond the SM that incorporate stronger CP violation and baryon number violation, which might become relevant at higher energies than those currently experimentally accessible.

Masses and Fine-tuning of Parameters

The masses of the fermions are parameters of the theory which are not known a priori and need to be added by hand from experimental data. Furthermore, the observed hierarchy of masses between the individual particles and between the three generations has not been explained. Such a model based on fundamental principles has long been sought within the physics community. As mentioned in section 1.1.1 even the neutrinos seem to have mass. They were previously thought to exist only in their left-handed state making it impossible to add a renormalisable mass term [23, p. 713-715].

Several parameters in the SM seem to be fine-tuned, including the bare Higgs mass (the mass at infinitesimal distances). It receives large quantum loop corrections from W , Z , H and top quark loop contributions that diverge quadratically with the renormalisation scale used in the procedure, which can be interpreted as the energy at which the SM breaks down. Although counterterms are guaranteed to cancel these divergencies, choosing values for the renormalisation scale near the Planck scale (10^{19} GeV) requires the tree diagram to have almost exactly the same value (difference needs to be order 10^{17} smaller) to counteract the loop contributions and arrive at a Higgs boson mass at the electroweak scale. This is an example of a ‘fine-tuning’ problem.

Several experiments have observed neutrino disappearance effects indicating that neutrinos can oscillate between the physically observable flavour states in direct contradiction with the SM prediction of massless neutrinos. The flux of electron neutrinos ν_e from the sun has been studied by the Homestake [44] and GNO [45] experiments, anti-neutrinos $\bar{\nu}_e$ generated by reactors are the focus of other experiments such as KamLAND [46], and atmospheric neutrinos can be detected by e.g.

Super-Kamiokande [47]. Disappearance effects have been observed in all of these cases where the observed neutrino flux is lower than the SM prediction. Appearance of surplus amounts of ν_e and ν_τ in beams of muon neutrinos has also been detected [48, 49].

Neutrino oscillations can be described theoretically by defining mass eigenstates $|\nu_i\rangle$ that are linear combinations of the flavour eigenstates $|\nu_\alpha\rangle$:

$$|\nu_i\rangle = \sum_j U_{\alpha i} |\nu_\alpha\rangle. \quad (1.42)$$

The probability of a particular flavour oscillation $|\nu_\alpha\rangle \rightarrow |\nu_{\alpha'}\rangle$ depends on the size of the mass difference $\Delta m_{\alpha,\alpha'}^2$, the distance travelled since the flavour eigenstate was created L , the energy of the system E and the mixing matrix U . Assuming there are three mass eigenstates, fits to experimental data are able to constrain the mass differences between each state resulting in [5]

$$\begin{aligned} |\Delta m_{12}^2| &\approx 7.5 \cdot 10^{-5} \text{ eV}, \\ |\Delta m_{31}^2| &\approx 2.3 \cdot 10^{-3} \text{ eV}, \\ |\Delta m_{21}^2|/|\Delta m_{31}^2| &\approx 0.032. \end{aligned} \quad (1.43)$$

Due to their lack of an electric charge it is still an open question whether neutrinos are their own antiparticle (Majorana fermion) or if they exhibit the same behaviour as the charged leptons having distinct antiparticles (Dirac fermion). In the former case lepton number is not a conserved quantity, which opens the possibility of experimentally determining their Majorana or Dirac nature through searches for e.g. neutrinoless double beta decay.

Unification of the Forces

The coupling constants of the three fundamental forces vary with the cut-off energy scale chosen when performing renormalisation. Plotting the coupling strengths as a function of the energy scale as seen in figure 1.6, the coupling strengths seem to coalesce into one unified force at very high energies of $\sim 10^{16}$ GeV. This has led to several Grand Unified Theories (GUTs) postulating that at energies above this GUT scale the three SM symmetries are replaced by a new single symmetry. Within the SM the strengths of the forces do not meet at exactly the same point. One way to make this happen is by introducing another symmetry which sets fermions and bosons on equal footing. This is called supersymmetry [50, 51] and is one out of

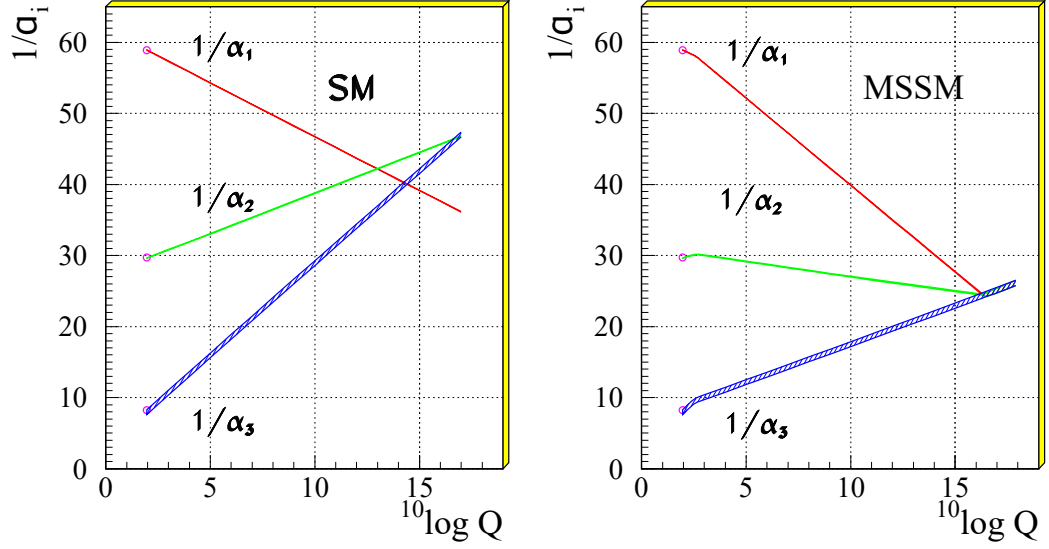


Figure 1.6: The evolution of the inverse SM coupling constants (left) and in the minimal supersymmetric extension of the SM (right). Unification of all three forces only happens with the inclusion of supersymmetry. The α_i represent, in numerically ascending order, the EM, weak and strong coupling constants [52].

many contenders for a theory beyond the SM.

Dark Matter

Evidence for the existence of dark matter has been found primarily in measurements of the properties of galaxies. Their rotational velocity as a function of the distance from their centre, the strength of gravitational lensing and the way galaxies cluster together cannot be explained by the presence of only visible matter. Dark matter is hypothesised to be a new form of matter that acts weakly or not at all through the electromagnetic force and which is stable over the age of the universe. Fluctuations in the cosmic microwave background (CMB) also provide evidence for dark matter. According to fits from the Planck Collaboration [53] the universe contains only 4.9% ordinary matter, while 26.8% is composed of dark matter. The remaining energy density is called dark energy and can be described as a cosmological constant that acts to accelerate the expansion of the universe. Models exist to describe the nature of dark matter including weakly interacting massive particles (WIMPs), axions and sterile neutrinos. Dark matter candidates can also be constructed through supersymmetric models that can be searched for at the LHC. Experimental evidence has not yet resolved which of these models, if any, are an accurate description of dark matter.

1.2 Higgs Phenomenology

Experimental analyses studying the Higgs boson are based on the predictions of the SM describing its production and decay modes and the properties and couplings described in the previous sections. Expected production cross sections and decay branching fractions are used in order to optimise search strategies and compare the results with the Standard Model predictions. The following sections describe the dominant Higgs production mechanisms in proton-proton collisions at the LHC as well as its decay channels.

1.2.1 Higgs Boson Production at the LHC

In proton-proton collisions the Standard Model Higgs boson is most abundantly produced in gluon fusion processes (ggF) for Higgs masses below 1 TeV due to the large gluon contribution to the proton PDF at low Bjorken x values. The main contribution comes from the top quark due to its large Yukawa coupling (coupling between a scalar and a Dirac field) to the Higgs boson. The leading order gluon fusion production is characterised by the decay products being closely back-to-back. In the case of decays to tau leptons, which is the focus of this thesis, this results in a low E_T^{miss} signature stemming from the undetected neutrinos in the τ decays. At NLO or higher the gluon fusion process can happen in association with jets, potentially boosting the Higgs in the transverse plane. This means that the E_T^{miss} from the neutrino coming from each τ will no longer cancel each other, and this higher E_T^{miss} signature can help discriminate events with Higgs production from background events. Another significant contribution arises through vector boson fusion (VBF), which has a cross section roughly one order of magnitude lower than ggF. The VBF process starts with the emission of a weak vector boson from a quark in each colliding proton that annihilate or ‘fuse’ into a Higgs boson. The VBF production channel involves some very pronounced effects that can be exploited to identify the events. The signature contains a high- p_T , high-invariant mass jet pair with large pseudorapidity separation. It is kinematically predicted that the Higgs boson will usually be situated in the central region of the detector with a high level of isolation from nearby jet activity. This makes VBF a viable channel to investigate even though the production cross section is considerably lower than ggF. Production in association with a vector boson $V = Z, W$ (VH), also called Higgs-strahlung, happens when a quark pair annihilates into a weak vector boson that subsequently radiates a Higgs boson. Features of this channel include a boost of the Higgs in the transverse plane and associated leptons ($l\nu/l^+l^-$) or jets being

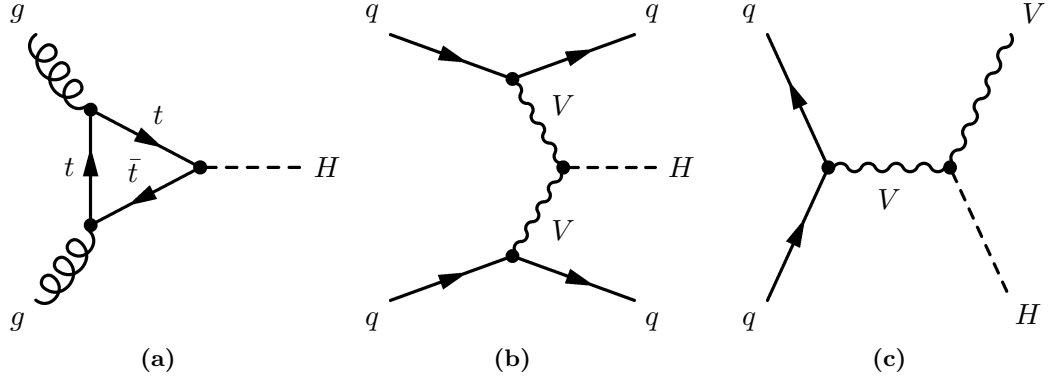


Figure 1.7: Feynman diagrams of the most common production channels of the Higgs boson at tree level at the LHC. Diagram (a) shows gluon fusion via a top loop (ggF). Diagram (b) gives the topology for Vector Boson Fusion (VBF), where $V = Z, W^+, W^-$. Diagram (c) shows associated production (VH).

decay products of the associated Z/W boson. The production cross section for this channel is much lower than both ggF and VBF and is therefore not expected to contribute extensively to the Higgs signal. It does however give access to additional Higgs couplings and is clean (low or isolated activity from other particles/jets in the detector) when the vector bosons decay leptonically ($Z \rightarrow ll$, $W \rightarrow l\nu$). Tree-level Feynman diagrams for these three production channels are shown in figure 1.7. Production in association with a top quark pair ($t\bar{t}H$) is also present but has a limited cross section as a consequence of the two heavy top quarks. It is not expected to contribute in a significant way to the signal region treated in this thesis and was previously determined to be negligible [2].

All cross section calculations used in this thesis follow the recommendations of the Higgs cross section working group, which provides theory calculations and SM predictions of Higgs phenomenology to all experiments at the LHC. The reports in [54–56] outline the current status of developments in higher order corrections on Higgs boson production cross sections and decay branching ratios.

The gluon fusion cross section is heavily dependent on QCD radiative corrections with NLO corrections in α_s contributing 80–100% relative to the LO cross section. The NNLO contribution is added in the high top quark mass limit including improvements by resumming the soft gluon contributions up to NNLL, resulting in a further increase of the cross section of approximately 30%. Additionally, electroweak corrections are applied at the two-loop level and strongly depend on the Higgs boson mass. Including these corrections a Higgs mass of $m_H = 125$ GeV yields an inclusive ggF production cross section of $\sigma_{\text{ggF}} = 19.27$ pb. Missing higher order corrections

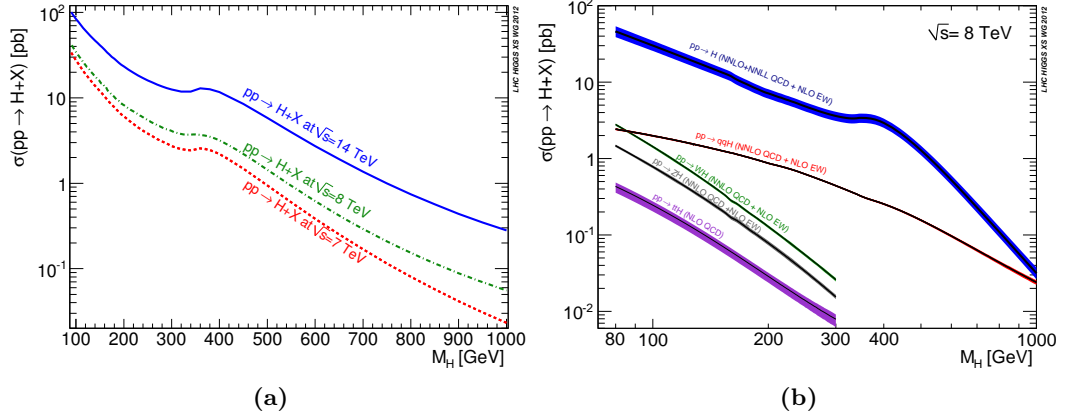


Figure 1.8: Higgs boson production cross sections at the LHC as a function of the Higgs boson mass m_H shown for (a) inclusive production at $\sqrt{s} = 7, 8$ and 14 TeV and (b) split into the main contributions at $\sqrt{s} = 8$ TeV including NNLO QCD and NLO EW corrections [54–56].

contribute to the theoretical uncertainty on this number, as do the uncertainties on the PDF, the numerical value of α_s and finite quark mass effects.

The VBF cross section is fully evaluated at NLO precision in both the strong and electroweak couplings. QCD corrections are of the order of 5–10%, while the electroweak corrections are evaluated to negatively affect the cross section by 5% [54]. Approximate NNLO QCD corrections are known using a structure function approach that is accurate up to the level of certain interference effects, and this additional correction reduces the remaining scale dependence to 1–2%.

The WH and ZH cross sections include full NNLO QCD corrections as well as NLO electroweak corrections. The associated scale uncertainties on the cross sections are in the 1–3% range.

The inclusive Higgs production cross section is shown in figure 1.8 (a) at

Process	$\sigma_{\text{incl.}}$ [pb]	QCD scale [%]	PDF+ α_s [%]
ggF	19.27	+7.2 – 7.8	+7.5 – 6.9
VBF	1.58	+0.2 – 0.8	+2.6 – 2.9
WH	0.705	± 1	± 2.3
ZH	0.415	± 3.1	± 2.5

Table 1.3: Inclusive Higgs boson production cross sections in proton-proton collisions and the uncertainties arising from variations in the QCD scale, the PDF and the numerical value of the strong coupling constant. The numbers are given at $\sqrt{s} = 8$ TeV and $m_H = 125$ GeV with corrections at NNLO in QCD and NLO in EW [56].

center-of-mass energies of 7, 8 and 14 TeV. Figure 1.8 (b) includes the cross sections of the individual production modes discussed above as a function of the Higgs boson mass. Table 1.3 summarises the same production cross sections specifically at $m_H = 125$ GeV and also lists their theoretical scale and PDF uncertainties.

1.2.2 Higgs Boson Decays

The 2012 discovery of the Higgs boson was made exclusively in the bosonic decay channels, mainly $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ(\rightarrow 4l)$, but also with a contribution from $H \rightarrow WW \rightarrow l^+\nu l^-\bar{\nu}$. If the newly discovered boson is truly the Higgs boson of the Standard Model, it should also decay to fermions for which the experimental conditions at the LHC are less ideal. At the experimentally determined Higgs boson mass of approximately 125 GeV the decay rates to electrons and muons are predicted to be very low. The most copious hadronic decay mode is $H \rightarrow b\bar{b}$, but this signal has to compete with the noisy environment of abundant jet production from background processes. $H \rightarrow \tau\tau$ has a relatively high branching ratio, but due to the short lifetime of the τ leptons they generally decay close to the beam pipe and not in the silicon detectors and need to be identified through the impact parameter of their displaced decay products consisting of one or more neutrinos and either a light lepton or hadrons. An extra reconstruction step therefore adds to the complexity of the analysis, and a considerable fraction of the decay energy will be in the form of neutrinos leaving only missing transverse energy E_T^{miss} . The achievable mass resolution is therefore greatly reduced compared to the majority of other channels.

Since the Higgs boson couplings to vector bosons and fermions are proportional to the masses of the particles as discussed in the previous sections, the branching fractions of Higgs boson decays are determined by the masses of the decay products. The total and partial decay widths are determined while including higher order corrections from QCD and electroweak processes. The partial widths are computed using the two software programs HDECAY [57, 58] and PROPHECY4F [59, 60] that include the highest order calculations available for each process. In particular, the $H \rightarrow \tau\tau$ branching ratio is calculated at NNNLO in QCD and NLO in EW corrections. At $m_H = 125$ GeV it contributes 6.3% to the total decay width and has a theoretical uncertainty of about $\pm 6\%$ [55]. Figure 1.9 (a) shows the SM branching ratios of the Higgs boson as a function of m_H including theoretical error bands. $H \rightarrow b\bar{b}$ decays dominate at low Higgs boson masses, while $H \rightarrow WW$ decays are most abundant when approaching $m_H \approx 2m_W$ and continue to dominate over a wide range up to 1 TeV.

The decay channels that are useful in experimental analyses depend on the

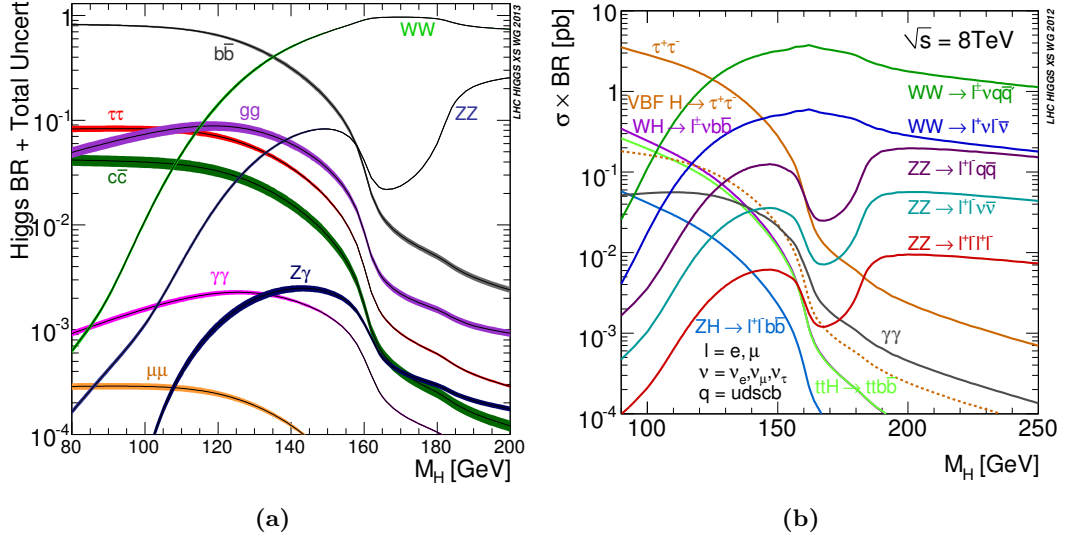


Figure 1.9: Standard Model branching ratios of the Higgs boson as a function of m_H (a) and the production cross section times the decay branching ratio $\sigma \times BR$ of the processes corresponding to the main experimental Higgs search channels at the LHC (b). The inclusive (solid orange line) and VBF (dashed orange line) modes are included for $H \rightarrow \tau\tau$, which is seen to have a significant contribution to $\sigma \times BR$ at a Higgs boson mass of $m_H = 125$ GeV [56].

specific final state particles and the ability to trigger efficiently on them as well as the signal to background ratio that can be achieved. Figure 1.9 (b) shows the product of the production cross section and the branching ratio $\sigma \times BR$ of the main experimental search channels at the LHC. Events in the $H \rightarrow ZZ^* \rightarrow 2l2l'$ channel with l, l' each denoting either an electron or a muon can be fully reconstructed, and the Z decays contain polarisation information that is useful in measurements of spin and CP quantum numbers, since the distributions of such variables depend on the spin of the Higgs and the presence of new BSM couplings in the Higgs Lagrangian. The channel provides a high mass resolution since both the muon momenta and electron energies can be reconstructed at high precision, and it has relatively low background levels since well-chosen selection criteria can effectively discriminate against hadronic processes. It does, however, require a large data sample due to the relatively low $\sigma \times BR$. A decay mode that can also be fully reconstructed is $H \rightarrow \gamma\gamma$. Although its $\sigma \times BR$ is relatively low, the clean signature of two photons in the EM calorimeter has excellent energy resolution and a low misidentification rate from hadronic processes. The discovered Higgs boson with $m_H = 125$ GeV lies serendipitously within the accessible mass range of 110–140 GeV in this channel and

it offers important contributions to the signal sensitivity and mass resolution.

Another important decay channel is $H \rightarrow WW^* \rightarrow l^+\nu l^-\bar{\nu}$, which has a large branching fraction but cannot be fully reconstructed due to the two neutrinos in the event leading to poor mass resolution. It is nonetheless much cleaner than final states with hadronic W decays due to the much higher rejection of $W + \text{jets}$ events in the final state where both W bosons decay leptonically. The remaining significant background contributions stem from diboson and top pair production. As with $H \rightarrow ZZ$ it is possible to exploit angular distributions and other kinematics to study spin and CP.

While being one of only two decay channels accessible at current luminosities that are able to probe the fermionic Higgs couplings, the fully hadronic decay channel $H \rightarrow b\bar{b}$ is complicated by the large backgrounds from multijet processes that reduce the achievable signal significance. Backgrounds can be reduced significantly in cases where it is possible to trigger on a leptonic signature such as in the production modes in association with a W or a Z boson.

The $H \rightarrow \tau\tau$ mode is the other fermionic decay channel with a sufficiently high branching ratio to be utilised in analyses at current luminosities. The VBF $H \rightarrow \tau\tau$ channel alone has a $\sigma \times \text{BR}$ on par with or greater than several of the bosonic decay channels at $m_H = 125$ GeV. The main background comes from $Z \rightarrow \tau\tau$ decays having similar kinematic properties to those of the signal process. However, the event topology of VBF production having two distinct and spatially separated jets aids in the suppression of this and other backgrounds. $H \rightarrow \tau\tau$ is the only decay channel with a sufficiently high branching ratio to be able to probe the Higgs couplings to leptons with the amount of data collected in Run 1 and is therefore the only channel that is currently able to probe the Higgs boson Yukawa couplings to leptons. For more details on the branching fractions of tau decays, see section 3.6.

1.3 CP Violation in VBF Higgs Production

If Higgs boson decays to tau leptons are discovered at a mass value consistent with the bosonic decay channels, it will still be necessary to investigate whether the rest of its properties are in agreement with the predictions of the SM. One such property is represented by the particle's CP quantum numbers. Violation of CP is already known to take place in certain processes such as neutral kaon decays [24], but the SM predicts that the Higgs should respect the symmetry. However, this needs to be verified.

It is theoretically possible to have CP-odd terms in both the production and

decay vertices of the Higgs. In this analysis only Higgs production is investigated, specifically the feasibility of detecting CP-odd contributions to the Vector Boson Fusion (VBF) production mode. By analysing e.g. angular distributions in the Higgs production it is possible to determine whether any BSM (beyond Standard Model) properties are involved. Ultimately, consistency of the CP properties in VBF production between $H \rightarrow \tau\tau$ and other decay channels also needs to be demonstrated in order to verify that the data originates from the same unique particle.

1.3.1 Theoretical Model of CP Violation

The theoretical model describing CP mixing in the Higgs system considers an effective Lagrangian built from the SM Lagrangian with additional CP-violating operators. These operators are chosen to be of mass dimension six, as this is the lowest order that can appear for new physics when lepton number is conserved [61]. The mass dimension is the dimensionality of a variable in mass/energy, offset by factors of $1/\Lambda^2$, where Λ the scale of new physics, to preserve the dimensionality of the Lagrangian ($[\mathcal{L}] = E^{+4}$). They are constructed using the Higgs doublet Φ and the $U(1)_Y$ and $SU(2)$ electroweak gauge fields B^μ and $W^{a,\mu}$ ($a = 1,2,3$) respectively. The construction implies that all other interactions between the Higgs boson and SM particles, apart from those involving electroweak bosons, adhere to the SM predictions. Consequently, Higgs production through gluon fusion and decays of the Higgs into a pair of τ leptons are both considered to be SM-like.

As described in [61] we can specify the effective $U(1)_Y$ - and $SU(2)$ -invariant Lagrangian in the following way:

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \frac{f_{\tilde{B}B}}{\Lambda^2} \mathcal{O}_{\tilde{B}B} + \frac{f_{\tilde{W}W}}{\Lambda^2} \mathcal{O}_{\tilde{W}W} + \frac{f_{\tilde{B}}}{\Lambda^2} \mathcal{O}_{\tilde{B}} \quad (1.44)$$

having the three dimension-six operators

$$\mathcal{O}_{\tilde{B}B} = \Phi^\dagger \hat{B}_{\mu\nu} \hat{B}^{\mu\nu} \Phi \quad (1.45)$$

$$\mathcal{O}_{\tilde{W}W} = \Phi^\dagger \hat{W}_{\mu\nu} \hat{W}^{\mu\nu} \Phi \quad (1.46)$$

$$\mathcal{O}_{\tilde{B}} = (D_\mu \Phi)^\dagger \hat{B}^{\mu\nu} D_\nu \Phi. \quad (1.47)$$

D_μ denotes the covariant derivative $D_\mu = \partial_\mu + \frac{i}{2}g'B_\mu + ig\frac{\sigma_a}{2}W_\mu^a$, $\hat{V}_{\mu\nu}$ ($V = B, W^a$) represents the field strength tensors and $\tilde{V}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}V^{\rho\sigma}$ the dual field strength tensors, with $\hat{B}_{\mu\nu} + \hat{W}_{\mu\nu} = i\frac{g'}{2}B_{\mu\nu} + i\frac{g}{2}\sigma^a W_{\mu\nu}^a$. Λ is the scale of new physics.

The operator $\mathcal{O}_{\tilde{B}}$ contributes to the CP-violating charged triple gauge couplings $\tilde{\kappa}_\gamma$ ($\tilde{\kappa}_Z$) via the relation $\tilde{\kappa}_\gamma = \cot^2 \theta_W \tilde{\kappa}_Z = -\frac{m_W^2}{2\Lambda^2} f_{\tilde{B}}$. The ALEPH [62]

and OPAL [63] experiments have previously constrained these CP-violating charged triple gauge couplings, and the contribution from $\mathcal{O}_{\tilde{B}}$ is neglected in the following. Only contributions from $\mathcal{O}_{\tilde{B}B}$ and $\mathcal{O}_{\tilde{W}W}$ are taken into account.

After electroweak symmetry breaking in the unitary gauge the effective Lagrangian in the mass basis of Higgs boson H , photon A and Z boson Z can be written as [64]

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \tilde{g}_{HAA} \tilde{A}_{\mu\nu} A^{\mu\nu} + \tilde{g}_{HAZ} \tilde{A}_{\mu\nu} Z^{\mu\nu} + \tilde{g}_{HZZ} \tilde{Z}_{\mu\nu} Z^{\mu\nu} + \tilde{g}_{HWW} \tilde{W}_{\mu\nu}^+ W^{-,\mu\nu}. \quad (1.48)$$

Only two of the four couplings \tilde{g}_i are independent due to constraints imposed by $U(1)_Y \times SU(2)$ invariance. They can be expressed in terms of two dimensionless couplings \tilde{d} and \tilde{d}_B as:

$$\tilde{g}_{HAA} = \frac{g}{2m_W} (\tilde{d} \sin^2 \theta_W + \tilde{d}_B \cos^2 \theta_W) \quad (1.49)$$

$$\tilde{g}_{HAZ} = \frac{g}{2m_W} \sin 2\theta_W (\tilde{d} - \tilde{d}_B) \quad (1.50)$$

$$\tilde{g}_{HZZ} = \frac{g}{2m_W} (\tilde{d} \cos^2 \theta_W + \tilde{d}_B \sin^2 \theta_W) \quad (1.51)$$

$$\tilde{g}_{HWW} = \frac{g}{m_W} \tilde{d}. \quad (1.52)$$

Hence in general WW , ZZ , $Z\gamma$ and $\gamma\gamma$ fusion contribute to VBF production. The relations between \tilde{d} and $f_{\tilde{B}B}$, and \tilde{d}_B and $f_{\tilde{W}W}$ are given by:

$$\tilde{d} = -\frac{m_W^2}{\Lambda^2} f_{\tilde{W}W} \quad \tilde{d}_B = -\frac{m_W^2}{\Lambda^2} \tan^2 \theta_W f_{\tilde{B}B}. \quad (1.53)$$

Since the different contributions from the various electroweak gauge boson fusion processes cannot be distinguished experimentally, the arbitrary choice $\tilde{d} = \tilde{d}_B$ is adopted. This yields the following relation for the \tilde{g}_i :

$$\tilde{g}_{HAA} = \tilde{g}_{HZZ} = \frac{1}{2} \tilde{g}_{HWW} = \frac{g}{2m_W} \tilde{d} \quad \text{and} \quad \tilde{g}_{HAZ} = 0. \quad (1.54)$$

The parameter \tilde{d} is related to the parameter $\hat{\kappa}_W$ used in the investigation of CP properties in the decay $H \rightarrow W^+ W^-$ [65] via $\tilde{d} = -\hat{\kappa}_W = \tilde{\kappa}_W / \kappa_{SM} \tan \alpha$. The choice $\tilde{d} = \tilde{d}_B$ yields $\hat{\kappa}_W = \hat{\kappa}_Z$ as assumed in the combination of the $H \rightarrow W^+ W^-$ and $H \rightarrow ZZ$ decay analyses [65].

The effective Lagrangian contains the following Lorentz structure for each vertex of the Higgs boson to two identical electroweak gauge bosons $HV(p_1)V(p_2)$

vertex ($V = W, Z, \gamma$), with $p_{1,2}$ denoting the momenta of the gauge bosons:

$$T^{\mu\nu}(p_1, p_2) = \sum_{V=W,Z} \frac{2m_V^2}{v} g^{\mu\nu} + \sum_{V=W,Z,\gamma} \frac{2g}{m_W} \tilde{d} \varepsilon^{\mu\nu\rho\sigma} p_{1\rho} p_{2\sigma}. \quad (1.55)$$

The first terms ($\propto g^{\mu\nu}$) are CP-even and describe the SM coupling structure; the second terms ($\propto \varepsilon^{\mu\nu\rho\sigma} p_{1\rho} p_{2\sigma}$) are CP-odd and arise from the CP-odd dimension six operators. Note that the choice $\tilde{d} = \tilde{d}_B$ gives the same coefficients multiplying the CP-odd structure for HWW , HZZ and $H\gamma\gamma$ vertices and a vanishing coupling for the $HZ\gamma$ vertex.

The matrix element \mathcal{M} for VBF production receives a CP-even contribution \mathcal{M}_{SM} from the SM and a CP-odd contribution $\mathcal{M}_{\text{CP-odd}}$ from the dimension six operators considered:

$$\mathcal{M} = \mathcal{M}_{\text{SM}} + \tilde{d} \cdot \mathcal{M}_{\text{CP-odd}}. \quad (1.56)$$

The differential cross section or squared matrix element has three contributions:

$$|\mathcal{M}|^2 = |\mathcal{M}_{\text{SM}}|^2 + \tilde{d} \cdot 2\Re(\mathcal{M}_{\text{SM}}^* \mathcal{M}_{\text{CP-odd}}) + \tilde{d}^2 \cdot |\mathcal{M}_{\text{CP-odd}}|^2. \quad (1.57)$$

The first $|\mathcal{M}|^2$ and third term $\tilde{d}^2 \cdot |\mathcal{M}_{\text{CP-odd}}|^2$ are both CP-even and hence do not yield a source of CP violation. The second term $\tilde{d} \cdot 2\Re(\mathcal{M}_{\text{SM}}^* \mathcal{M}_{\text{CP-odd}})$, stemming from the interference of the two contributions to the matrix element, is CP-odd and is a possible new source of CP-violation in the Higgs sector. The interference term integrated over a CP-symmetric part of phase space vanishes and does therefore not contribute to the total cross section and observed event yield after applying CP-symmetric selection criteria. The third term quadratic in \tilde{d} leads to an increase of the total cross section. This signal rate information can increase the sensitivity to the BSM couplings at the cost of a loss of generality, since it is difficult to guarantee an accurate normalisation of the cross section in an EFT such as the one employed here. This analysis does not exploit the prediction of the signal rate to constrain the BSM couplings and instead only relies on differences in the shapes of distributions.

1.3.2 Observables Sensitive to CP Violation

It is possible to construct several variables that are sensitive to CP-odd couplings of the Higgs boson in the VBF production channel. For example, an effect can be observed on the angular distribution of the emitted jets. A more mathematically involved approach is to construct a variable using the matrix element of the process. These two observables will be defined and discussed below.

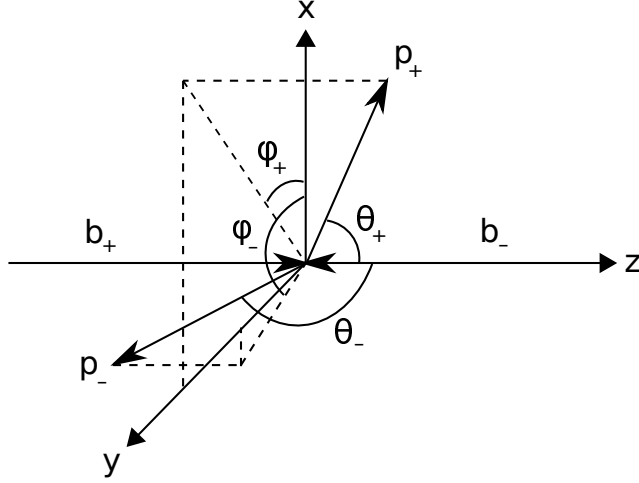


Figure 1.10: Illustration of the jet angles in a VBF event.

Signed Azimuthal Angle $\Delta\phi_{jj}^{\text{sign}}$

The forward and backward tagging jets are characteristic features of the VBF production process, and their distributions can be exploited to gain information about the tensor structure of the HVV vertex without being dependent on the Higgs decay mode. The signed azimuthal angle between the two tagging jets in a VBF event is approximately symmetric around zero for an exclusively SM Higgs boson, while CP-odd couplings will add a degree of asymmetry depending on the strength of these couplings. The azimuthal angle was suggested in [66] and is formally defined as

$$\epsilon_{\mu\nu\rho\sigma} b_+^\mu p_+^\nu b_-^\rho p_-^\sigma = 2p_+^T p_-^T \sin(\phi_+ - \phi_-) = 2p_+^T p_-^T \sin \Delta\phi_{jj}. \quad (1.58)$$

Here b_+^μ and b_-^μ denote the normalised four-momenta of the two proton beams and p_+^μ and p_-^μ denote the four-momenta of the two tagging jets, where p_+ (p_-) points into the same detector hemisphere as b_+^μ (b_-^μ). ϕ_+ and ϕ_- are the azimuthal angles of the tagging jets, denoted in the same way (see figure 1.10). This definition is invariant under the interchange $(b_+, p_+) \leftrightarrow (b_-, p_-)$, since this leaves the ordering of the angles unchanged, thereby avoiding the sign ambiguity otherwise encountered in $\Delta\phi_{jj}$. This ordering is kept explicit in the variable by naming it $\Delta\phi_{jj}^{\text{sign}}$. The $\Delta\phi_{jj}^{\text{sign}}$ distribution of a CP-odd coupling introduced using the second term of the tensor structure in equation 1.55 can be qualitatively evaluated by noticing the presence of a Levi-Civita tensor which only produces nonzero results when the four outgoing parton momenta in the process are independent. Back-to-back or collinear tagging jets in the transverse plane will therefore produce a vanishing matrix element, leading to a distinct signature compared to the SM coupling. A mixture of SM and

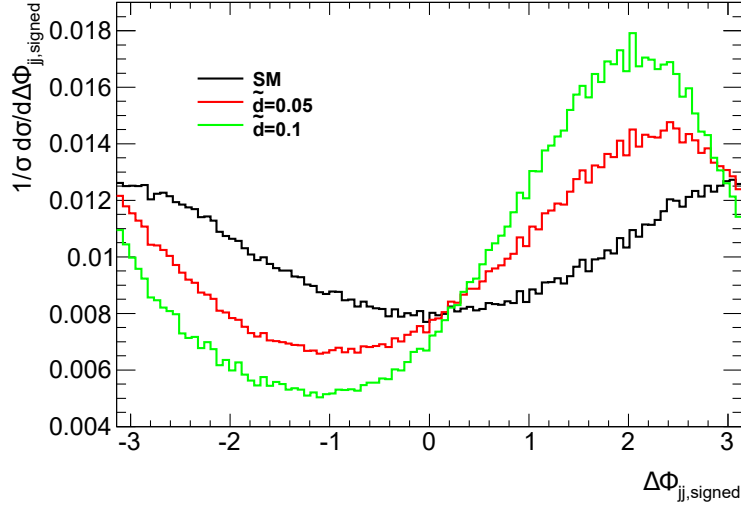


Figure 1.11: $\Delta\phi_{jj}^{\text{sign}}$ distributions of VBF Higgs signal events in the case of pure CP-even SM couplings and two CP-mixed models with $\tilde{d} = 0.05$ and $\tilde{d} = 0.1$.

CP-odd couplings will produce a distribution with characteristics of both, depending on the size of the CP-odd admixture. A measurement of \tilde{d} can be performed either by comparing the mean value of $\Delta\phi_{jj}^{\text{sign}}$ with the prediction for various \tilde{d} values or by a maximum-likelihood fit to the full distribution of $\Delta\phi_{jj}^{\text{sign}}$. Figure 1.11 shows distributions of $\Delta\phi_{jj}^{\text{sign}}$ in simulated VBF events at $\sqrt{s} = 8$ TeV in a VBF-enriched region. Distributions are shown for $\tilde{d} = 0$ (SM), 0.05 and 0.1. In the SM case the distribution is symmetric around zero, while BSM contributions show up as a progressively more asymmetric distribution according to the relative size of the BSM and SM couplings, with the direction of the asymmetry determined by the sign of \tilde{d} .

The Optimal Observable

Another observable is the so-called ‘Optimal Observable’ [67], which can be constructed from the pure SM cross section σ_{SM} and the new contribution to the cross section from CP-odd terms σ_{CP} in the following way:

$$\mathcal{O} = \frac{\sigma_{\text{CP}}}{\sigma_{\text{SM}}}. \quad (1.59)$$

The Optimal Observable approach yields a completely model-independent way of testing CP invariance. If CP invariance holds, then the mean value has to vanish, i.e. $\langle \mathcal{O}_{\text{CP}} \rangle = 0$. An observation of a non-vanishing mean value is a clear sign of CP

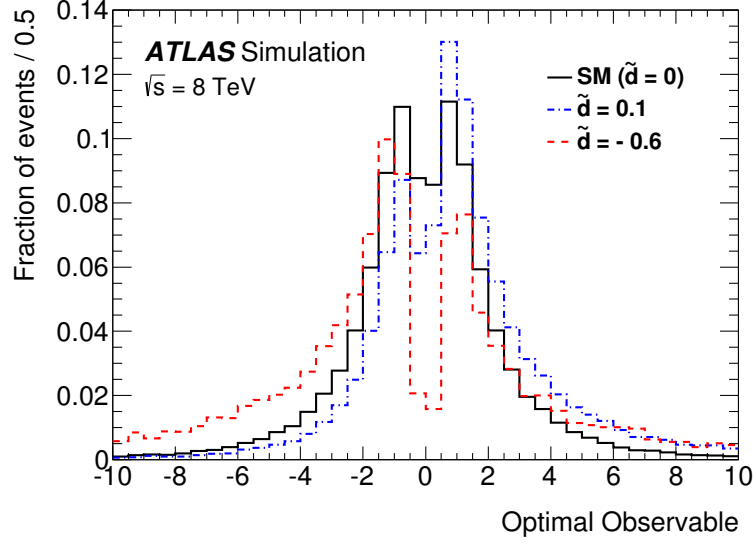


Figure 1.12: Distribution of O_1 at generator level for $\tilde{d} = 0, 0.1$ and -0.6 . The SM sample was generated using AMC@NLO [68] at leading order, and then re-weighted to different \tilde{d} values using the procedure described in section 4.3. A typical VBF selection has been applied – see table 4.3.

violation. In terms of matrix elements the Optimal Observables of first and second order are defined as follows:

$$O_1 := \frac{2\Re(\mathcal{M}_{\text{SM}}^* \mathcal{M}_{\text{CP-odd}})}{|\mathcal{M}_{\text{SM}}|^2}, \quad (1.60)$$

$$O_2 := \frac{|\mathcal{M}_{\text{CP-odd}}|^2}{|\mathcal{M}_{\text{SM}}|^2}. \quad (1.61)$$

To understand the Optimal Observable, we assume we can split the matrix element of the VBF vertex into an even and an odd part, including the parameter \tilde{d} to control the strength of the odd admixture. Squaring this matrix element gives us equation 1.57. We see that O_1 is the ratio between the SM term and the CP-odd term linear in \tilde{d} , whereas O_2 is the ratio between the SM term and the CP-odd term quadratic in \tilde{d} . The first order Optimal Observable, O_1 , has a mean value of 0 if CP is conserved, whereas if CP is violated ($\tilde{d} \neq 0$) its mean value is shifted from zero in the positive or negative direction, depending on the sign of \tilde{d} . Figure 1.12 shows the distribution of the first order Optimal Observable, both in the Standard Model case and for non-vanishing $\tilde{d} \neq 0$ values, which introduce an asymmetry in the distribution and yield a non-vanishing mean value. The second order Optimal Observable gives sensitivity to CP not only through shape differences, but also through

an increase in the signal cross section, i.e. on the signal normalisation. However, as mentioned in section 1.3.1, the use of an EFT complicates the exploitation of the increase in cross section.

The final state consisting of the Higgs boson and the two tagging jets can be characterised by seven phase space variables if the Higgs mass is specified, jet masses are neglected and momentum conservation in the plane transverse to the beam pipe is exploited. The Optimal Observable combines the information of the multidimensional phase space into a single observable, which can be shown to be the optimal variable choice in terms of sensitivity for small values of the parameter of interest when neglecting the terms in the matrix element quadratic in it. The method was first suggested for the estimation of a single parameter using the mean value only [69] and via a maximum-likelihood fit to the full distribution [70] using the first order Optimal Observable. The extension to several parameters and exploiting also the matrix element contributions quadratic in the parameters by additionally using the second order Optimal Observable was introduced in [71–73]. The technique has been applied in various experimental analyses, e.g. [65, 67, 74–78].

The values of the leading order matrix elements needed for the calculation of the Optimal Observable are extracted from HAWK [79–81]. The evaluation requires the four-vectors of the Higgs boson and the two tagging jets. The Bjorken x values of the incoming partons x_1 (x_2) in positive (negative) z -direction can be derived from the Higgs boson and tagging jet four-momenta exploiting energy-momentum conservation as:

$$x_{1/2}^{\text{reco}} = \frac{M_{\text{Hjj}}}{\sqrt{s}} e^{\pm y_{\text{Hjj}}} \quad (1.62)$$

where M_{Hjj} and y_{Hjj} are the mass and rapidity of the vectorial sum of the tagging jets and the Higgs boson four-momenta. Since the flavour of the incoming and outgoing partons cannot be determined experimentally, the sum over all possible flavour configurations $ij \rightarrow klH$ weighted by the parton distribution functions (the probability of finding a particular flavour of parton, see section 2.2.2) is calculated separately for the matrix elements in numerator and denominator:

$$\begin{aligned} |\mathcal{M}_{\text{SM}}|^2 &= \sum_{i,j,k,l} f_i(x_1) f_j(x_2) |\mathcal{M}_{\text{SM}}|^2(ij \rightarrow klH), \\ 2\Re(\mathcal{M}_{\text{SM}}^* \mathcal{M}_{\text{CP-odd}}) &= \sum_{i,j,k,l} f_i(x_1) f_j(x_2) 2\Re(\mathcal{M}_{\text{SM}}^* \mathcal{M}_{\text{CP-odd}})(ij \rightarrow klH). \end{aligned} \quad (1.63)$$

At reconstruction level, the inputs to the Optimal Observable calculation need to be replaced by their reconstructed equivalents. The following inputs are used:

- The reconstructed Higgs boson four-vector, being the sum of the two tau four-vectors obtained using the Missing Mass Calculator (MMC) [82] algorithm. The presence of neutrinos in the tau decays means that only the visible component of the tau decay can actually be reconstructed. To obtain an estimator of the full tau four-vectors, the MMC algorithm scans over each possible configuration (accounting also for the E_T^{miss} resolution), weighting it by its probability, and finally returns for each component the most probable value of the scan points. For more information see section 4.1.3.
- The leading and sub-leading reconstructed jets. The jets are reconstructed using the anti- k_t algorithm with a distance parameter $R = 0.4$, and are subject to channel-dependent selections on their p_T and pseudo-rapidity separation, discussed in section 4.6.
- The reconstructed Bjorken x obtained from the equations

$$x_{1/2}^{\text{reco}} = \frac{M_{\text{final}}}{\sqrt{s}} e^{\pm y_{\text{final}}}, \quad (1.64)$$

where M_{final} (y_{final}) is the mass (rapidity) of the vector sum of the two leading jets and the two τ four-vectors obtained as discussed in the preceding items.

In this analysis only the first order Optimal Observable is used. This is primarily due to the increased complexity resulting from trying to combine O_1 and O_2 (e.g. through a two-dimensional fit), compared to the relatively small gains obtained by including the second order Optimal Observable. More specifically, the low signal statistics in the signal region implies that a two-dimensional fit requires a coarser binning, which in turn reduces the sensitivity compared to a one-dimensional fit of the first order Optimal Observable with a finer binning. In the future (e.g. Run 2), when higher signal statistics can be expected in the data, this point can be revisited and the gains of including the second order Optimal Observable re-evaluated. Unless explicitly specified, the expression ‘Optimal Observable’ will refer to the first order Optimal Observable, O_1 .

Chapter 2

The ATLAS Experiment

The ATLAS detector (A Toroidal LHC ApparatuS) [83] is one of four main experiments situated around the ring of the Large Hadron Collider [84], located at the European Organisation for Nuclear Research (CERN) in Geneva, Switzerland. This chapter starts by describing the Large Hadron Collider and the sequence of accelerators that feed protons into the machine in section 2.1, followed by a description of variables and concepts necessary to understand the phenomenology of proton-proton collisions in section 2.2. An overview of the constituents of ATLAS can be found in section 2.3, and a summary of the data taking conditions in 2012 is given in section 2.4. The chapter is concluded by section 2.5 containing a description of the author’s service work on muon calorimeter isolation in the ATLAS trigger system.

2.1 The Large Hadron Collider

In 1994 the construction of the world’s currently highest-energy particle accelerator was approved by the council of The European Organization for Nuclear Research (CERN). The machine was named the Large Hadron Collider (LHC) due to its function as a proton-proton collider. It was installed in the existing 27 km circular tunnel at the border between Switzerland and France in which the Large Electron-Positron Collider (LEP) was operational until 2000. The LHC is a high-energy, high-luminosity accelerator designed to reach an energy of 7 TeV per proton per beam, resulting in a center-of-mass collision energy of $\sqrt{s} = 14$ TeV, at an instantaneous luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Much higher energies are achievable when accelerating protons compared to electrons due to the much lower energy loss due to synchrotron radiation, which is proportional to $1/m^4$. Alternatively, lead ions can be accelerated to an energy of 1.38 TeV per nucleon to study heavy ion collisions. The first proton-proton runs at $\sqrt{s} = 900$ GeV used for physics analysis were completed at the end of 2009. Subsequent datasets were collected in 2011 at $\sqrt{s} = 7$ TeV and 2012 at $\sqrt{s} = 8$ TeV, collectively called Run 1. The 2012 dataset is the basis for the work

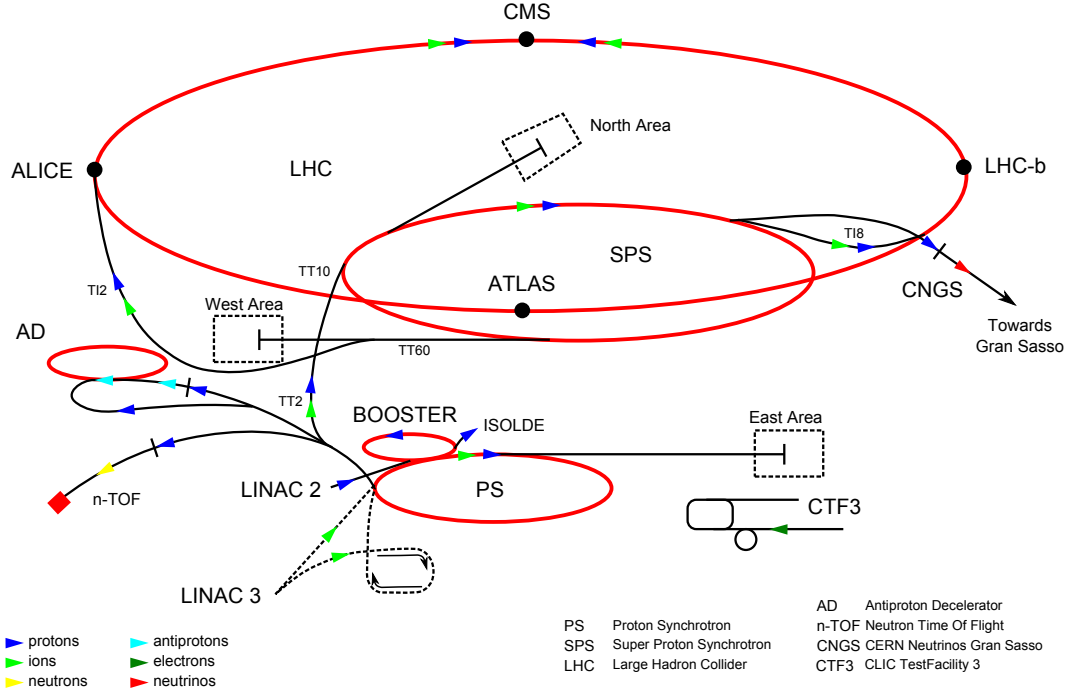


Figure 2.1: Illustration of the the CERN accelerator complex [85] including the various machines used to prepare protons for injection into the LHC.

in this thesis.

The particle beams circulate in opposite directions inside two evacuated beam pipes and are bent by 1232 super-conducting dipole magnets cooled to a temperature of 1.9 K. The dipoles produce field strengths of up to 8.33 T in order to successfully bend the proton beams at the full design energy of 7 TeV per beam. Higher-order magnets in the form of 392 quadrupole magnets are needed to focus the beams, and eight superconducting cavities operating at 400 MHz generate the electric fields needed to accelerate the particles. Once the beams have reached their target energy they will gradually lose luminosity, mainly due to collisions at the interaction points of the four LHC experiments, limiting the useful lifetime of the beams to approximately half a day.

The LHC storage ring is the final destination in a long line of steps needed to produce the protons and accelerate them from rest. These include previous CERN flagships, now repurposed to serve the LHC with high-energy protons and ions. Some of these machines were built in the mid 20th century. The protons are extracted by ionising stored hydrogen and are subsequently injected into the Alvarez Proton Linac, also called Linac 2, a linear accelerator operational since 1978 capable of accelerating the protons to 50 MeV (see figure 2.1) while grouping

them into bunches using radiofrequency (RF) quadrupoles. The protons are already travelling at roughly 30% of the speed of light ($0.3c$). The next stop on the protons' itinerary is the Proton Synchrotron Booster raising the energy to 1.4 GeV ($0.9c$). The oldest of the accelerators, the Proton Synchrotron (PS) from 1959, followed by the Super Proton Synchrotron (SPS) then take over to further raise the energy to 25 GeV ($0.999c$) and 450 GeV ($0.99997c$) respectively. The PS additionally generates the bunch train structure used at the LHC. The design pattern specifies bunch trains consisting of 72 bunches with a temporal spacing of 25 ns followed by empty buckets taking up 320 ns. However, in the 2012 dataset analysed in this thesis a bunch spacing of 50 ns was used. 39 bunch trains are filled into the LHC at the design conditions, with each bunch containing approximately 10^{11} protons. Once the protons have been accelerated by the SPS and having already travelled more than 6 million kilometers in the process, they are ready to be injected into the LHC storage ring. Here they are accelerated to their final energy and collided at the detectors placed around the ring, of which the four main experiments are ALICE, ATLAS, CMS, and LHCb.

The choice of a hadron collider as the next big CERN endeavour should be seen in the context of the state of particle physics in general. When the LHC was commissioned, the main goals of the physics community were to either complete the Standard Model picture by discovering the Higgs boson or perhaps extend knowledge beyond the Standard Model by finding evidence for theories such as supersymmetry, technicolor or extra dimensions. Colliding hadrons is ideal for this purpose due to their composite nature, resulting in a wide range of energies accessible at the same operating parameters of the machine and at the same center-of-mass energy. This makes hadron colliders able to scan a large phase space and a wide interval of particle masses. Hadron colliders are often called 'discovery machines' for this reason. The disadvantage is the chaotic conditions of the collisions generating a large amount of background, which sets high demands on the granularity and time resolution of the detectors and lowers the possible precision of the measurements. Had the basic particles already been discovered, precision measurements would be the preferred choice. This is easier with electron-positron colliders such as CERN's Compact Linear Collider (CLIC) or the proposed International Linear Collider (ILC). Using leptons (elementary particles) provides a very clean environment to perform precision measurements at a narrow energy range. It is however entirely possible that the conditions at a hadron collider could prevent the discovery of certain types of particles, where elementary particle colliders could end up being the first to discover them.

2.2 Proton-Proton Collider Physics

This section describes a number of concepts that are necessary in order to understand and work with the collisions of high-energy protons, from calculations of the beam intensity as a function of time to a mathematical description of the parton interactions inside the protons and the subsequent evolution of the resulting final state particles.

2.2.1 Luminosity

High-energy collision experiments are usually focussed on measuring event rates in a predefined region of phase space. The instantaneous luminosity of a particle beam is defined as the relation between the number of observed events of a given kind of interaction per unit time and the cross section σ for that interaction to occur:

$$N_{\text{evt}} = \mathcal{L}\sigma. \quad (2.1)$$

Given that a particle beam has been circulating for some time interval T with a time-dependent luminosity $\mathcal{L}(t)$, the integrated luminosity during that time interval is then

$$L = \int_0^T \mathcal{L} dt, \quad (2.2)$$

which can be calculated in a collider experiment as

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma}{4\pi \varepsilon_n \beta^*}. \quad (2.3)$$

The variables used in the calculation are the number of particles per bunch N_b , the number of bunches per beam n_b , the frequency of revolution f_{rev} , the relativistic gamma factor γ , the emittance of the beam in the transverse direction ε_n and the beta function at the collision point β^* . The emittance is a constant determined by the initial conditions when the beam was formed and describes how closely packed the protons are in the beam. The beta function varies along the collider ring and is a measure of the focussing strength of the quadrupole magnets in a particular point. β^* can be interpreted as the distance from the interaction point where the beam width is twice as wide as the focus point. This value should be scaled by a geometric factor depending on the crossing angle between the beams.

2.2.2 Parton Distribution Functions

The non-elementary nature of hadrons complicates the theory of their interactions. The individual partons inside the protons in a collider beam can in principle have momentum in all three spatial directions, but the extremely high momentum in the beam direction and the effort to collimate the beam makes the perpendicular momentum components negligible, and they will not be treated in the following. In a collision between two protons it is possible for one or more partons from each proton to interact with each other. The interacting partons can be valence quarks, sea quarks or gluons. Two energy scales are usually identified as the *soft* scale at energies below ~ 200 MeV where the running strong coupling constant is of order 1 or higher, meaning QCD is non-perturbative, and the *hard* scale at considerably higher energies, where the partons can be considered free particles during their interaction. The present analysis is focused on the hard interactions, which are the only kind of interactions involving the required amount of energy to produce an SM Higgs boson. In case the hard process produces a coloured particle, it may emit a gluon as final state radiation (FSR), or one of the interacting partons may emit a gluon before the hard interaction, which is called initial state radiation (ISR). Gluons with sufficient energy can change the topology of the event in the detector by producing jets. These arise in a subsequent process where the gluons undergo parton showering, creating a swarm of quarks and gluons, and finally hadronise into a jet of colour neutral particles. If the partons participating in the interaction are denoted i and j , then the fraction of the hadron momentum P_{hadron} carried by each of them, called the Bjorken x values, can be written as

$$x_{i,j} = \frac{P_{i,j}}{P_{\text{hadron}}}. \quad (2.4)$$

This determines how much momentum/energy is available to the produced particles following the collision. The probability of a collision taking place involving partons with certain x -values is contained in the parton distribution functions (PDFs) $f_i(x)$, where i is the particular parton involved. The PDFs are crucial elements in the theoretical calculations of particle production cross sections. According to the factorisation theorem for deep inelastic scattering [86,87] the total cross section can be calculated from two independent contributions describing the characteristics of interactions over relatively long and short distances respectively, with the PDFs taking the role of the long distance part. The hard process (partonic) cross sections that are calculable from perturbative QCD constitute the short distance part. The

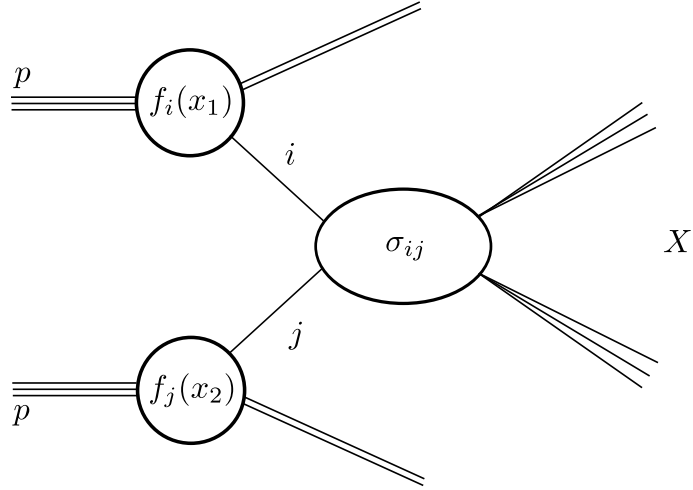


Figure 2.2: Illustration of a parton interaction in a proton-proton collision. The partons i, j have probability densities $f_{i,j}(x)$ to have a fraction x of the proton momentum, and the cross section of the hard interaction producing some final state X is σ_{ij} .

factorised cross section can then be calculated as

$$\sigma_{pp \rightarrow X} = \sum_{ij} \int_0^1 dx_1 \int_0^1 dx_2 f_i(x_1, \mu_f) f_j(x_2, \mu_f) \times \hat{\sigma}_{ij \rightarrow X}(x_1 P_1, x_2 P_2, \mu_f), \quad (2.5)$$

where μ_f is the factorisation scale that describes the proton model and $\hat{\sigma}$ is the partonic cross section that is theoretically calculable from perturbative QCD. According to the value of μ_f the proton can either be approximated by the valence quark model or the more complicated picture where sea quarks and gluons are also prevalent. As the PDFs describe the proton binding mechanism in the non-perturbative QCD energy region, these cannot be calculated from first principle but must be obtained as experimental fits to large sets of data. Several groups are dedicated to obtaining these fits, among which the results of the Coordinated Theoretical-Experimental Project on QCD (CTEQ) [88, 89] are widely used. The main experimental inputs come from data collected in lepton-nucleon deep-inelastic scattering experiments such as HERA that are able to constrain the quark PDFs, as well as processes including vector boson fusion and single-inclusive jet production. Another collaboration supplies the Martin-Stirling-Thorne-Watt parton distribution functions (MSTW) [90]. Major uncertainties in the PDF fits include the experimental uncertainties on the fitted data, uncertainties on the strong coupling parameter α_s and the parametrisation of the fit function. The probability functions are depen-

MSTW 2008 NLO PDFs (68% C.L.)

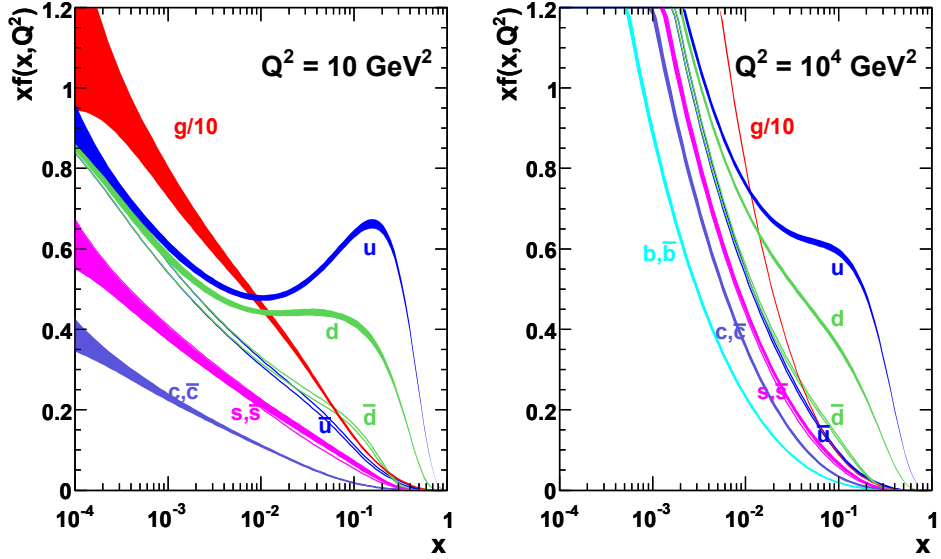


Figure 2.3: The MSTW 2008 Parton Distribution Functions [90] at a momentum transfer scale of $Q^2 = 10 \text{ GeV}^2$ (left) and $Q^2 = 10^4 \text{ GeV}^2$ (right). Both valence quark, sea (anti-)quark and gluon functions are shown.

dent on the energy scale, also called the momentum transfer scale, involved in the interaction. This is the product of the energies of both partons, $Q^2 = x_i x_j s$. Since radiative processes become more prevalent at higher Q^2 the PDFs will be biased towards lower x -values when the energy scale increases. An illustration of a proton-proton collision with a single parton interaction is given in figure 2.2. Figure 2.3 shows the MSTW 2008 PDFs at two values of Q^2 .

2.2.3 Partonic Cross Sections

In the high-energy perturbative limit of QCD the total partonic cross section of an interaction can be expanded in powers of α_s :

$$\hat{\sigma}_{ij}(\mu_f) = \sum_{n=1}^{\infty} \alpha_s^n(\mu_r) \hat{\sigma}_n(\mu_f, \mu_r), \quad (2.6)$$

where μ_r is the renormalisation scale factor on which the running strong coupling constant α is dependent. The SM is a renormalisable theory, which allows a redefinition of the coupling constants that absorbs divergencies that would otherwise lead to infinities, and the final observable cross section does not depend on μ_r . In practice it is only possible to calculate the partonic cross sections up to a certain order,

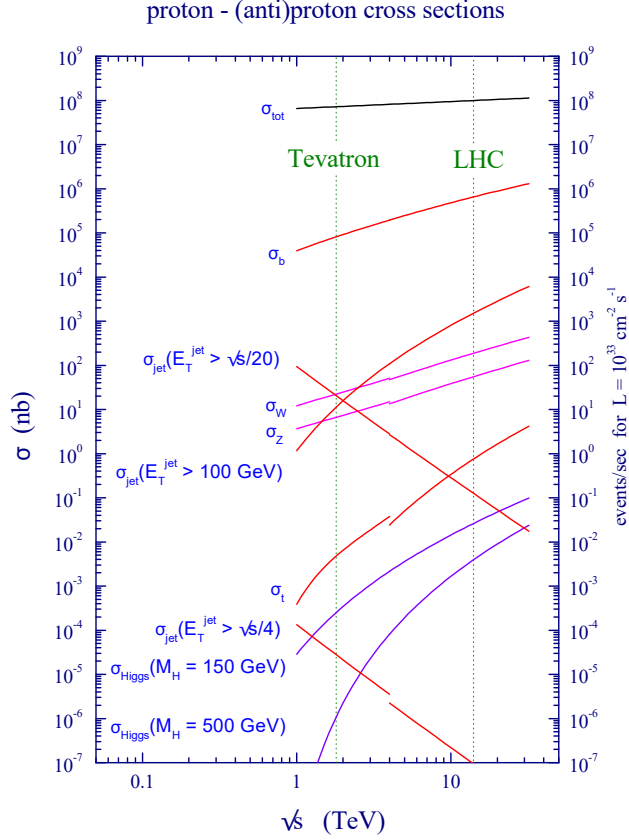


Figure 2.4: Standard Model cross sections as a function of the centre-of-mass energy at the Tevatron (proton-antiproton) for $\sqrt{s} < 4\text{TeV}$ and at the LHC for higher energies [87].

leaving a residual dependency in the result. Uncertainties due to the missing higher orders in the calculation are therefore usually estimated by varying both μ_r and μ_f . At each order in α_s the cross sections can be calculated through the Feynman rules given by the theory. The leading order calculation is followed by higher orders describing virtual (loop) corrections as well as real emissions. Figure 2.4 shows cross sections for the production of common particles at the LHC as a function of the centre-of-mass energy.

It is usually necessary to perform higher order calculations using numerical methods due to their complexity. In order to simulate event kinematics, differential distributions as functions of the phase space variables are needed. Dedicated event simulation programs called Monte Carlo generators [91] are employed to perform the numerical integration and generate particles according to the resulting kinematic distributions. As the LHC collides protons containing strongly interacting particles,

events with multiple jets in the event, also called QCD multijet events, are common and greatly contribute to the background in many physics analyses.

2.3 The ATLAS Detector

ATLAS is a general purpose detector designed to have excellent resolution on all final state objects needed to reconstruct high-energy collisions. Its design produces the required sensitivity to search for the SM Higgs boson within the first few years of running and to search for new particles beyond the SM at the TeV scale. These requirements necessarily make ATLAS a complex combination of many subdetectors, each with its own specialised purpose. The operating parameters of the LHC also impose constraints on the detector parts. ATLAS is a forward-backward symmetric detector and contains a solenoid magnet with a strong magnetic field of 2 T required to bend high- p_T charged particles and enable momentum and charge measurements. Three toroid magnets provide a separate momentum measurement for muons. The high nominal luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ demands a high time resolution of the various detectors as well as fine-grained tracking due to high particle multiplicities. As the physically largest of the four main detectors at the LHC, ATLAS is 44 m long and 25 m high while weighing 7,000 tonnes. It is installed at Point 1 in the

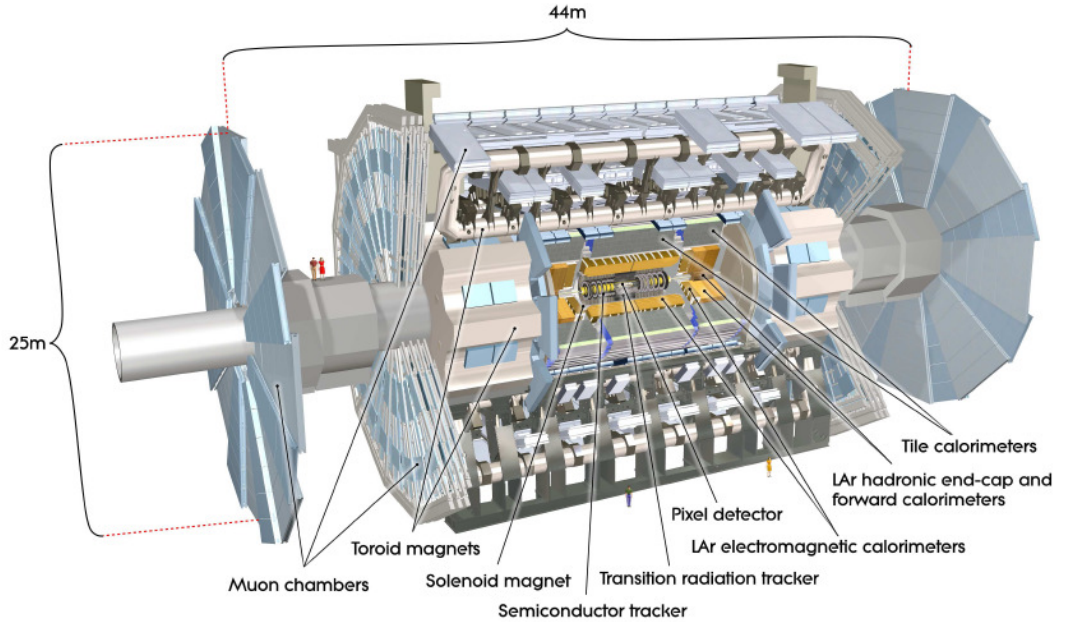


Figure 2.5: Full view of the ATLAS detector illustrating the placement of tracking systems, calorimetry, muon spectrometer and magnets.

LHC ring, situated very close to CERN's Meyrin site. Figure 2.5 gives a schematic overview of the ATLAS detector and its subcomponents, which will be discussed in the following sections.

2.3.1 Definition of Coordinates and Variables

The basis for all spatial calculations in ATLAS is a right-handed cartesian coordinate system with its origin placed at the nominal interaction point at the center of the detector. The z -axis is oriented in the direction of the beam pipe, while the positive x -axis is defined to point towards the center of the LHC ring, and the positive y -axis points upwards. Thus the xy -plane is orthogonal to the direction of the particle beams. Alternatively, a cylindrical coordinate system can be used, which is more convenient when describing rotationally symmetric properties. This defines the radial distance from the beam pipe in the xy -plane as $r = \sqrt{x^2 + y^2}$ and the azimuthal angle in the xy -plane as $\phi = \arctan(y/x)$, while the polar angle from the beam direction is $\theta = \arctan(r/z)$.

A widely used variable in particle physics is the rapidity $y = \frac{1}{2} \ln[(E + p_z)/(E - p_z)]$. It relates different relativistic frames of reference by describing the hyperbolic angle associated with the Lorentz boost necessary to transition from one frame to the other. Its additive nature under such Lorentz boosts makes it useful compared to the complex rules of velocity addition. The polar angle θ can be substituted with the so-called pseudorapidity η defined as the low mass limit of rapidity: $\eta = \lim_{m \rightarrow 0} y$. This quantity is useful for highly relativistic particles where $\eta \sim y$ since the kinetic energy is dominant. The pseudorapidity can be written as

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \quad (2.7)$$

The initial momentum of the interacting partons along the z -direction is not known due to the nature of the collisions described in section 2.2.2. Only the transverse component is known. A very good approximation is that the initial transverse momentum is zero. This is not an exact statement since transverse fluctuations can also occur, but at such low scales that the above approximation is possible. The natural way to express quantities such as momentum and energy is thus by using their transverse component:

$$p_T = \sqrt{p_x^2 + p_y^2} = p \sin(\theta), \quad (2.8)$$

$$E_T = \sqrt{m^2 + p_T^2 \cosh \eta}. \quad (2.9)$$

Another important concept is that of missing transverse energy, E_T^{miss} . It refers to energy that is not detected but is necessary to uphold energy and momentum conservation in a collision. Particles such as neutrinos are not directly detected due to their feeble interaction strength, so the only practical way to infer their existence is by applying conservation laws and calculating the direction and amount of missing energy. The missing transverse energy is defined as

$$E_T^{\text{miss}} = - \sum_i (p_T)_i. \quad (2.10)$$

In isolation studies and other analyses where it is important to know how close the produced particles are to each other when traversing the detector, the quantity ΔR is defined to describe the angle between the trajectories of any two particles. It is defined in the η - ϕ plane as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (2.11)$$

This definition takes into account the fact that on average the distribution of particles is uniform in both ϕ and η (for $|\eta|$ below a certain threshold depending on the collision energy, defining a ‘rapidity plateau’). The detector granularity also reflects this.

2.3.2 The Inner Detector

The innermost part of ATLAS, the Inner Detector (ID), is designed to provide high precision tracking of electrically charged particles for $|\eta| < 2.5$ and enables the reconstruction of particle momenta, primary collision vertices as well as secondary vertices of decaying particles. It is designed to have a relative transverse momentum resolution of $\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$ [92], where the term proportional to p_T is due to the spatial resolution of the detector and the constant term is due to multiple scattering. This is achieved by a combination of three distinct detection systems, each utilising different technologies. Placed closest to the beam pipe are two silicon based subsystems named the Pixel and SCT detectors providing pixel tracking and silicon strip detection. A detector based on straw tubes (TRT) is positioned as the third system, which provides both tracking and electron identification based on transition radiation detection at $|\eta| < 1.9$. Figure 2.6 shows the layout of the ID and its three tracking subdetectors.

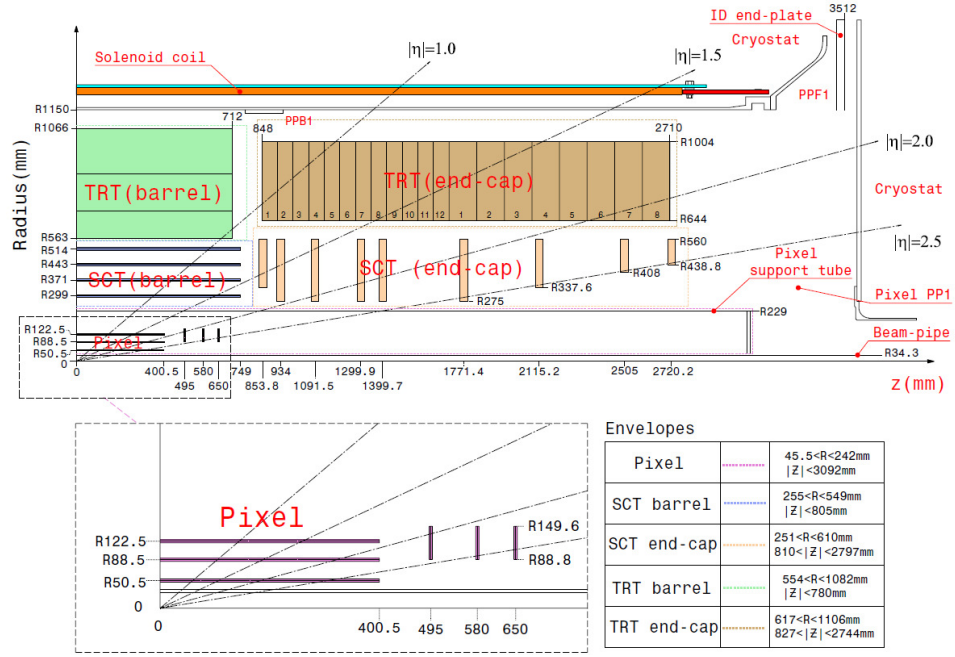


Figure 2.6: An illustration of one half of the Inner Detector showing the relative placement of the layers in the Pixel, SCT and TRT with respect to the beam pipe (radius) and central collision point (z -axis).

As a tracking system, the Inner Detector is designed to minimise the interaction of the produced particles with the detector services and support structure to let particles reach the outer detector parts while losing as little energy as possible. The material budget is a trade-off between the desired low interaction rate and the need for a stable mechanical structure, readout electronics, cooling etc. The amount of material that particles must traverse is given in units of interaction lengths, which is the mean distance traversed by a high energy electron through the material before it only possesses $1/e$ of its initial energy. The ID material budget is displayed graphically in figure 2.7.

The Pixel Detector

The detector system positioned closest to the beam pipe is the Pixel detector. It consists of 1744 pixel modules, each having an active detection area of $63 \times 19 \text{ mm}^2$, a thickness of $250 \text{ }\mu\text{m}$ and each containing 46080 pixels. The modules are arranged in three layers parallel to the beam pipe at $0 < |\eta| < 1.7$ (barrel) complemented by three vertical disks on each side covering the interval $1.7 < |\eta| < 2.5$ (endcap). All three layers have identical tracking precision with each module having a spatial resolution of $10(R - \phi) 115(z) \text{ }\mu\text{m}$.

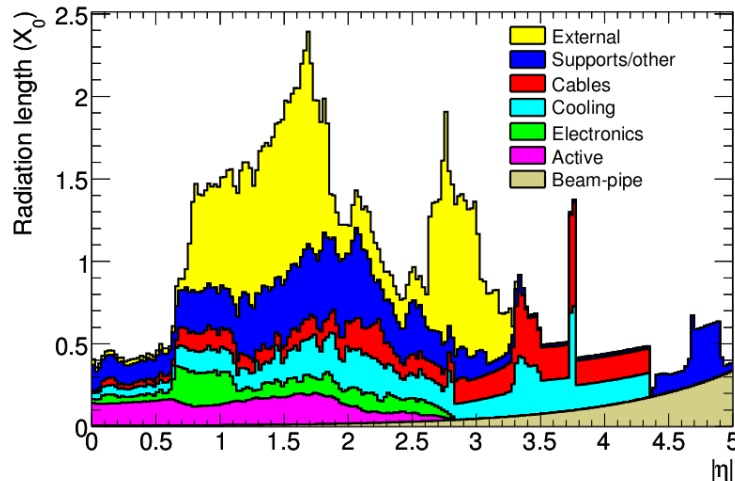


Figure 2.7: *The material budget for various parts of the Inner Detector given in radiation lengths (X_0) as a function of $|\eta|$. The distribution has been averaged over ϕ . The central region (low $|\eta|$) is most sensitive to new physics, and the material budget is accordingly very low in this region.*

With its first layer positioned at a radius of just 50.5 mm from the beam pipe the Pixel detector is necessarily designed using highly radiation resilient materials. The proximity to the proton interactions makes the Pixel detector essential for reconstructing secondary vertices of short-lived particles, which is used extensively in e.g. b-physics.

For each triggered collision event a list of hits is provided by the Pixel detector. The discrimination between noise and signal in each pixel cell is done using a voltage threshold. The initial analogue charge signal first needs to be amplified and is then compared to the discrimination threshold. The signal is subsequently digitised to contain the spatial coordinates of the hits as well as the Time over Threshold (ToT). Since the pixel cells have been calibrated to have a linear ToT response versus charge deposition, the energy loss dE/dx of particles traversing the cell can be obtained. The response is tuned such that minimum ionising particles (MIPs) correspond to a specific ToT.

The Semi-Conductor Tracker

The SCT resembles the Pixel detector in function and layout with barrel and endcap sections but relies on narrow silicon strips instead of pixels. A module contains two sensors, each containing 768 active strips with a thickness of $285\ \mu\text{m}$ that are mounted back-to-back on the modules at a slight angle of 40 mrad. This enables

the SCT to measure position also in the z -direction and thereby provide three-dimensional space points with an accuracy of $17\ \mu\text{m}$ in the R - ϕ plane and $580\ \mu\text{m}$ along the z -axis. The total number of 4088 modules are placed in four coaxial layers in the barrel region and 9 disk layers in each of the two endcaps. The modules cover an area of $63\ \text{m}^2$ which gives almost hermetic coverage with potential for a least four space-point measurements at all pseudorapidities $|\eta| < 2.5$.

The Transition Radiation Tracker

The Transition Radiation Tracker (TRT) exploits the phenomenon of transition radiation emitted by charged particles when they pass a boundary between two dielectric materials. The intensity of the emitted transition radiation is proportional to the γ -factor of the particle, which enables good discrimination between particles such as electrons and pions due to their mass difference. The TRT is constructed from proportional chambers consisting of straws within a radiator material in such a way that the probability of transition radiation by relativistic particles is maximised. The emitted photons have a typical energy of a few keV, corresponding to X-rays. The TRT thus has finely developed particle identification (PID) capabilities for the purposes of discriminating e.g. pions from electrons. The straw resolution is $140\ \mu\text{m}$ (R - ϕ), and the total number of hits per track averages at 35. The readout is performed using two threshold values with the low threshold (LT) tuned for minimum ionising particles at 250 eV, while the high threshold (HT) triggers at $\sim 6\ \text{keV}$ for identification of transition radiation.

The straws are made of mylar coated kapton and have a radius of 2 mm. The cathode consists of a $0.2\ \mu\text{m}$ aluminium layer in the form of a coating on the inside of the straws, while the central anode is a $31\ \mu\text{m}$ wide tungsten wire with a gold coating. The medium in the straws consists of 70% Xe , 27% CO_2 and 3% O_2 . The TRT amounts to a total number of 298,304 straws arranged in two endcaps, each with 122880 straws, and a barrel region with 52544 straws, covering pseudorapidities up to $|\eta| < 2$.

2.3.3 Calorimeters

Outside the Inner Detector the particles reach electromagnetic and a hadronic calorimeter systems designed to measure particle energies through absorption. The electromagnetic calorimeter is designed to fully absorb electrons and photons and measure the resulting cascades of electromagnetic bremsstrahlung and e^+e^- pair production. The goal of the hadronic calorimeter is to contain and measure hadronic

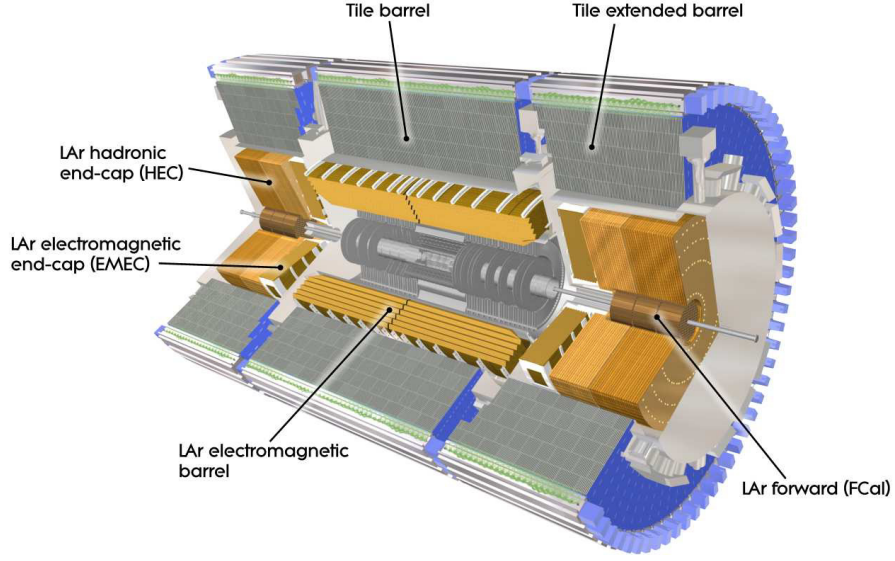


Figure 2.8: *Cut-away view of the ATLAS calorimeter system [83].*

showers of high-energy hadrons. The combined calorimeter system covers pseudo-rapidities up to $|\eta| < 4.9$. It is of a sampling design with alternating layers of passive material that initiates showering and active scintillating material that is able to detect the resulting particles. The calorimeters are divided into segments in both the longitudinal and azimuthal direction, which provides spatial resolution and the ability to reconstruct the three-dimensional position of a particle shower. This can be used in photon reconstruction and detection of missing transverse energy from weakly interacting particles such as neutrinos that escape the calorimeters. Contrary to the tracking purposes of the ID, the calorimeters are designed to absorb the full energy of the particles. The combined depth should therefore correspond to a large number of radiation lengths X_0 , the mean length after which the energy of an electron is reduced by a factor of $1/e$ from bremsstrahlung. The placement of the calorimeter systems is detailed in figure 2.8, and the pseudo-rapidity coverage of each calorimeter system can be seen in figure 2.9.

The Electromagnetic Calorimeter

The electromagnetic calorimeter system is built as a liquid argon (LAr) sampling calorimeter with lead absorber plates. It is constructed with a hermetic and fine-grained accordion geometry with a high energy precision of $\sigma_E/E = 10\%/\sqrt{E/\text{GeV}} \oplus 0.7\%$. The lead plates have a thickness of 1.13 to 1.53 mm depending on the module. The electromagnetic barrel (EMB) calorimeter covers the region $0 < |\eta| < 1.475$ and

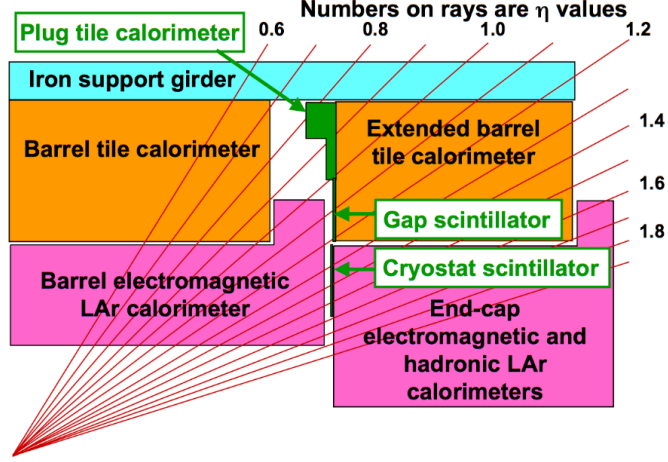


Figure 2.9: Overview of the Tile and LAr calorimeters and their η coverage.

consists of 4 layers. The electromagnetic endcap (EMEC) calorimeter also consists of 4 layers and covers the region $1.375 < |\eta| < 3.2$.

The layered design provides additional information about the radial development of particle showers. The first layer is called the strip layer and has a granularity of $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$ in the barrel. The second and thickest layer has a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$ and covers 16 radiation lengths X_0 . The third layer helps measure the tails of electromagnetic showers that can help distinguish between electromagnetic and hadronic processes due to the longer interaction length of the latter. A single 11 mm thin LAr layer called the presampler is placed in front of the main EM calorimeter in the region $0 < |\eta| < 1.8$ to provide an additional measurement before the first sampling layer to help correct for energy losses occurring in the ID and the support structure.

The Hadronic Calorimeter

The hadronic calorimeter is based on different technologies depending on the pseudorapidity and generally has a lower granularity than the EM calorimeter. The resolution of the hadronic barrel and endcap calorimeters is typically $\sigma_E/E = 50\%/\sqrt{E/\text{GeV}} \oplus 3\%$, rising to $\sigma_E/E = 100\%/\sqrt{E/\text{GeV}} \oplus 10\%$ in the forward calorimeter. In the barrel region it is based on active plastic scintillating material interleaved with steel tiles as absorbant, earning it the name ‘tile calorimeter’. It covers radii from 2280 mm to 4230 mm while the η coverage extends to $|\eta| \lesssim 1.7$. The calorimeter is subdivided into a central barrel covering $|\eta| \lesssim 1.0$ and an extended barrel covering $0.8 \lesssim |\eta| \lesssim 1.7$. Both barrel parts are divided into 64

modules spaced evenly in azimuthal angle ϕ . The cell granularity in the barrel region of the tile calorimeter is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the first two layers, while the third layer has half this granularity. The radial depth of the tile calorimeter is approximately 7.4 interaction lengths resulting in a very low probability of hadronic particles punching through to the muon system.

In the endcap region a LAr-based system has been chosen for its radiation hardness and linearity as well as the possibility of replacing the argon over time as it gets ionised. The two disks cover the region $1.5 < |\eta| < 3.2$ and are built using interleaving sections of copper and liquid argon. For $|\eta| < 2.5$ the calorimeter cell granularity is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. At higher pseudo-rapidities the granularity is twice as coarse. An additional forward calorimeter, also LAr-based, covers the far forward region $3.1 < |\eta| < 4.9$ and acts as both an electromagnetic and hadronic calorimeter. Of the three layers, the first uses copper as absorber material and is designed for electromagnetic interactions, while the two following layers are based on tungsten to absorb hadronic showers.

2.3.4 The Muon Spectrometer

The Muon Spectrometer (MS) is the outermost and hence the last part of ATLAS that particles can potentially meet on their trajectory. It is designed to detect charged particles that are able to penetrate the calorimeters, and to measure their momentum for $|\eta| < 2.7$. The momentum measurement is enabled by large toroid magnets. Depending on the radius r and azimuthal angle ϕ the field strength has values from 0.15 T to 2.5 T. To give a sense of the momentum resolution achievable in the MS, a track with three MDT hits has a relative uncertainty of

$$\frac{\sigma_p}{p} = \frac{\Delta S \times p}{500 \text{ } \mu\text{m}}, \quad (2.12)$$

with an assumed sagitta resolution ΔS of 45 μm . This corresponds to a resolution of $\sim 10\%$ for a 1 TeV track. A wide range of detector technologies are used in the MS including monitored drift tubes (MDT), multi-wire proportional chambers (cathode strip chambers, CSC), resistive plate chambers (RPC) and thin gap chambers (TGC), where the last two technologies are used for triggering. For an illustration see figure 2.10.

The trajectories of muons are measured by MDTs, CSCs and TGCs with varying response times and resolutions. The MDTs are drift chambers with aluminium tubes filled with a gas mixture containing argon and CO_2 at a pressure of about 3 bar. Ionised electrons are collected by tungstenrhenium wires at a voltage

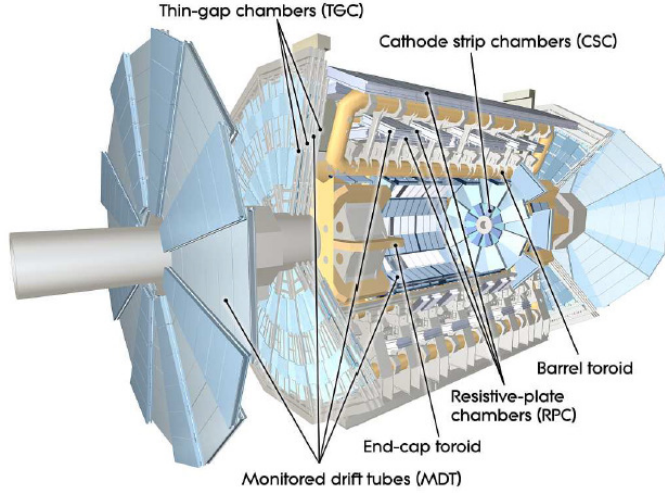


Figure 2.10: *The positions of the various parts of the ATLAS muon system.*

of 3080 V. The chamber sizes increase in direct relation to their distance from the interaction point. The MDT chambers cover the region $|\eta| < 2.7$, except for the innermost layer where they are replaced by CSC chambers for $2 < |\eta| < 2.7$. In the barrel region they are arranged in three layers at radii between 5 and 10 m from the beam pipe, while the endcap has four wheels placed at $|z| = 7.4$ m, 10.8 m, 14 m and 21.5 m. The large restitution time inherent to the MDT technology of close to a μs makes it unfit for use at high pseudorapidities with high particle occupancy, but in the central detector region it provides very precise spatial resolution.

The CSC technology is used in the forward areas $2.0 < |\eta| < 2.7$ where the MDTs can no longer cope with the high particle flux. The CSCs measure both track coordinates simultaneously through charge collected on two planes of orthogonal strips. Tracks are measured in four CSC planes, which is less than in the MDTs, but the faster response time helps regain tracking precision. The total area covered by CSC chambers is $\sim 65\text{m}^2$.

Two complementary triggering systems with high time resolution are in place in the MS. RPCs are in place in the range $|\eta| < 2.4$. These consist of two parallel electrode plates with a gap of 2 mm filled with a gas mixture. The plates have an electric field of 4.9 kV/mm and are orthogonally segmented to provide two-dimensional hits. The chambers are mounted in front of and on the back of the middle MDT layer, as well as on the back of the outer MDT layer. The two innermost layers of the end cap wheels are also equipped with TGCs that provide triggering information and an additional hit for the muon momentum measurement. The time resolution of the triggering technologies is on the order of nanoseconds to successfully

identify the bunch crossing that each track belongs to.

2.3.5 The ATLAS trigger system

Given the extremely high luminosity at which the LHC is designed to run, it is impossible to record every single event. At the 50 ns bunch spacing used in 2012, bunch crossings occur at a rate of $2 \cdot 1/(50 \cdot 10^{-9}\text{s}) = 40$ MHz. The collision rate is several times higher if one includes pile-up. This should be compared to the capabilities of the storage system, which could only record approximately 400 events per second in the same time period. It is therefore necessary to rely on complex algorithms to quickly select only the events that are interesting for physics analysis. For this purpose the ATLAS trigger system has been developed as custom-built hardware placed locally at the detector followed by a data farm based on commercially available computers.

The trigger system used in 2012 is designed to run in three stages/levels¹. The hardware based Level 1 (L1) trigger starts by selecting signatures from high- p_T particles in the muon spectrometer or calorimeters at reduced granularity. Within a decision time of $2.5 \mu\text{s}$ it accepts or rejects events made up of simplified detector information. Energy deposits in the calorimeters are searched for maxima using sliding window algorithms defining so-called trigger towers, and for each accepted event the L1 trigger defines a region of interest (RoI) containing the identified trigger object. L1 electron and photon objects are defined as windows of 2×2 trigger towers in the EM calorimeters, while trigger objects for the identification of hadronic tau decays consider towers from both EM and hadronic calorimeters. The process is illustrated in figure 2.11 (a). The L1 muon trigger makes use of coincidence requirements on the trigger chambers. When a hit is found in the first layer, it defines a geometrical coincidence window inside which additional hits are searched for in the outer layers. The centre of the window is placed on the expected impact point of a straight (infinite momentum) track with its origin at the interaction point of the event. The momentum threshold of the specific trigger defines the size of the window in order to include tracks with the corresponding amount of bending in the magnetic field. Figure 2.11 (b) gives a visual representation of this procedure. An estimate of the missing transverse energy is also calculated at L1 using the scalar and vectorial sum of the energy deposits. The central trigger processor (CTP) then takes decisions based on the reconstructed information using logical expressions stored in look-up tables. The L1 trigger is designed to reduce the output rate to the order of

¹The trigger system has since evolved into a two-stage system, unifying the L2/EF software levels.

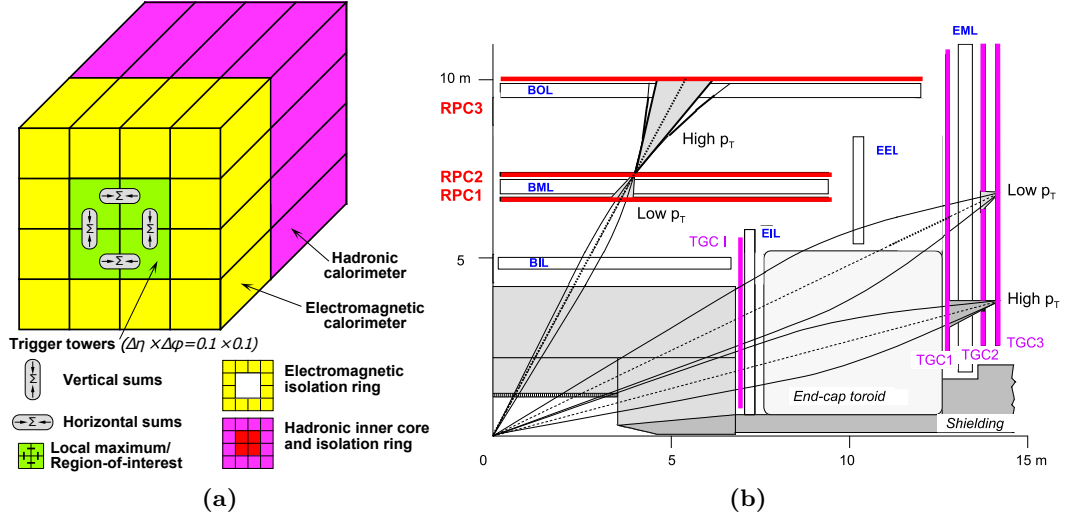


Figure 2.11: Illustration of (a) the L1 electron, photon and hadronic tau trigger algorithms and (b) the L1 muon trigger based on coincidences in momentum threshold dependent windows [83].

75 kHz.

Events that pass the L1 trigger have their detector information within the L1 RoI passed on to the Level 2 (L2) software trigger. Depending on the identified objects in the RoI, different dedicated L2 chains are called. The L2 analyses the data at full granularity and precision within the RoI, which corresponds to roughly 2% of the total event data. It has access to more detailed reconstruction algorithms including information about ID tracks. This reduces the rate output to approximately 3.5 kHz. The final stage of the event selection is carried out by the event filter (EF), which performs a full event reconstruction at full detector granularity. The EF algorithms are closely related to those used offline, requiring a processing time of about 4 seconds. This step again reduces the event rate to about 400 Hz, double the design output rate of 200 Hz due to the availability of extra computing and storage resources [93]. Events passing the complete trigger system are then stored permanently and reconstructed offline.

2.4 Data Taking with ATLAS in 2012

The LHC completed its first run period, called Run 1, in 2013. A gradual increase in the injection energy brought the centre-of-mass energy from 900 GeV starting in November 2009 up to 8 TeV in 2012. In parallel, the luminosity was increased such that the 2012 dataset comprised 21.3 fb^{-1} of recorded data compared to just 45

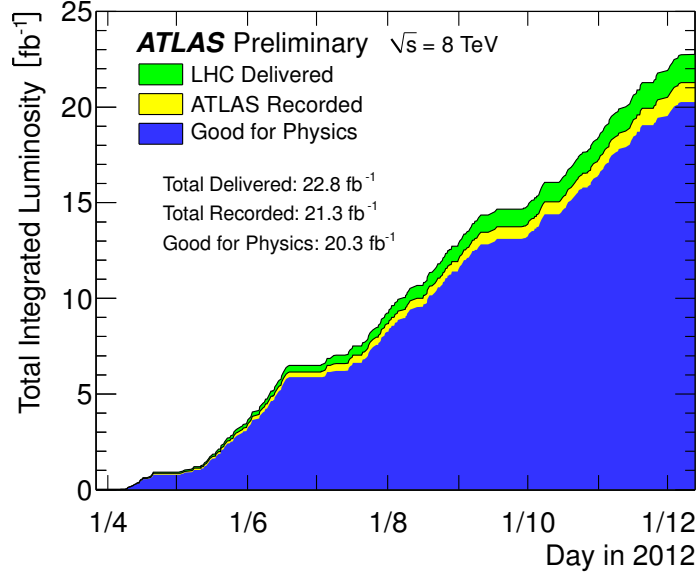


Figure 2.12: Cumulative integrated luminosity recorded by ATLAS (yellow) as a function of date in 2012, compared to the amount of luminosity delivered by the LHC (green) and the fraction of this luminosity that passed basic quality requirements and was deemed ‘good for physics’ (blue) [94].

pb^{-1} in 2010 [94]. The data taking efficiency during the entire run was above 90%. A further small reduction in the size of the dataset used in actual physics analyses is caused by periods with detector modules being offline or otherwise performing suboptimally. Figure 2.12 shows the integrated luminosity as a function of time as well as the efficiency of the ATLAS data collection and the fraction of this data passing basic quality criteria in 2012.

As the luminosity increased, so did the number of proton-proton collisions per bunch crossing, from roughly 3–5 at the start of 2011 to a maximum of about 35 in 2012. The phenomenon on multiple interactions within one bunch crossing is usually termed *pile-up*. The contributions from pile-up to the event reconstruction can be divided into *in-time* pile-up if the additional interactions happened within the same bunch crossing, and *out-of-time* pile-up if interactions from an earlier bunch crossing have an effect on the reconstruction of the current event, which can happen if the bunch spacing is smaller than the time it takes for the detector readouts to return to their neutral state after receiving a signal. Figure 2.13 shows the mean number of interactions per bunch crossing in the 2012 dataset.

The operation of the LHC is organised in *runs* with a duration of the order of a day or less. Hence, the ATLAS data taking follows this run structure by defining

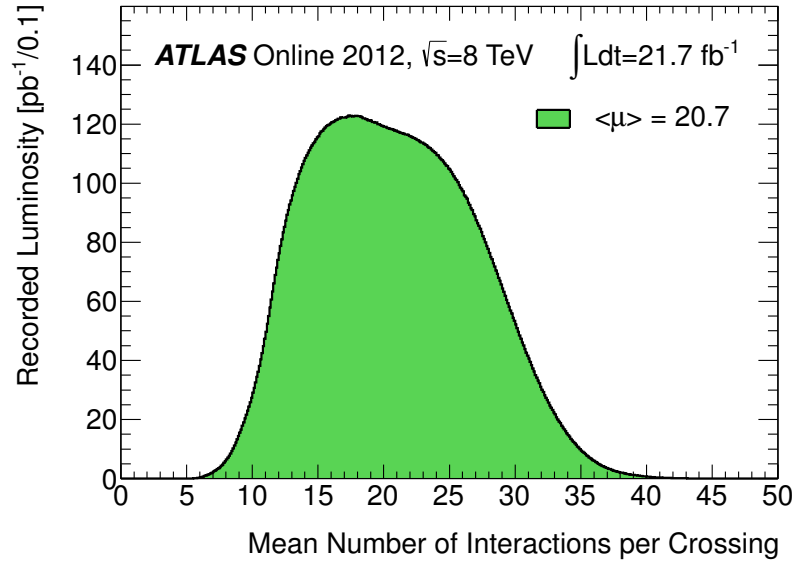


Figure 2.13: The mean number of interactions per bunch crossing at $\sqrt{s} = 8$ TeV in 2012 [94].

periods where the data acquisition system is active. A run is further subdivided into *luminosity blocks*, which define the smallest unit of data taking in ATLAS. The aim is for each luminosity block to contain the same amount of data, making their duration luminosity dependent. Data quality is monitored during data taking in all detector subsystems by both automated systems and personnel to ensure that the detector performance in each run is suitable for physics analysis. Smaller shortcomings in a run such as individual defective calorimeter cells can be flagged before reconstruction, while more severe errors such as entire subsystems being offline are catalogued in a database recording the data quality in each luminosity block. Analysers may then use this database to create Good Run Lists (GRLs) containing the runs and luminosity blocks with good detector performance in the subsystems relevant for their analysis.

2.5 Muon Calorimeter Isolation

The calorimeter isolation energy of a muon is defined as the transverse energy within a ΔR cone centred on the muon track, typically of size 0.2. Extra activity in the event will deposit energy that does not come from the muon, and the amount of energy depends on the process from which the muon originates. For example, muons from W and Z decays are typically fairly isolated with little extra activity in close

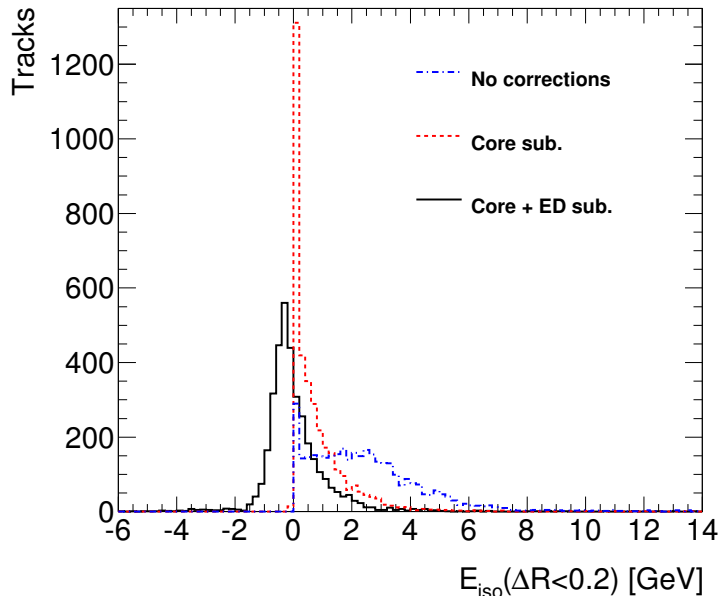


Figure 2.14: Muon calorimeter isolation energy within a cone of width $\Delta R < 0.2$ for simulated muons from $Z \rightarrow \mu\mu$ events at $\sqrt{s} = 13$ TeV. Distributions are shown for the uncorrected energy (blue), the result of subtracting the core energy within $\Delta R < 0.1$ (red) and the final correction subtracting both the core energy and the ambient energy in the detector (black).

proximity to the muon track, while muons from b - and c -jets or from kaon and pion decays will contain larger amounts of energy within the muon isolation cone. The isolation energy can therefore help discriminate between different decay types. The isolation variable should ideally contain only the energy stemming from the remaining decay products other than the muon, so the muon self-energy within a smaller core cone as well as energy stemming from pile-up and the underlying event need to be subtracted to the extent possible.

Online calorimetric muon isolation in Run 1 was performed at the L2 trigger level and was based on simple addition of calorimeter deposits within conical sections. This proved to have some separation power, but is insufficient in the high pile-up environment of Run 2 and beyond, where the efficiency of the algorithm is severely reduced. To ensure compatibility with the new high level trigger design for Run 2, the algorithm was rewritten to run in the combined L2+EF high level trigger (HLT), and more advanced methods of calorimeter energy measurements were employed to increase accuracy and reduce the impact of pile-up. In the intervening time between the initial development of the L2 algorithm and the start of Run 2 several tools and algorithms that were previously only accessible offline had been

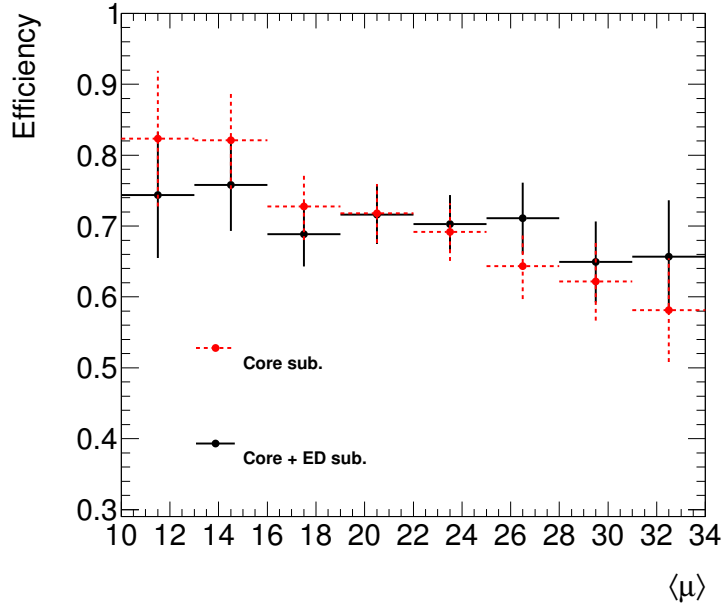


Figure 2.15: Efficiency of the muon calorimeter isolation trigger algorithm in simulated $Z \rightarrow \mu\mu$ events as a function of the mean number of interactions per bunch crossing $\langle\mu\rangle$, plotted at a total signal efficiency of 70%. Efficiencies are shown for the core energy subtraction including (black) and excluding (red) the energy density subtraction.

made available for online use. Among these is a common isolation tool interface that can access pre-calculated isolation variables based on topological clusters (see section 3.2) and with the ability to correct for the mean ambient energy density in the central and forward calorimeter regions separately (central region: $|\eta| < 1.5$, forward region: $|\eta| > 1.5$ and $|\eta| < 3.0$) as an estimate of the energy contribution stemming from the underlying event and pile-up. Figure 2.14 shows the deposited energy estimated using topological clusters within a cone of $\Delta R < 0.2$ around muons in simulated $Z \rightarrow \mu\mu$ events. Collisions were simulated at $\sqrt{s} = 13$ TeV with a range of pile-up conditions corresponding to an average number of interactions per bunch crossing of roughly 10–40. Energy distributions are shown for the total energy within the isolation cone and after the subtraction of the muon core cone energy and the ambient energy. Both absolute (E_T) and relative (E_T/p_T^{muon}) isolation variables are used in isolation studies. Since the muon self-energy is always a fraction of the full energy in the cone, the resulting isolation variable after the core cone subtraction will always be positive or zero. The ambient energy estimate is independent of the isolation energy and therefore has the potential to result in a negative final isolation energy.

The trigger efficiency as a function of pile-up is shown in figure 2.15 using relative isolation including the corrections subtracting the muon self-energy alone and in combination with the ambient energy subtraction. The isolation threshold was chosen separately in the two cases such that the total algorithm efficiency was 70%. A stronger dependence on pile-up is observed when the ambient energy is not subtracted, indicating that this correction could help increase the trigger efficiency and robustness in future ATLAS runs with high pile-up conditions.

Chapter 3

Event Reconstruction

Following the online trigger event selection, for both real and simulated data, events are stored as raw readout data from the detector. They are then run through chains of algorithms in an offline environment designed to reconstruct particles, spatially connected energy deposits and other event kinematics that the ATLAS detector is sensitive to, including charged particle trajectories (tracks), electrons, muons, tau leptons, photons, collimated jets of particles and missing transverse energy. Analyses investigating the $H \rightarrow \tau\tau$ decay are slightly unusual in the sense that they require the reconstruction of every object type except photons. Algorithms at the start of a chain usually consist of routines that reconstruct tracks with associated four-momenta from individual hits in the tracking detectors or build clusters of calorimeter cell energies from topologically connected energy deposits. Variables can also be constructed to aid in the discrimination between different types of particles. This higher-level information can then be exploited in the subsequent particle reconstruction and identification algorithms. As is the case with the high-level software-based online trigger levels, the offline event reconstruction algorithms are implemented and organised using the ATHENA [95] framework.

This chapter will describe the relevant object reconstruction in $H \rightarrow \tau\tau$ events and its performance in both the 2012 ATLAS data and the corresponding simulated samples. In cases where details differ between $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$, the emphasis will be on $\tau_\ell\tau_{\text{had}}$. Tracks and vertices are introduced in section 3.1. Together with clusters of calorimeter energy deposits described in section 3.2 these form the building blocks used to reconstruct physics objects such as light leptons, jets and hadronically decaying taus that are treated in sections 3.3 to 3.6. Although neutrinos cannot be directly reconstructed in the detector due to their negligible interaction cross section, their presence in an event can be inferred through the identification of missing transverse energy as described in section 3.7. These are not meant as exhaustive descriptions of the algorithms, but rather as a summary of the reconstruction and how the objects are used in the analysis.

3.1 Tracking and Vertexing

Particles with an electric charge can leave signals, also called hits, in the tracking detectors. Algorithms [96] are employed to combine the spatial coordinates of these hits into tracks consistent with particle trajectories and estimate their momenta. The default reconstruction of tracks with $p_T > 400$ MeV is an inside-out procedure starting from seed tracks in the pixel and SCT detector layers, extrapolating the resulting tracks to the TRT according to the magnetic field strength and detector material present in the ID. Compatible track segments in the TRT are then combined with these to form improved trajectory and momentum estimates. In the case of secondary tracks originating from particle decays at larger radial distances, an outside-in procedure can be used to produce TRT seed tracks that are extrapolated towards the interaction vertex and combined with track segments in the pixel and SCT detectors. The precision with which a track is reconstructed depends on the number and placement of hits in the various subdetectors, and quality criteria are imposed to reject tracks with poor reconstruction. These can include a minimum number of hits in the silicon detectors as well as the requirement of a hit in every layer of the pixel detector. Track quality criteria can differ slightly between reconstruction algorithms for different types of particles (electrons, muons, hadronic taus etc.) and are defined by their respective ATLAS performance groups according to the detector conditions in a specific data taking period.

By comparing the z -positions of reconstructed tracks at the beamline they can also be used to identify primary interaction vertices. The z -position of a track is used as a vertex seed in an iterative χ^2 fit to nearby tracks [97]. The tracks have associated weights depending on the χ^2 compatibility. Tracks positioned more than 7 standard deviations away from a previously found vertex will seed a new vertex, and this procedure is repeated until all vertices have been found. Events will generally have more than one reconstructed vertex. The primary vertex of an event is defined as the vertex having the largest sum of squared transverse momenta of its associated tracks, i.e. $\sum p_T^2$, with the remaining vertices counting as pile-up. Secondary decay vertices can also be found in the case where the primary particle has a sufficiently long lifetime to enter the pixel detector before decaying. The compatibility of a track stemming from the primary vertex of an event can be estimated by computing the transverse and longitudinal components d_0 and z_0 of the track's impact parameter. The impact parameter is defined as the distance between the point of closest approach of a track and the primary vertex. This is useful for example in b -tagging due to the relatively long lifetimes of hadrons

containing b -quarks leading to decays inside the silicon detectors, which is exploited in the identification of tau leptons.

3.2 Topological Clusters

Topological clustering is a method of grouping spatially connected energy deposits in calorimeter cells into three-dimensional clusters originating from the same particle shower while suppressing calorimeter noise. The calorimeter cells used for clustering are calibrated to the electromagnetic scale (EM scale) [98] defined by calibration constants obtained from test beam measurements. Topological clusters are the primary input objects used in many hadronic object reconstruction algorithms in ATLAS such as those used to reconstruct jets and hadronically decaying taus. Clusters are seeded by cells with an energy significance (signal to noise ratio) $\Gamma = E/\sigma$ above a high threshold $\Gamma > t_{\text{seed}}$ and subsequently grow by iteratively including neighbouring cells having energy significance above a lower threshold $t_{\text{neighbour}}$ [99]. Finally, all direct neighbours on the outer perimeter are included if they have a threshold above t_{cell} . The lower threshold applied to the surrounding cells ensures that tails of showers are not discarded, while the high threshold for seed cells suppresses noise from both electronics and pile-up. The level of noise in a given cell depends on η , the amount of pile-up and the type of calorimeter module as shown in figure 3.1. In the standard ATLAS reconstruction two types of topological clusters are built: electromagnetic and combined. The electromagnetic clusters are built exclusively from EM calorimeter layers, while combined clusters use both EM and hadronic layers. The calorimeters typically have a lower energy response to hadronic showers, and lower thresholds are therefore chosen for combined clusters, $\{t_{\text{seed}}, t_{\text{neighbour}}, t_{\text{cell}}\} = \{4, 2, 0\}$, compared to electromagnetic clusters where the same sequence is $\{6, 3, 3\}$. $t_{\text{cell}} = 0$ for combined clusters indicates that all perimeter cells are included, which gives a more complete shower containment.

The above algorithm is adequate only in the case of isolated particle showers which is not typical for most ATLAS events at the current levels of pile-up, especially at high $|\eta|$. A cluster splitting algorithm is added to identify overlapping showers by separating the previously reconstructed clusters if they have identifiable local energy maxima. Local maximum cells are defined as those cells having $E > 500$ MeV with energy greater than that of any neighbouring cell and with more than a threshold number of neighbours included in the parent cluster, typically four. Each local maximum cell acts as a seed for a new cluster, which is grown according to a similar iterative process to the one described above, but only considering cells

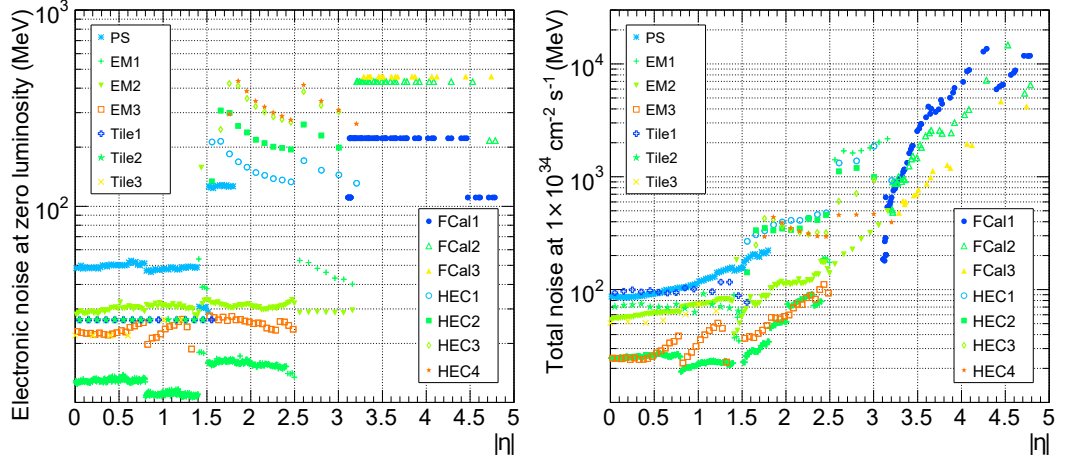


Figure 3.1: Simulated per-cell calorimeter noise for each calorimeter subdetector as a function of pseudo-rapidity. The left plot includes only electronics noise while the right plot shows the combined noise from electronics and pile-up at a collision luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ [99].

included in the parent cluster and with no cluster merging. If a particular cell is included in more than one child cluster, the two clusters with the most energetic neighbour cells will share the cell, weighted according to the cluster energies and the spatial distance between the cell and the cluster centres.

As an alternative correction to the EM-scale calibration a local cluster weighting calibration (LC) can be applied, which accounts for the shower shape and the probability p of the cluster originating from a hadronic interaction [100, 101]. A hadronic cell weight w_{HAD} is calculated according to the total cluster energy and the cell energy density by comparing these quantities to the true energy deposits found from MC simulation of charged and neutral pions. The final applied weight is then

$$w_{\text{cell}} = w_{\text{HAD}} \cdot p + w_{\text{EM}} \cdot (1 - p), \quad (3.1)$$

where $w_{\text{EM}} = 1$. Jets reconstructed from LC calibrated clusters will generally have an improved energy resolution compared to jets reconstructed from EM-scale clusters since the former take shower shape variations into account. Jets and hadronic tau decays used in this thesis are both using LC calibrated clusters.

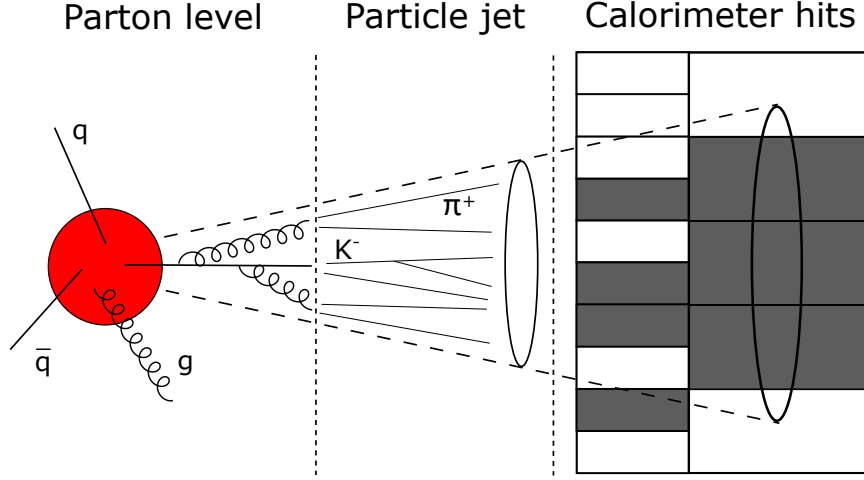


Figure 3.2: Schematic presentation of the evolution of a jet starting with partons from the hard scattering (red circle) through parton showering and hadronisation to calorimeter energy deposits.

3.3 Jets

Jets are the result of fragmentation of high-energy quarks and gluons that turn into directional sprays of particles due to colour confinement in QCD (see section 1.1.4). Observing jets in an event signifies the presence of high- p_T quarks and gluons in the final state, and jets are therefore important in a wide range of analyses including studies of the Higgs boson. Following its initial formation from scattered partons the jet undergoes the processes of parton showering and hadronisation to become a particle jet consisting of colour-neutral particles, typically mesons such as kaons and pions. Finally, it goes through showering as it enters the calorimeters and interacts with the detector material. This evolution is illustrated in figure 3.2. These steps need to be accurately reproduced in simulation to connect the particles produced in the hard scattering with the jets reconstructed in data in the calorimeters. Jets are reconstructed using topological clusters as outlined in section 3.2, and their energy is calibrated to the jet-energy scale described below. Jets containing long-lived hadrons with b -quarks, also called b -jets, can be identified through b -tagging algorithms that will be introduced in section 3.3.2. The hadronisation of highly boosted partons into jets of colour neutral particles is a non-perturbative QCD process, which prevents a direct matrix element calculation of the properties of the resulting jets such as the particle multiplicity and their distributions in space and energy. Instead, hadronisation algorithms based on string and cluster fragmentation have been developed and are used in MC generators such as PYTHIA [102] and

HERWIG [103] respectively. These methods allow the software to approximate the underlying processes and describe the characteristics of observed data from the detector.

3.3.1 Jet Reconstruction and Calibration

The following section highlights relevant aspects of jet reconstruction algorithms described in more detail in e.g. [104, 105]. The energy deposits in the calorimeters have to be analysed and collected into individual jets in order to obtain a comparison between experimental data and simulation. Many such jet clustering algorithms have been developed with various robustness criteria. The ideal jet clustering algorithm should be insensitive to non-perturbative effects such as hadronisation and underlying event contamination while being stable under infrared and collinear radiation of additional partons (IRC safe) [104]. Accurate determination of the jet's mass and energy requires a sufficiently large jet area to be included. On the other hand, a smaller jet area reduces the amount of underlying event and pile-up picked up by the jet, making the choice of jet size a trade-off between accuracy and rejection of contamination. A collinear unsafe algorithm will be affected by the splitting of a hard particle by changing the number and properties of the jets. An infrared unsafe algorithm is likewise unstable under the emission of soft gluons. The above mentioned criteria guarantee that cross sections of hadronic jet measurements are comparable to those calculated at parton level.

Many jet algorithms exist and can be categorised as either *cone* or *clustering* algorithms. Algorithms of the cone variety assume that jets form within conical regions, and the clustering is therefore performed in $\eta - \phi$ space with jets having circular boundaries. Their implementation is relatively easy, but they are generally IRC unsafe and will not be discussed further. Clustering algorithms instead work by grouping particles based on energy and geometrical closeness. The analysis presented in this thesis reconstructs jets based on the anti- k_T [106] clustering algorithm, which is implemented as part of the FASTJET software package [107]. It clusters objects (particles, topological clusters) sequentially in such a way that soft objects will tend to cluster with their hard counterparts first, resulting in jets with one or more hard objects surrounded by softer objects. This is achieved by defining the distance measures

$$d_{ij} = \min(p_{T,i}^{2n}, p_{T,j}^{2n}) \frac{\Delta_{ij}^2}{R^2}, \quad (3.2)$$

$$d_{iB} = p_{T,i}^{2p}, \quad (3.3)$$

where $\Delta_{ij} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and y_i , ϕ_i and $p_{T,i}$ are respectively the rapidity, azimuthal angle and transverse momentum of particle i . A radius parameter R is added to control the radial size of the reconstructed jets (which will not necessarily be perfectly conical), and the parameter p determines the relative power of the momentum versus geometrical scales. In the case of the anti- k_T algorithm this is set to $p = -1$. All jets in this thesis are reconstructed using $R = 0.4$. The algorithm computes the distance measures of all input objects and identifies the minimum value. If this is of type d_{ij} the objects i and j are clustered together, while a minimum value of type d_{iB} means object i is defined as a jet and is no longer considered for further clustering. The procedure is repeated until all objects are included in a jet. A useful property of the anti- k_T algorithm is its stable passive jet area¹ of πR^2 which is independent of Δ_{ij} . Corrections such as energy subtractions accounting for underlying event and pile-up are jet area sensitive and therefore benefit from this property. The algorithm is also IRC safe and relatively light in terms of computing power.

Additional detector information is saved including information about associated tracks in the central detector. This information can be used to calculate a jet vertex fraction (JVF) [109] for jet_i with respect to the vertex PV_j as

$$\text{JVF}(\text{jet}_i, \text{PV}_j) = \frac{\sum_k p_T(\text{track}_k^{\text{jet}_i}, \text{PV}_j)}{\sum_n \sum_l p_T(\text{track}_l^{\text{jet}_i}, \text{PV}_n)}, \quad (3.4)$$

where k runs over all tracks originating from PV_j matched to jet_i , n over all primary vertices in the event and l over all tracks originating from PV_n matched to jet_i . This ratio is useful in the suppression of vertices originating from pile-up. The 2012 data set used in this thesis uses a requirement of $|\text{JVF}| > 0.5$ on jets with $p_T < 50$ GeV and $|\eta| < 2.4$. Jets with higher p_T very rarely originate from pile-up. The systematic uncertainty on the selection efficiency was found to be negligible at less than 1% [109]. Additional jet cleaning criteria are applied in order to remove jets from extraneous background sources such as cosmic ray muons, calorimeter noise and so-called beam-gas events, where a proton from the beam interacts with the residual gas inside the beam pipe [110]. The efficiency of jets from actual hard scattering events passing these quality criteria exceeds 99.8% at the standard loose working point. An efficiency working point defines a set of criteria with a tradeoff between signal efficiency and purity.

In addition to the local cluster weighting calibration performed on topological

¹Jet area is usually defined as either the *passive* area measuring a jet's susceptibility to point-like radiation, or the *active* area measuring its susceptibility to diffuse radiation [108].

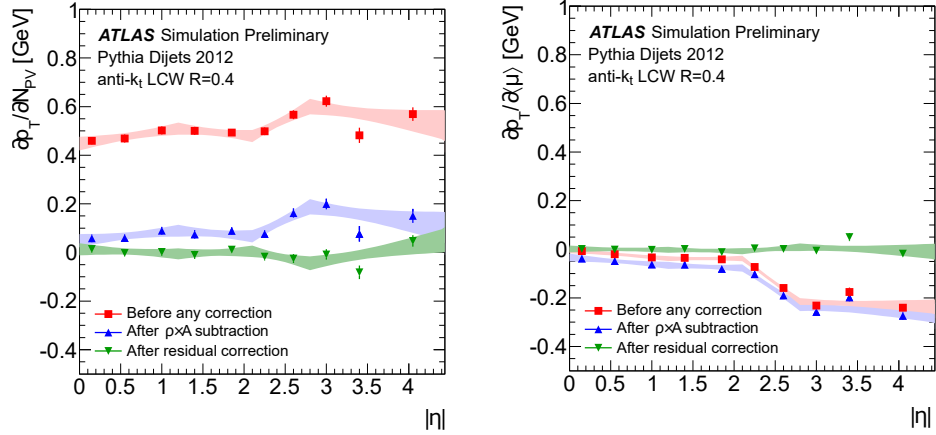


Figure 3.3: A comparison of truth particle jets and reconstructed jets in simulated dijet events showing the dependence of the reconstructed jet p_T on in-time pile-up (left) and out-of-time pile-up (right) at the different correction stages: before any correction, after $\rho \cdot A$ subtraction and after the residual correction. The values are plotted in bins of jet $|\eta|$ in addition to the 68% error bands of a fit to the same functional form as the residual correction [109].

clusters, a jet level calibration is applied to more accurately describe the discrepancies between parton level jets and the jets reconstructed in the calorimeters. A correction is applied to subtract the diffuse energy stemming from the underlying event and in-time pile-up according to the equation [109]

$$p_T^{\text{jet,corr}} = p_T^{\text{jet}} - \rho \cdot A, \quad (3.5)$$

where ρ is the median p_T density in the event and provides a direct estimate of the pile-up activity, while A is the jet area that gives an estimate of a jet's sensitivity to pile-up. The correction reduces the dependence of the jet energy on the number of primary vertices from several hundred MeV per vertex to under 200 MeV per vertex in all cases and under 100 MeV for central jets. Additional residual corrections proportional to the number of reconstructed pile-up vertices $N_{\text{PV}} - 1$ (to account for in-time pile-up) and proportional to $\langle \mu \rangle$ (to account for out-of-time pile-up) further improve this performance. Figure 3.3 shows the dependence of the reconstructed jet p_T on pile-up at various correction stages. The last correction stage is a set of data-driven corrections designed to account for small differences in the jet response between the simulated and physical calorimeter. The components of this correction step are listed in section 5.2 alongside estimated combined systematic uncertainties resulting from the various calibration steps described above.

3.3.2 b -tagging

Hadrons containing b -quarks have relatively long lifetimes leading to typical decay lengths of $\beta\gamma c\tau \approx 5$ mm at energies around 50 GeV [5]. b -tagging is the process of identifying b -hadrons through their secondary decay vertex or properties of their decay products. This is of great importance in the $H \rightarrow b\bar{b}$ search channel but is also useful in e.g. $H \rightarrow \tau\tau$ through the ability to reduce backgrounds from $t\bar{t}$ events. ATLAS has several b -tagging algorithms that identify b -jets through different means [83, 111, 112]:

- *Impact parameter-based:* Algorithms of this type take as input the transverse and longitudinal impact parameters of all the tracks in a jet. Particle tracks originating from secondary vertices located a significant distance away from the primary vertex will likely have large impact parameters that can be used to identify them. The IP3D algorithm uses a likelihood ratio technique where the input is compared to pre-defined distributions of d_0/σ_{d_0} and z_0/σ_{z_0} for b -jet and light jet hypotheses obtained from simulation. Actual reconstruction of the secondary vertices is not required.
- *Secondary vertex-based:* The secondary vertex formed by the decay products of the b -hadron can be analysed if it is possible to reconstruct it. The SV1 algorithm exploits three of the vertex properties: the invariant mass of the associated tracks, the ratio of the sum of energies of the tracks in the vertex to the sum of energies of all tracks in the jet, and the number of two-track vertices. These variables are combined using a likelihood ratio technique.
- *Decay chain reconstruction:* b -hadrons mostly decay through an intermediate c -hadron. The JETFITTER algorithm searches for decay chains of this type by attempting to find a common line on which the decay vertices of the b - and c -hadron lie. A likelihood discriminant is then built using similar variables to those used in the SV1 algorithm but also including decay topology information.

Due to the likelihood discriminants used in IP3D, SV1 and JETFITTER they can easily be combined to take advantage of the different efficiencies and rejection rates of the algorithms in various jet kinematic regions. ATLAS uses a combination of the algorithms called MV1 that significantly improves both the tagging efficiency and the mistag rate over each of the individual algorithms. It makes use of a neural network discriminant constructed using the outputs of SV1, IP3D and the combination IP3D+JETFITTER. The present analysis uses the MV1 algorithm at a working point having a selection efficiency of 70%, measured in $t\bar{t}$ MC events.

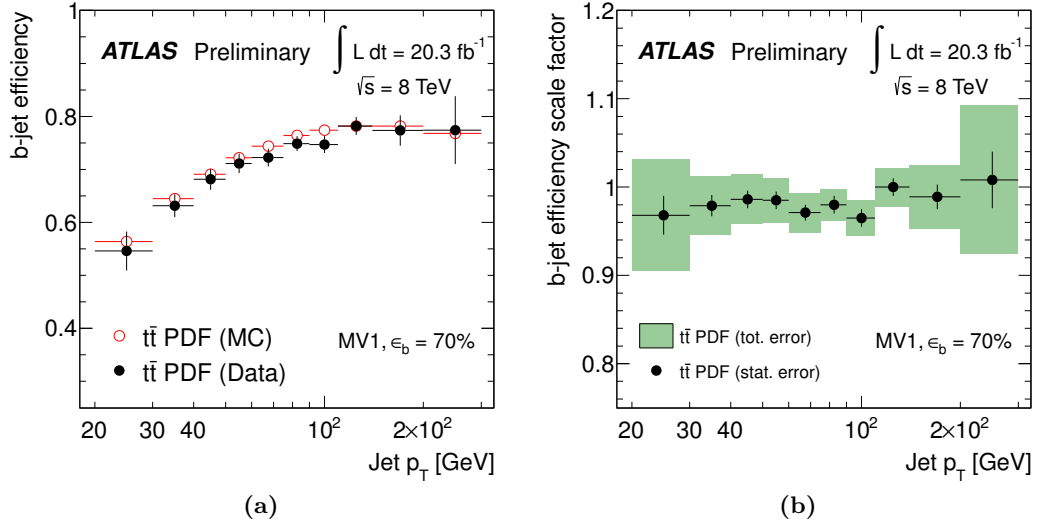


Figure 3.4: The efficiency of the MV1 b -tagging algorithm at the 70% efficiency working point measured in dileptonic $t\bar{t}$ data events and in MC simulation (a) and the scale factors found as the ratio of the two, including their uncertainties (b) [113].

Several analyses have measured the tagging efficiencies and mistag rates of the various b -tagging algorithms. Both dileptonic $t\bar{t}$ events [113] and muons from semi-leptonic b -hadron decays [114] have been used to extract tagging efficiencies. An example is shown for $t\bar{t}$ events in figure 3.4 comparing the measured b -tagging efficiencies in data and Monte Carlo for the MV1 algorithm. Scale factors are derived as the ratio between MC and data values and are used to correct the efficiencies in simulation. Systematic uncertainties on these scale factors are described in section 5.2.

3.4 Electrons

ATLAS has the ability to identify electrons with high purity and obtain precise measurements of both their momenta in the tracking detectors and their energies in the electromagnetic calorimeter, making them useful experimental signatures in a wide range of physics analyses. Their typical calorimeter shower shape also allows for their efficient identification as well as rejection of hadronic jets. These characteristics are also present at the trigger level, enabling the use of single electron triggers in the collection of data samples in the $H \rightarrow \tau_e \tau_{\text{had}}$ decay channel.

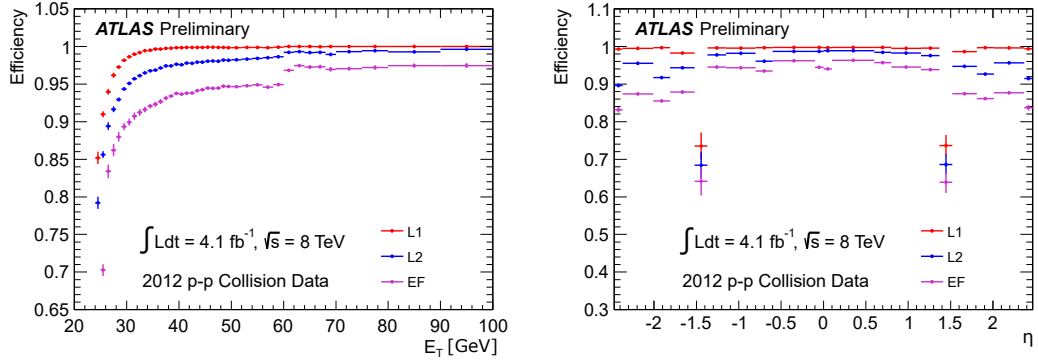


Figure 3.5: *Efficiencies of the single electron trigger at the three levels of the ATLAS trigger system in bins of transverse energy (left) and pseudo-rapidity (right) of the reconstructed electron. Measurements were made in a sample of $Z \rightarrow ee$ events using a tag-and-probe method [94].*

3.4.1 Electron Trigger in $\tau_\ell \tau_{\text{had}}$

Online trigger chains used to select electrons follow the ATLAS Run 1 three-stage design as described in section 2.3.5. The hardware L1 trigger defines regions of interest (ROI) with a sliding window algorithm using trigger towers as input with a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The L2 software trigger takes as input the ROIs and further analyses them at full calorimeter granularity to form clusters that are combined with ID tracks using a track-to-cluster matching algorithm. The EF level trigger has access to algorithms similar to the full offline reconstruction and identification algorithms. Loose identification criteria and isolation requirements are imposed to preserve a high selection efficiency.

The $\tau_\ell \tau_{\text{had}}$ analysis channel makes use of single-electron trigger items that require the presence of at least one electron candidate above a p_T threshold determined by the specific trigger. The technical names of the applied triggers are *e24vhi_medium1* and *e60_medium1*. The number following *e* denotes the p_T threshold in GeV above which the trigger accepts an electron candidate, while *medium* refers to the chosen efficiency of the applied identification criteria for clusters and tracks. An additional veto (*vh*) is applied on activity in the hadronic calorimeter region beyond the electron EM calorimeter energy for the 24 GeV electron trigger, designed to reduce the acceptance of misidentified jets. The associated track must also be isolated (*i*) with respect to other tracks. The two triggers are combined using a logical OR in order to increase the efficiency at high p_T where the isolation requirements and hadronic veto of the 24 GeV trigger become inefficient. Figure 3.5 shows the efficiency of this trigger combination in bins of the reconstructed electron trans-

verse energy and pseudo-rapidity measured in $Z \rightarrow ee$ events using a tag-and-probe method. The tag electron is required to pass tight identification criteria whereas the probe electron has no such requirements. At invariant masses of the di-electron system close to the Z peak the electron purity is high, and the sample can be used to determine the trigger efficiency. An abrupt increase in the trigger efficiency is observed at $p_T \approx 60$ GeV where the turn-on curve of the non-isolated trigger is located.

3.4.2 Electron Reconstruction and Identification

Only electrons within the tracker acceptance of $|\eta| < 2.47$ are included in the analysis presented in this thesis, and hence electron reconstruction in the forward detectors will not be covered. Offline reconstruction of electrons uses a sliding window algorithm similar to the one used at the EF trigger level but with a slightly different cluster size of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ [99] with a transverse energy requirement of 1.5 GeV and loose shower shape criteria. The electron track reconstruction is performed in an ROI of size $\Delta R = 0.3$ around each cluster in two steps. The first step uses pattern recognition with an initial pion hypothesis in terms of energy loss at material surfaces. If unsuccessful, an alternative algorithm allowing a maximum of 30% energy loss with an electron hypothesis is attempted. This improves the electron reconstruction efficiency [115]. The second step associates tracks with clusters under various geometric criteria related to the position of a track compared to the location of the EM cluster. The tracks passing this matching step are subsequently refitted using a Gaussian Sum Filter algorithm that optimises the electron track parameters by accounting for non-linear bremsstrahlung effects [116].

The electron energy is determined from the energy inside the reconstructed cluster and subsequently corrected by estimating the energy lost before reaching the calorimeter, the energy outside the reconstructed cluster as well as beyond the EM calorimeter [117, 118]. The EM scale of the cells was calibrated through testbeam measurements, and the uncertainty on the absolute scale is reduced as a function of η using data-driven fits to di-electron invariant mass spectra from $Z \rightarrow ee$ and $J/\Psi \rightarrow ee$ decays. The sources of uncertainty include imperfect knowledge of the material in front of the calorimeter, the energy scale of the presampler layer and the background estimation and fitting method used in the data-driven calibration. These are far below $\pm 4\%$ in the barrel region and up to $\pm 1\%$ in the endcap region [118]. The electron energy resolution is also found through $Z \rightarrow ee$ and $J/\Psi \rightarrow ee$ data events. The resolution in data is approximately 1% worse than in MC, and the simulation is smeared accordingly.

Backgrounds from jets as well as electrons originating from heavy flavour hadron decays and photon conversions that are not part of the hard scatter have a probability of contaminating the reconstructed electron samples [117]. A series of identification criteria are introduced to reject these backgrounds with minimal impact on the electron efficiency. These rely on several observables, one of which is the shower shape that allows a comparison with the expected development of an electromagnetic shower. Quality measures are also applied to track reconstruction as well as the track matching to clusters. High pile-up activity in an event can influence the efficiency of the electron identification due to tracks from additional primary vertices entering the shower and potentially altering the energy fraction in the core of the shower and the hadronic activity beyond the electromagnetic calorimeter. High priority is given to reducing the dependence of the electron identification efficiency on pile-up. Several working points are defined containing varying collections of criteria giving different tradeoffs between background rejection rate and signal efficiency. The standard working points are denoted *loose*, *medium* and *tight* corresponding to the amount of background rejection having signal efficiencies of roughly 95%, 85% and 75% respectively for reconstructed electron candidates with $E_T > 20$ GeV [119].

$Z \rightarrow ee$ and $J/\Psi \rightarrow ee$ events are used to measure the electron identification efficiencies using a tag-and-probe method [115]. The final efficiency is composed of a reconstruction efficiency and a separate identification efficiency. The reconstruction efficiency is nearly a constant 99% for electrons with transverse energy $E_T > 20$ GeV, while the identification efficiency is dependent on E_T . Figure 3.6 compares the electron identification efficiency between data and simulated events as a function of E_T and the number of reconstructed primary vertices. The ratio is used to correct the simulated samples to match the efficiency found in data. The plot vs. η reveals a dip in the identification efficiency in the transition region between barrel and endcap ($1.37 < |\eta| < 1.52$). Electrons are only selected outside this region.

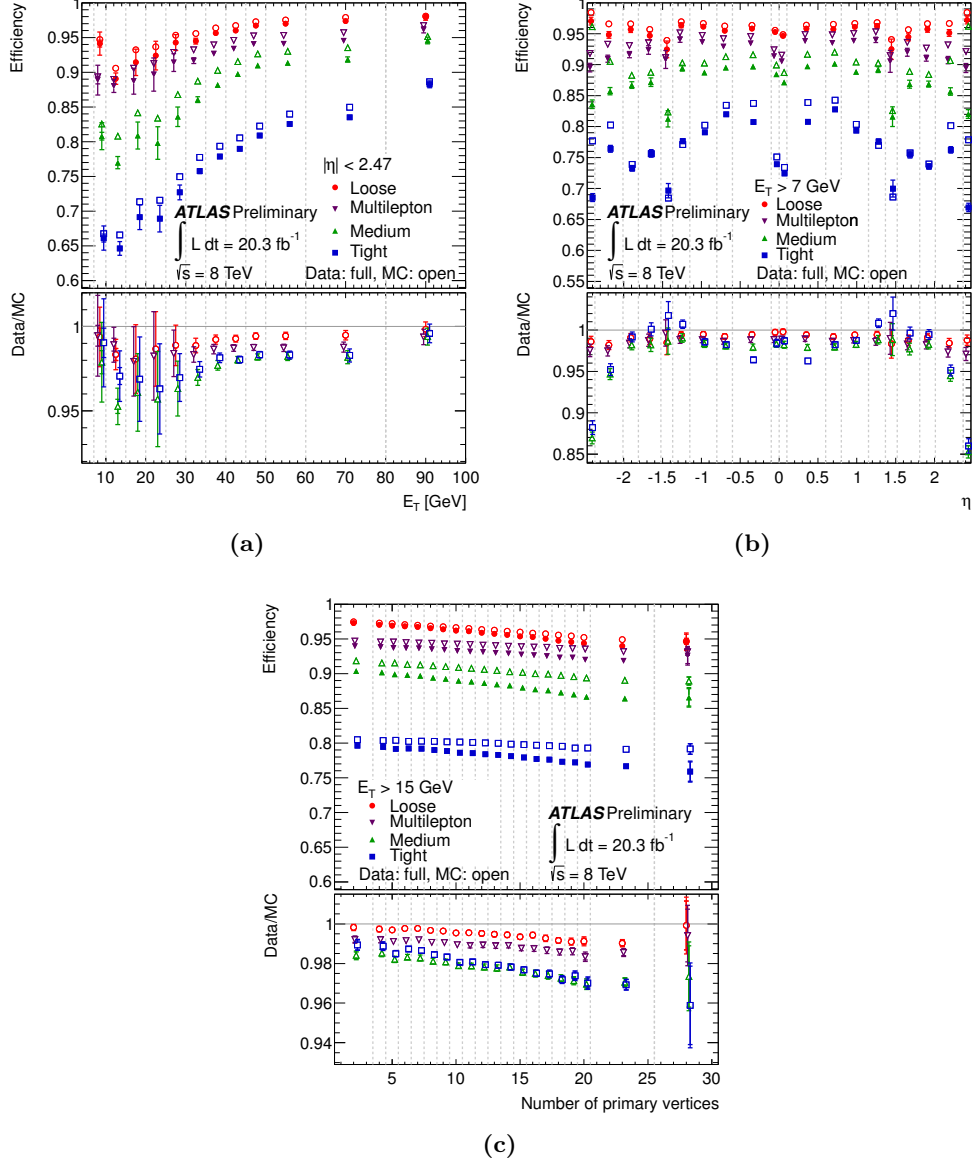


Figure 3.6: Electron identification efficiencies in bins of transverse energy (a), pseudo-rapidity (b) and number of reconstructed primary vertices (c) at efficiency working points loose, multilepton, medium and tight. Efficiencies are measured in $Z \rightarrow ee$ events with a tag-and-probe method for both data and simulation. The ratios in the lower part of the plots are used to correct the simulated samples [115].

3.5 Muons

Muons are easily distinguishable from electrons and jets due to their relatively weak interaction with the detector material. At typical production energies their lifetime in the lab frame permits them to traverse the entire detector before decaying, leaving a unique signature in the form of an ID track, minimal energy losses in the calorimeters and a track in the muon spectrometer located beyond the calorimeters. Toroid magnets immerse the muon spectrometer in a magnetic field that allows for a separate momentum measurement (see section 2.3.4). Background rejection rates and momentum precision are both excellent, making muons ideal trigger signatures.

3.5.1 Muon Trigger in $\tau_\ell\tau_{\text{had}}$

The muon trigger follows the ATLAS three-level trigger design [120]. The L1 hardware level triggers on coincident hits in the RPC inside $|\eta| < 1.05$ and TGC in the endcap region $1.05 < |\eta| < 2.4$ of the muon spectrometer (MS). The muon momentum is estimated using the width of the coincidence window and is passed on to the subsequent trigger layers together with the geometrical location of the hits. The data output from L1 is reduced by defining regions of interest (ROI) around the set of hits. Cables and other equipment servicing the ID and calorimeters occupy the region at $\eta = 0$ producing a ‘crack’ that limits the acceptance of the MS and hence the L1 trigger. The software L2 trigger has access to precision tracking information from the MDT modules within the ROI. By referring to fast lookup-tables containing pre-defined track shapes matched to momentum values it is possible to quickly assign a transverse momentum estimate to the tracks. The tracks from the inner detector are also matched to the tracks from the MS to further increase the precision. Background rejection can be added to a chain through isolation variables calculated from neighbouring tracks to the muon candidate as well as reconstructed calorimeter energy deposits in cones around the muon candidate track (see also section 2.5). At EF level the employed algorithms are functionally very close to their offline counterparts and have access to the full detector information. Combined ID and MS tracks are reconstructed using two algorithms that use either an ID or MS track as the seed and subsequently extrapolate it outward to the MS or inward to the ID. These are called inside-out and outside-in algorithms respectively.

The $\tau_\ell\tau_{\text{had}}$ decay channel uses a single muon trigger with the chain name *mu24i_tight*. The number 24 following *mu* refers to the online transverse momentum threshold imposed by the trigger in GeV, which has been chosen such that unprescaled operation is possible within the bandwidth of the trigger system. The *i*

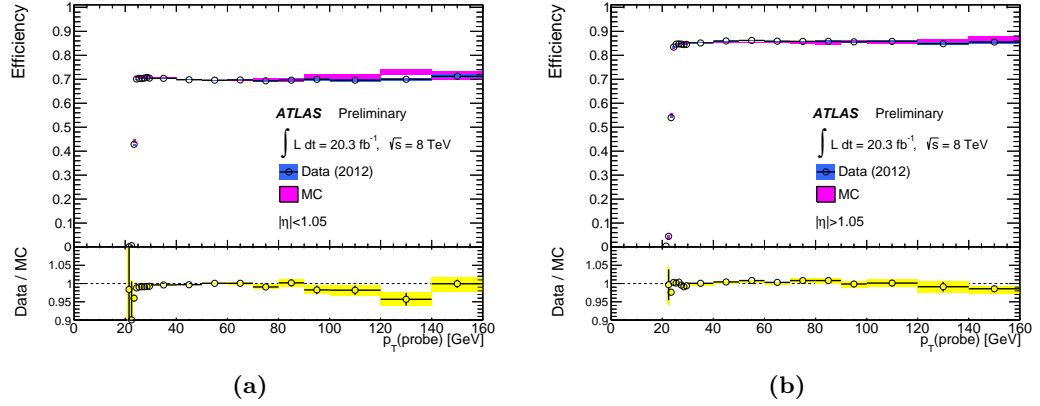


Figure 3.7: Efficiencies of passing the single muon triggers μ_{24i_tight} or μ_{36_tight} measured in $Z \rightarrow \mu\mu$ events using a tag-and-probe method in the barrel (a) or endcap (b) regions plotted as a function of the probe muon p_T . The efficiency ratios between data and MC are in good agreement within the momentum region shown [94].

indicates the use of isolation criteria to reduce backgrounds. A procedure similar to the electron case is used to determine the trigger efficiency by analysing a sample of $Z \rightarrow \mu\mu$ events using a tag-and-probe method [120]. An isolated tag muon candidate track is required to be geometrically matched with the initial trigger object, while an isolated probe muon candidate must have opposite charge. Requiring that the invariant mass of the muon system must be compatible with the Z mass results in a high-purity sample with backgrounds from other processes representing less than 1% of the events. Trigger efficiencies at $\sqrt{s} = 8$ TeV are shown for both data and simulated samples in figure 3.7 as a function of the muon transverse momentum. Differences of a few percent between data and MC are used to correct the simulated samples used in the analysis.

3.5.2 Muon Reconstruction and Identification

The low-background environment for muons in ATLAS allows the reconstruction and identification algorithms to be designed with precision momentum measurements as the main goal. Three different reconstruction algorithms are defined depending on the available data from the detector subsystems [121]:

- *Stand-alone (SA)*: Only MS information is included in the muon reconstruction. Tracks are extrapolated back to the interaction point while taking into account the expected energy loss in the calorimeters.

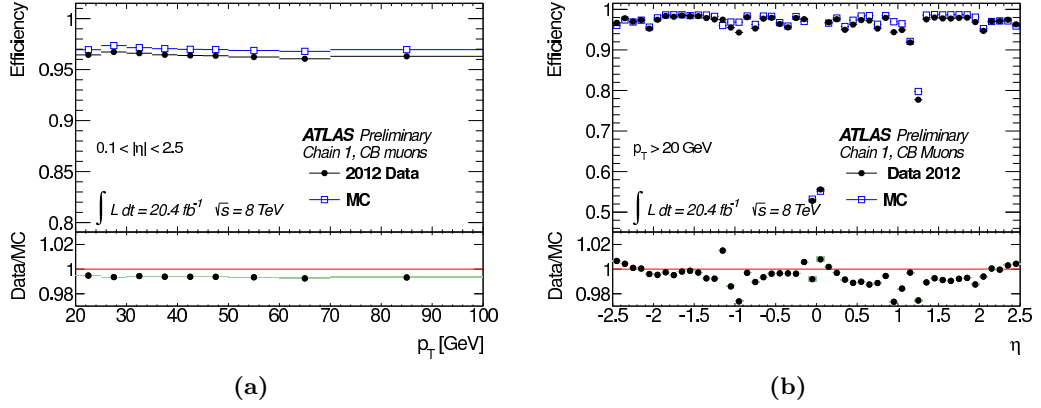


Figure 3.8: Muon reconstruction efficiencies as a function of transverse momentum (a) and pseudo-rapidity (b) measured in a $Z \rightarrow \mu\mu$ sample using a tag-and-probe method. Discrepancies of roughly 1–2% between data and simulated samples are observed [121].

- *Segment-tagged (ST)*: These objects are constructed from ID tracks with at least one possible associated MS track.
- *Combined (CB)*: A subset of segment-tagged tracks, the combined muon objects combine ID and MS tracks while imposing quality criteria on the covariance matrices of the two track fits. The inclusion of both subdetectors and covariance matrix requirements ensures a higher momentum resolution and signal purity than the other reconstruction algorithms.

The analysis presented in this thesis uses only combined muons. An important prerequisite to optimal muon reconstruction performance is detailed knowledge of the misalignment of the muon chambers. This can be measured in studies of cosmic ray events or in separate data runs where the toroid magnets are switched off. Technical limitations of the detector result in pseudo-rapidity intervals where the muon reconstruction efficiency is significantly lower than in the rest of the tracking acceptance inside $|\eta| < 2.5$. As previously mentioned, the region at $|\eta| = 0$ has support structures and service equipment that limit the availability of the MS, while in the region $1.1 < \eta < 1.3$ some chambers were not installed until Run 2, making them unavailable in the 2012 dataset. The efficiencies of muon reconstruction and identification are measured in $Z \rightarrow \mu\mu$ events following the same tag-and-probe method used to find the trigger efficiencies. Figure 3.8 shows the reconstruction efficiency as a function of transverse momentum and pseudo-rapidity, where the latter has visible drops at the mentioned regions with sub-optimal coverage. The efficiency

is well-described in simulation, having relative errors of 1–2% [121] compared to data. Scale factors are derived from these ratios and applied as corrections to the simulated samples.

In addition to efficiency discrepancies between data and MC, the scale of the muon momentum also needs to be investigated. The relative momentum resolution can be parametrised as

$$\sigma(p_T)/p_T = a \oplus (b \times p_T), \quad (3.6)$$

where a denotes a constant contribution originating from multiple scattering and b is a contribution proportional to the transverse momentum due to the spatial resolution of the detector. The parameters are derived in bins of pseudo-rapidity through studies using samples of $Z \rightarrow \mu\mu$, $J/\Psi \rightarrow \mu\mu$ and $\Upsilon \rightarrow \mu\mu$ events. The momentum scale itself is accurate to within the order of a permille, while the constant and p_T -dependent resolution terms are of the order $\pm 2\%$.

3.6 Hadronically Decaying Taus

The tau lepton is the only lepton capable of decaying hadronically due to its relatively high mass of $m_\tau \sim 1.8$ GeV. This gives the particle a rich spectrum of decay channels with a myriad of different final states. Having a mean lifetime of $290.3 \cdot 10^{-15}$ s [5] corresponding to a proper decay length of 87 μm means that τ leptons typically decay before reaching the tracking systems of ATLAS, and can therefore only be identified through their decay products. Figure 3.9 shows examples of the most common decay channels. The branching ratios of $\tau^- \rightarrow \nu_\tau e^- \bar{\nu}_e$

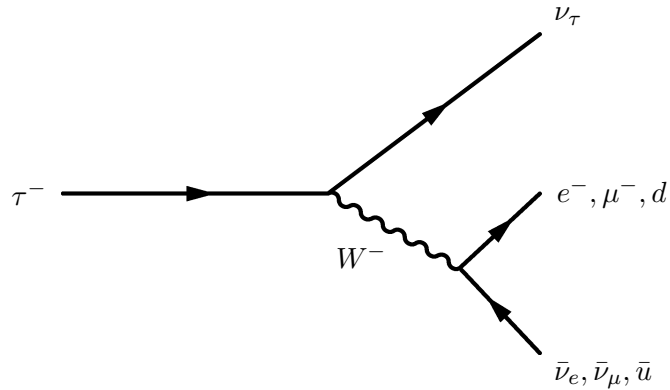


Figure 3.9: Feynman diagram depicting common decays of the τ lepton through emission of a W boson.

Decay mode	Branching fraction [%]
$\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$	17.83 ± 0.04
$\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$	17.41 ± 0.04
$\tau^- \rightarrow \pi^- \pi_0 \nu_\tau$	25.52 ± 0.09
$\tau^- \rightarrow \pi^- \nu_\tau$	10.83 ± 0.06
$\tau^- \rightarrow \pi^- \pi^0 \pi^0 \nu_\tau$	9.30 ± 0.11
$\tau^- \rightarrow \pi^- \pi^0 \pi^0 \pi^0 \nu_\tau$	1.05 ± 0.07
$\tau^- \rightarrow K^- \nu_\tau$	0.700 ± 0.010
$\tau^- \rightarrow K^- \pi^0 \nu_\tau$	0.429 ± 0.015
$\tau^- \rightarrow \pi^- \pi^- \pi^+ \nu_\tau$	8.99 ± 0.06
$\tau^- \rightarrow \pi^- \pi^- \pi^+ \pi^0 \nu_\tau$	2.70 ± 0.08

Table 3.1: Most common tau lepton decay modes and their branching fractions. These are the result of a combined fit to data [5].

and $\tau^- \rightarrow \nu_\tau \mu^- \bar{\nu}_\mu$ are both approximately 17-18%, leaving a large fraction of the total width to hadronic decays. These decays have a wide array of possible final states due to the many ways for the resulting quarks to hadronise. However, they can be categorised according to the number of charged decay products. Charge conservation limits these to final states with 1, 3 or 5 charged particles, defining the number of ‘prongs’, e.g. a decay with 3 charged particles in the final state is a three-prong decay. One-prong and three-prong decays to charged pions represent 72% and 22% [5] of hadronic decays respectively, with charged kaons being present in the majority of the remaining hadronic decays. Five-prong decays are rare and will not be discussed further in this thesis. Table 3.1 lists the most common leptonic and hadronic decay modes.

No dedicated reconstruction algorithms have been defined for the leptonic decay modes, denoted τ_{lep} , since these are effectively indistinguishable from prompt electron or muon production with missing energy. Existing lepton reconstruction chains are used in these cases. The hadronic decay modes, denoted τ_{had} , result in experimental signatures with an odd number of charged hadrons, mostly in the form of pions, possibly a number of neutral hadrons as well as missing transverse energy due to the additional invisible tau neutrino. The visible part of the tau decay will be denoted $\tau_{\text{had-vis}}$ in the following. Electroweak tau production typically leads to kinetic energies much higher than the tau mass, resulting in narrower hadronic tau decay showers compared to the showers produced by jets from the hadronisation of quarks and gluons. Identification criteria to distinguish between hadronic jets and hadronic tau decays can be defined from these shower shape differences and the specific number of charged hadrons present. Separate identification criteria are

employed to discriminate against electrons that also typically have relatively narrow shower shapes and a single charged track. The reconstruction and identification of tau leptons is described in more detail in section 3.6.1. The subsequent section outlines an energy calibration performed to correct the reconstructed energy from the calorimeters to the true energy of the visible decay products.

3.6.1 Reconstruction and Identification of Hadronic Tau Decays

Anti- k_T jets with a distance parameter of $R = 0.4$ are used in the reconstruction of hadronic tau decays [122,123]. Analogously to the quark- and gluon-initiated jet reconstruction the input is in the form of topological clusters with LC calibration applied, with the additional acceptance criteria $p_T > 10$ GeV and $|\eta| < 2.5$, corresponding to the tracking acceptance of ATLAS. Each tau candidate is associated with a vertex by a dedicated Tau Jet Vertex Association algorithm (TJVA) [124] that identifies the vertex with the highest jet vertex fraction. This vertex is used as the origin when calculating cell- and cluster-based variables. The four-momentum of a $\tau_{\text{had-vis}}$ candidate is defined in terms of the three variables p_T , η and ϕ while the mass is defined to be zero, making the transverse momentum and the transverse mass E_T identical. The $\tau_{\text{had-vis}}$ barycentre is found as the sum of the four-vectors of its constituent topological clusters and is used to define the corresponding $\tau_{\text{had-vis}}$ axis in (η, ϕ) by using clusters within $\Delta R < 0.2$ around the barycentre. The narrower cone was introduced in 2012 to exploit the collimated behaviour of hadronic tau decays to suppress clusters originating from pile-up that would be included at larger cone sizes. The energy scale of $\tau_{\text{had-vis}}$ candidates is calibrated separately from the jet energy scale due to the specific composition of hadronic tau decays in terms of charged and neutral hadrons [122]. Track association with the $\tau_{\text{had-vis}}$ candidate is performed within two concentric cones around the $\tau_{\text{had-vis}}$ direction. Charged tracks subject to quality criteria [123] are counted within a core region of width $\Delta R < 0.2$, while an isolation region within $0.2 < \Delta R < 0.4$ associates further tracks used for the calculation of identification variables and is subject to the same quality criteria.

The reconstruction steps outlined above provide basically no rejection against backgrounds from jets, electrons and muons. Several discriminating variables are defined during reconstruction and are used in a separate step for background rejection. A selection of these variables is shown in figure 3.10. The fraction of the total tau energy located within the central core region $\Delta R < 0.1$ is denoted $f_{\text{core}}^{\text{corr}}$ and can be used to exploit the narrow shower shape of typical hadronic tau decays. Tracking information is also useful in jet discrimination. Quark and gluon initiated jets often

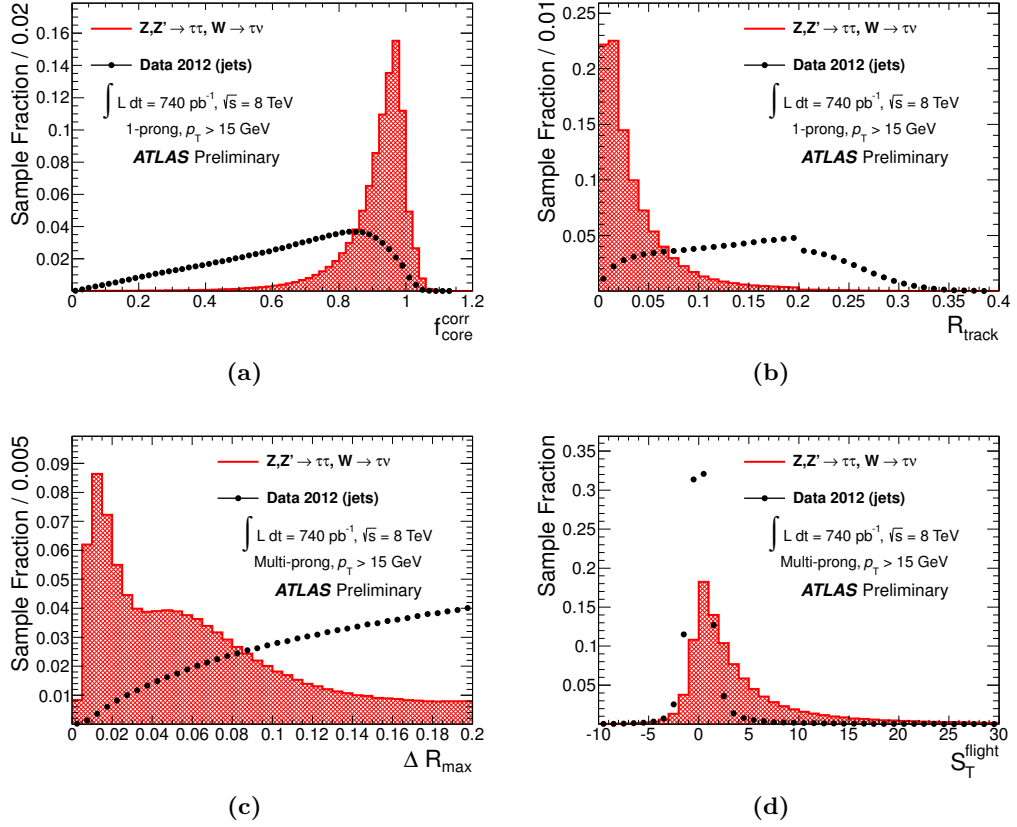


Figure 3.10: Distributions of a selection of variables used to discriminate against jets in the reconstruction of hadronic tau decays. The distributions are extracted from simulated $Z, Z' \rightarrow \tau\tau$ and $W \rightarrow \tau\nu$ signal samples and a jet background sample from the 2012 data set [123].

contain many charged tracks with a wide spread in $\eta - \phi$. The average p_T -weighted track distance from the tau axis, R_{track} , as well as the distance to the track furthest from the tau axis, ΔR_{max} , both have significant discrimination power. Despite the tau decay happening before it reaches the tracking detectors, the significance of the decay length of a reconstructed secondary vertex S_T^{flight} defined as the ratio of the decay length and its uncertainty can provide discrimination, especially in the three-prong case. A π^0 reconstruction algorithm has also been developed measuring the number of reconstructed neutral pions in the core region of the $\tau_{\text{had-vis}}$ candidate. The substructure of the $\tau_{\text{had-vis}}$ candidate can then be probed by combining kinematic information about tracks and clusters likely originating from π^0 decays, thereby improving the reconstructed four-momentum and energy resolution. The number of reconstructed tracks is not directly used in the $\tau_{\text{had-vis}}$ identification but is instead exploited at the analysis level as part of the requirements listed in table 4.7.

Variable	one-prong	three-prong
f_{cent}	•	•
f_{track}	•	•
R_{track}	•	•
$S_{\text{leadtrack}}$	•	
$N_{\text{track}}^{\text{iso}}$	•	
ΔR_{max}		•
$S_{\text{T}}^{\text{flight}}$		•
m_{track}		•
$m_{\pi^0+\text{track}}$	•	•
N_{π^0}	•	•
$p_T^{\pi^0+\text{track}}/p_T$	•	•

Table 3.2: List of discriminating variables included in the multivariate BDT discriminators used to discriminate against jets for one-prong and three-prong $\tau_{\text{had-vis}}$ candidates [122].

Jet Discrimination

Quark- and gluon-initiated jets are the main background in the reconstruction of hadronic tau decays due to the high production cross section and their similar experimental signature. A multivariate discriminator called a boosted decision tree (BDT) is trained for one- and three-pronged tau candidates individually [122] using eight and nine variables respectively (see section 4.6.1 for more details on BDTs). The variables have been chosen according to their pile-up robustness and include the discriminating variables mentioned above in one or both cases. Table 3.2 lists which variables are used for discrimination against jets for one-prong and three-prong taus respectively in the offline reconstruction. Further utilised variables are:

- f_{track} : Transverse momentum of the highest- p_T charged particle in the core region of the $\tau_{\text{had-vis}}$ candidate divided by the transverse energy sum.
- $S_{\text{leadtrack}}$: 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.
- $N_{\text{track}}^{\text{iso}}$: Number of tracks associated with the $\tau_{\text{had-vis}}$ in the region $0.2 < \Delta R < 0.4$.
- m_{track} : Invariant mass calculated from the sum of the four-momenta of all tracks in the core and isolation regions, assuming a pion mass for each track.

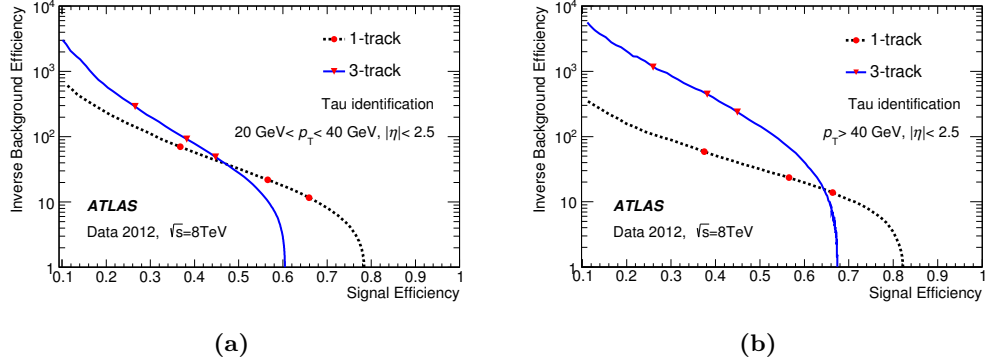


Figure 3.11: Inverse background efficiency as a function of signal efficiency in the offline tau identification in (a) a low- p_T and (b) a high- p_T $\tau_{\text{had-vis}}$ sample. Simulated $Z, Z' \rightarrow \tau\tau$ and $W \rightarrow \tau\nu$ samples are used for signal, while data from multi-jet events is used for background. The three efficiency working points described in the text are shown as red markers separately for one-prong and three-prong candidates [122].

- $m_{\pi^0+\text{track}}$: Invariant mass of the system composed of the tracks and π^0 mesons in the core region.
- N_{π^0} : Number of π^0 mesons reconstructed in the core region.
- $p_T^{\pi^0+\text{track}}/p_T$: Ratio of the p_T estimated using the track + π^0 information to the calorimeter-only measurement.

The BDTs are trained using simulated signal samples of W, Z and Z' decays to tau leptons and a QCD background sample of jet-enriched data. The BDT trained with three-prong candidates is used to classify all $\tau_{\text{had-vis}}$ candidates with two or more tracks. Three working points – *loose*, *medium* and *tight* – are defined corresponding to different identification efficiencies. The BDT working points are determined in bins of true p_T of the $\tau_{\text{had-vis}}$ candidate in order to achieve a stable efficiency as a function of transverse momentum. Good efficiency stability versus the number of reconstructed primary vertices is also achieved due to the choice of identification variables. Figure 3.11 shows background and signal efficiency curves for one-prong and three-prong $\tau_{\text{had-vis}}$ candidates in a low- p_T and a high- p_T region with included markers indicating the efficiencies at the three working points. The observed signal efficiency saturation point in both plots corresponds to the efficiency for a true $\tau_{\text{had-vis}}$ with one or three charged decay products to be reconstructed as a one-prong or three-prong $\tau_{\text{had-vis}}$ candidate. In the analysis presented in this thesis the medium identification working point is used corresponding to a signal (inverse background) efficiency of approximately 57% (25) for one-prong and 38% (500) for three-prong $\tau_{\text{had-vis}}$ objects at transverse momenta above 40 GeV.

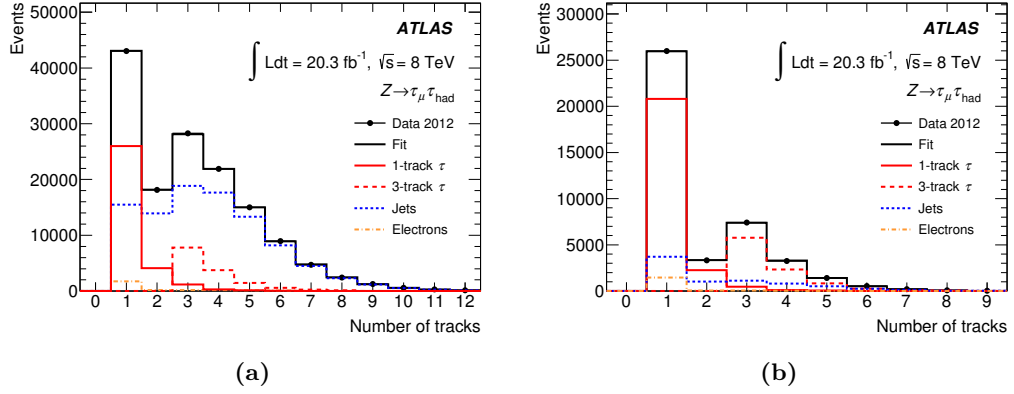


Figure 3.12: Number of tracks within $\Delta R < 0.6$ around the $\tau_{\text{had-vis}}$ axis from a tag-and-probe study in $Z \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ data in the muon channel. A fit of signal and background is performed (a) before and (b) after applying medium tau identification criteria [122].

Corresponding signal efficiencies in data can be measured using a tag-and-probe method in $Z \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ events selected by single electron and muon triggers. Trigger thresholds were chosen to be $p_T > 24$ GeV with a subsequent offline requirement on the tag lepton of $p_T > 26$ GeV. The probe objects are $\tau_{\text{had-vis}}$ candidates requiring one or three associated tracks in the core region, an electric charge of one and no geometrical overlap with other leptons. Additionally, the tag lepton and the probe $\tau_{\text{had-vis}}$ candidate should have opposite charge. A very loose requirement is added to the tau identification BDT score that strongly suppresses jets while keeping the $Z \rightarrow \tau\tau$ signal efficiency at more than 99%. Further requirements ensure suppression of $Z \rightarrow \ell\ell$ and $W \rightarrow l\nu_l$ events. The remaining background, mostly coming from multi-jet and W +jets events, can be estimated using a template fit to the distribution of the track multiplicity within a cone of $\Delta R < 0.6$ around the $\tau_{\text{had-vis}}$ axis. The shape of the background contributions can be found using various control regions including a jet-enriched region applying the signal region selection, but requiring same-sign charges of the tag and probe objects. Performing the template fit before and after applying the $\tau_{\text{had-vis}}$ identification criteria then allows to extract the signal efficiency. Figure 3.12 shows the track multiplicity of $\tau_{\text{had-vis}}$ candidates in the muon channel including the result of the template fit before and after applying the medium tau identification criteria. Scale factors are extracted as the ratio between the signal efficiencies obtained in simulation and in data and are compatible with one.

Electron and Muon Discrimination

Secondary background sources to hadronic tau decays are electrons and, in rare cases, muons mimicking a one-prong tau signature consisting of a single charged track with associated narrow energy deposits in the calorimeter. In most cases such backgrounds can be significantly reduced through overlap removal by identifying reconstructed light leptons with similar kinematics. The high reconstruction efficiency of both electrons and muons means that a large fraction of such backgrounds is removed in this manner. However, additional steps are taken to minimise the backgrounds, especially in detector regions with poor track reconstruction and particle identification efficiencies.

Discrimination against electrons is improved through an electron veto BDT (e-veto BDT) [123] trained on simulated $Z \rightarrow \tau\tau$ signal events and $Z \rightarrow ee$ background events. Training was performed separately for different τ_{had} pseudo-rapidity regions using slightly differing sets of input variables including several describing the shower profile, the fraction of transverse energy of the $\tau_{\text{had-vis}}$ candidate deposited in the EM calorimeter and the ratio of high-threshold to low-threshold hits in the TRT. The working points are defined corresponding to signal efficiencies of 95%, 85% and 75%, where the medium working point is used in the $\tau_\ell\tau_{\text{had}}$ analysis. Although the BDTs are trained exclusively on one-prong tau decays, they do provide some discrimination against electrons faking three-prong tau decays through e.g. the emission of bremsstrahlung, which is also exploited in the analysis. A $Z \rightarrow ee$ tag-and-probe method is used to measure the efficiency for electrons reconstructed as $\tau_{\text{had-vis}}$ objects to pass the electron veto in data. An isolated electron passing tight identification criteria is selected along with a probe τ_{had} candidate. Differences in efficiency between simulation and data are used to define correction factors in bins of η of the $\tau_{\text{had-vis}}$ candidate. Uncertainties on the correction factors are also η -dependent with values of $\pm 10\%$ at the loose working point and increasing slightly for medium and tight [122].

As minimum ionising particles, muons are unlikely to deposit enough energy in the calorimeters to be reconstructed as τ_{had} objects. However, if an unrelated but sufficiently energetic energy cluster can be associated with a muon, it could create a $\tau_{\text{had-vis}}$ candidate. This is generally avoided through overlap removal using the standard muon track reconstruction algorithms, leaving only cases where the muon track reconstruction fails. These cases include muons passing through an inefficient region of the MS, muons losing sufficient amounts of energy in the calorimeter to skew their tracks and very low-energy muons that are stopped in the calorimeter. Muons depositing significant amounts of energy in the calorimeter are likely to de-

posit the majority in the hadronic calorimeter, creating $\tau_{\text{had-vis}}$ candidates with a very low electromagnetic energy fraction. Such muons can also have a track momentum higher than the calorimeter energy, hinting at the fact that the track-cluster pair does not originate from just one object. A muon veto is defined that exploits these characteristics by imposing requirements on the electromagnetic energy fraction and the transverse momentum/energy of the track/cluster. The efficiency of the muon veto is greater than 96% for true τ_{had} while reducing the muon fakes by roughly 40%. However, the muon veto is not used in the present analysis due to the very low probability of misidentification.

3.6.2 Tau Energy Calibration

The use of LC calibrated topological clusters improves the energy estimate of $\tau_{\text{had-vis}}$ candidates compared to using clusters at the EM scale, but several additional effects are still not accounted for. Energy can be lost before the calorimeters, pile-up and underlying event have potential energy contributions and nearby particles can lead to out-of-cone effects. The tau energy scale (TES) calibration is designed to correct for these additional effects [125, 126]. Simulated $Z, Z' \rightarrow \tau\tau$ and $W \rightarrow \tau\nu$ samples including both in-time and out-of-time pile-up are used to derive the TES. The reconstructed and true tau energies are compared in events with at least one reconstructed $\tau_{\text{had-vis}}$ candidate and no reconstructed jets with $p_T > 15$ GeV within $\Delta R < 0.5$ of the $\tau_{\text{had-vis}}$ candidate, which is additionally required to pass medium identification criteria. Calibration constants are derived on the tau momentum and pseudo-rapidity to bring the reconstructed $\tau_{\text{had-vis}}$ candidate to the true momentum scale and direction. The calibrated momentum p_{cal}^τ is defined as

$$p_{\text{cal}}^\tau = \frac{p_{\text{LC}}^\tau}{R(p_{\text{LC}}^\tau, |\eta_{\text{reco}}^\tau|, n_p)}, \quad (3.7)$$

where R is a calibration term determined from simulated events that is a function of the reconstructed $\tau_{\text{had-vis}}$ momentum at the LC scale p_{LC}^τ , the reconstructed pseudo-rapidity $|\eta_{\text{reco}}^\tau|$ and the number of prongs (one- or multi-prong) [126]. The τ response is defined as the ratio of p_{LC}^τ to the true $\tau_{\text{had-vis}}$ momentum $p^{\tau\text{-truevis}}$ and is derived in bins of the true visible momentum and $|\eta_{\text{reco}}^\tau|$. In each bin the response is fitted with a Gaussian to obtain a mean value that is associated with the average reconstructed momentum in the bin, and the response as a function of the reconstructed momentum is then fitted with an empirically derived functional form. This brings the reconstructed momentum to the true scale within 1–2%. Figure 3.13 shows the momentum response curves for 1-prong and multi-prong $\tau_{\text{had-vis}}$ objects,

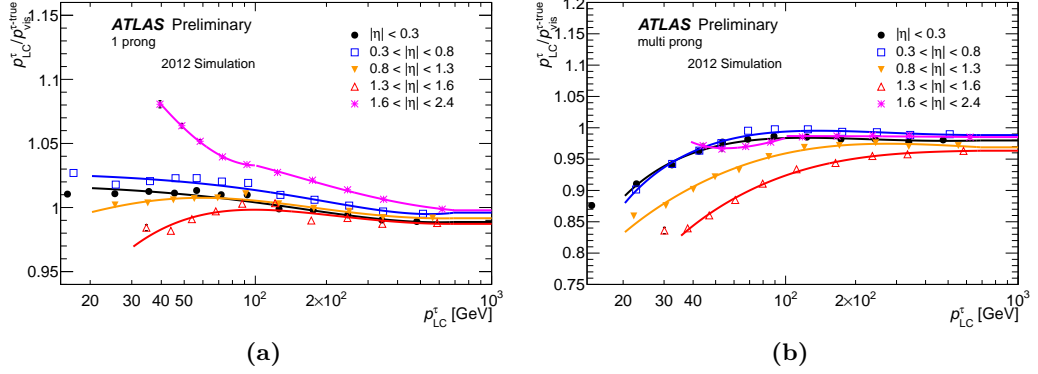


Figure 3.13: Tau momentum response curves as a function of the reconstructed visible momentum at the LC scale p_{LC}^τ for (a) one-prong and (b) multi-prong $\tau_{\text{had-vis}}$ objects [126].

where multi-prong refers to candidates with at least two reconstructed tracks.

An η -dependent discrepancy in the response in transverse momentum at the level of a few percent is still observed in areas such as the transition region between the EM barrel and endcap. Poorly reconstructed clusters in this region lead to underestimated energies that can potentially introduce a bias in the reconstructed pseudo-rapidity. The η value of the $\tau_{\text{had-vis}}$ is corrected to account for this effect based on the same simulated samples used in the momentum correction. The size of the correction does not exceed 0.01 units in any region of the detector.

In addition to the corrections on momentum and pseudo-rapidity described above it is also necessary to correct for changes in the energy response due to pile-up conditions. A correction depending on the number of reconstructed primary vertices N_{PV} is introduced which further corrects the tau momentum such that

$$p_{\text{pile-up}}^\tau = A(|\eta_{\text{reco}}^\tau|, n_p)(N_{\text{PV}} - \langle N_{\text{PV}} \rangle), \quad (3.8)$$

where $\langle N_{\text{PV}} \rangle$ is the average number of reconstructed primary vertices in the sample used to derive the correction. A is a parameter determined in bins of $|\eta_{\text{reco}}^\tau|$ and n_p , the number of prongs, from a linear fit as a function of the number of primary vertices.

The tau momentum resolution is shown in figure 3.14 as a function of the true $\tau_{\text{had-vis}}$ momentum for both one-prong and multi-prong objects. The resolution is obtained by dividing the mean value from a Gaussian fit by the mean value of the true visible momentum. It ranges between values of 20% at low momenta to values of 5–10% at higher momenta. One-prong tau candidates generally have a better

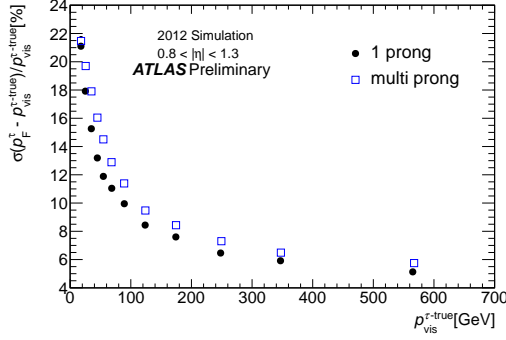


Figure 3.14: Momentum resolution for one-prong and multi-prong τ_{had} objects in the pseudo-rapidity range $0.8 < |\eta^{\tau}| < 1.3$ [126].

momentum resolution due to the higher fraction of electromagnetic energy deposits from π^0 decays compared to the hadronic energy deposits.

3.7 Missing Transverse Energy

In hadron collisions the initial momenta of the partons participating in the hard interaction are unknown due to the composite nature of the hadrons, and hence the final state momentum along the beam axis also cannot be predicted. However, since the protons in the LHC ring have a negligible transverse momentum the initial state can also be assumed to have a vanishing transverse momentum. This is exploited by defining an observable called the *missing transverse energy*, $\mathbf{E}_T^{\text{miss}}$, given by the momentum vector corresponding to the transverse momentum that is needed to counteract the observed transverse momentum of the final state in order to achieve a total transverse momentum of zero. A large $\mathbf{E}_T^{\text{miss}}$ vector indicates the production of particles that are not reconstructed in the detector, e.g. neutrinos that interact weakly with the detector material. $\mathbf{E}_T^{\text{miss}}$ is an important tool in $H \rightarrow \tau\tau$ analyses due to the large fraction of the momentum from tau decays carried by neutrinos. It offers additional background suppression and enables the reconstruction of the ditau mass.

The ATLAS $\mathbf{E}_T^{\text{miss}}$ calculation uses reconstructed and calibrated physics objects as described in the previous sections. Calorimeter energy deposits are associated with physics objects and added together in the x and y direction separately as follows [127]:

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss},e} + E_{x(y)}^{\text{miss},\gamma} + E_{x(y)}^{\text{miss},\tau} + E_{x(y)}^{\text{miss,jets}} + E_{x(y)}^{\text{miss},\mu} + E_{x(y)}^{\text{miss,SoftTerm}}, \quad (3.9)$$

where each term is the negative sum of the energy of the reconstructed objects corresponding to electrons, photons, hadronic tau decays, jets and muons. In the following the transverse energy will be treated as a scalar value E_T^{miss} . Any energy contributions not matched to one of the mentioned physics objects generally originate from soft particles and are grouped into an additional soft term named $E_T^{\text{miss,SoftTerm}}$. Electrons are calibrated to the electron energy scale as described in section 3.4.2, photons are reconstructed at the EM scale, while hadronic tau decays are corrected to the TES discussed in section 3.6.2. Jets clustered with the anti- k_T algorithm with $R = 0.4$ and $p_T > 20$ GeV are included with the pile-up corrections detailed in section 3.3.1 applied. Muons of both the combined and segment-tagged variety (see section 3.5.2) are included in the muon term to also include low momentum muons. Special care needs to be taken with respect to soft particles since the default ATLAS object reconstruction algorithms are optimised on particles with higher energy. Such low-momentum particles can reach relatively high multiplicities in high-energy hadron collisions. To increase the accuracy of the E_T^{miss} estimate, a separate $E_T^{\text{miss,SoftTerm}}$ term is therefore added to account for these particles through the inclusion of further energy deposits and tracks. Only topological clusters are considered in order to reduce the impact of noise, while tracks are added only in cases where they are matched to a physics object or a topological cluster.

The resolution of the reconstructed E_T^{miss} is impacted by pile-up, especially the jet and soft terms. Jets with $p_T < 50$ GeV are therefore required to have a jet vertex fraction $|\text{JVF}| > 0$ (see section 3.3.1). A similar variable called the soft term vertex fraction [128] is defined as the momentum fraction of tracks matched to the primary vertex in the event and is used to scale the $E_T^{\text{miss,SoftTerm}}$ to correct for pile-up effects. The pile-up corrections are shown in figure 3.15 in terms of their impact on two different variables. Figure 3.15 (a) shows the mean pull value between the reconstructed and true E_T^{miss} , also called the linearity, in bins of the true transverse missing energy. Due to resolution effects and the fact that E_T^{miss} is always defined to be positive, a bias toward positive values is seen at low $E_T^{\text{miss,True}}$. The pile-up correction improves the linearity over a wide range of $E_T^{\text{miss,True}}$ with a slight overcorrection at high values. Figure 3.15 (b) shows the E_T^{miss} resolution as a function of the scalar sum of all transverse energy contributions in the event ($\sum E_T$). The red markers indicate the resolution after the soft term correction,

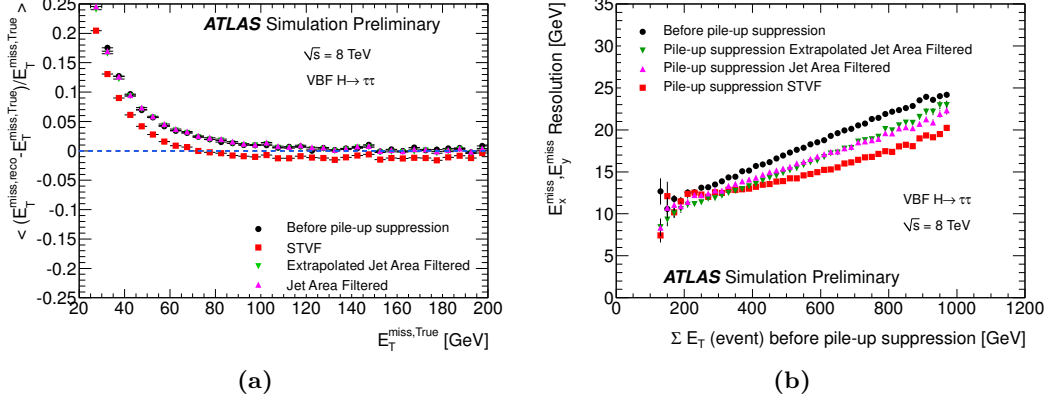


Figure 3.15: E_T^{miss} linearity as a function of $E_T^{\text{miss, True}}$ (a) and E_T^{miss} resolution as a function of ΣE_T (b) in VBF $H \rightarrow \tau\tau$ events at various stages of the pile-up suppression, including the correction applied on the soft term (red) [127].

which significantly improves the resolution in events with high ΣE_T despite the smaller relative role of the soft term in VBF events due to the topology including neutrinos and two high-energy jets.

In addition to systematic uncertainties on energy scales and resolutions of the established physics object types that all contribute to the uncertainty on E_T^{miss} it is also necessary to estimate the systematic uncertainty on the soft term. The resolution and energy scale was studied in $Z \rightarrow \mu^+\mu^-$ events without jets. In such events only the muon term and the soft term have significant contributions. From these studies, including the STVF pile-up correction on the soft term, systematic uncertainties were found to be 7.9% on the scale and 4.8% on the resolution of the soft term [127].

Summary

The above sections have discussed the reconstruction of all the kinds of objects necessary in an analysis of $H \rightarrow \tau\tau$ events with ATLAS. Tracks and vertices can be combined with calorimeter deposits to reconstruct electrons, muons, taus and jets of particles as well as missing transverse energy due to e.g. neutrinos. The calibration of these objects and comparisons between simulation and data have also been discussed. The next step is to define the analysis in terms of the signal and background processes of interest and how the reconstructed objects are used to identify their corresponding experimental signatures.

Chapter 4

$H \rightarrow \tau\tau$ Analysis Method

The high centre-of-mass energy and instantaneous luminosity of the LHC have enabled the discovery of the Higgs boson in data collected by the ATLAS and CMS experiments. The 2012 discovery was made in the bosonic decay channels, while the fermionic couplings are still to be discovered. At the experimentally determined Higgs boson mass of 125 GeV the SM decay branching ratio into a pair of b quarks dominates at 58% but has a complicated experimental signature. The second largest branching ratio comes from decays into W bosons at 22%, and decays into tau leptons constitute approximately 6.3% of the total decay width. The 2012 $H \rightarrow \tau\tau$ couplings analysis [2] was able to produce evidence for the coupling with a signal significance of 4.5σ , strongly indicating the presence of $H \rightarrow \tau\tau$ decays. The present analysis exploits the collected data sample of events with $H \rightarrow \tau\tau$ decays to attempt an identification of BSM Higgs couplings to vector bosons, which can be done by studying the vector boson fusion (VBF) production mode. This analysis therefore only counts $H \rightarrow \tau\tau$ events produced through VBF as signal events, while the remaining production modes such as gluon fusion (ggF) are defined as background. Although the associated VH production mode also includes a Higgs boson coupling to a vector boson, the production cross section is too low to warrant a separate analysis category.

This chapter outlines the signal and background processes involved in the analysis, how these are modelled, and the kinematic variables employed in the event selection and categorisation. All processes involved are presented in section 4.1, while the utilised MC samples are listed and described in section 4.2. The signal re-weighting to include CP-odd coupling terms is presented in section 4.3. The experimentally collected data samples included in the analysis and the triggers used to collect them are given in section 4.4 along with a description of the methods for data-driven background estimation. An algorithm called the missing mass calculator (MMC) provides an improved estimate of the mass of the full di-tau decay and is described in section 4.1.3. The event selection is detailed in section 4.6 including

a description of the discriminating variables included in the boosted decision tree (BDT) classification used to discriminate signal from background events.

4.1 Experimental Signatures of Signal and Background

The analysis studies di-tau events with both taus decaying leptonically ($\tau_\ell\tau_\ell$) as well as events with one tau decaying leptonically and the other hadronically ($\tau_\ell\tau_{\text{had}}$). The experimental signature of a leptonic tau decay consists of one reconstructed electron or muon plus missing momentum from two invisible neutrinos. A hadronic tau decay consists of a reconstructed τ_{had} object as well as missing momentum from one neutrino. The $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ decay channel signatures will then be the appropriate combination of two of these. In $\tau_\ell\tau_{\text{had}}$ the light lepton and the τ_{had} object carry opposite electric charge. The E_T^{miss} from the neutrinos can be used to extract further event information. Additional jets can be present in the event depending on the specific Higgs boson production mode. In VBF production two high- p_T jets are expected and incorporated in the event selection criteria described in section 4.6. Background estimation is performed in different ways depending on the type of background. Major background contributions are determined using data-driven methods, while other contributions are simulated. A procedure called the Fake Factor method is used to estimate the fake tau background, which is based on deriving a relation between the number of events that fail or pass the τ identification. The $Z \rightarrow \tau\tau$ background estimation employs a hybrid procedure that embeds simulated tau decays in real data events which have reconstructed $Z \rightarrow \mu\mu$ candidates to obtain an accurate description of the underlying event and the additional jet modelling.

4.1.1 Higgs Boson Processes

The analysis presented in this thesis studies the Higgs boson production modes described in section 1.2.1. The production of Higgs bosons in association with a pair of top quarks is considered to be negligible and is excluded from the analysis. The only production mode that is considered as signal is VBF production, since the investigated Higgs boson couplings are those to vector bosons. It is a purely electroweak process leading to relatively low QCD corrections at higher orders. Two jets are present at tree-level stemming from the two outgoing partons seen in figure 1.7 (b). They generally have high transverse momentum and have a high probability of being well-separated in pseudo-rapidity with little hadronic activity in the region between them due to the absence of colour exchange between the jets, leading to a distinct

event topology that is very different from that of the major background processes. Preliminary signal-enriched phase space regions are therefore relatively easy to define through simple kinematic selection criteria on e.g. jet p_T and $\Delta\eta_{jj}$, i.e. the separation between the two jets in pseudo-rapidity.

The ggF process has no additional partons in the final state at tree-level. It does, however, involve an initial state with coloured gluons and a heavy quark loop leading to large contributions from higher-order QCD diagrams that can include final state partons that subsequently hadronise into jets. Due to the recoil of the Higgs boson against the jet system the Higgs boson will be boosted with respect to the centre-of-mass frame. This can potentially lead to event kinematics similar to those of a VBF event with two or more high- p_T jets and a boosted Higgs boson. Therefore a sizeable amount of ggF events, roughly 50% of the VBF signal yield, are expected to leak into the VBF selection category.

The VH process leads to a variety of final state topologies depending on the decay of the vector boson. The primary contribution to the VBF selection category is expected to come from VH events where the vector boson decays hadronically to form one or more jets that match the VBF selection criteria.

A non-negligible amount of Higgs production events with $H \rightarrow WW$ decays enter the VBF selection category in the $\tau\ell\tau\ell$ decay channel in the case where both W bosons decay leptonically. The event contribution is of the order of 10% of the VBF signal yield. These processes are also included in the event simulation.

4.1.2 Background Processes

The most significant backgrounds in the analysis stem from processes with the same or similar final states as the $H \rightarrow \tau\tau$ signal including two real tau leptons. Processes with the same final state are called *irreducible*, although event kinematics including the reconstructed mass of the intermediate resonance may still provide some level of discrimination power. Other processes have the possibility to fake light leptons or τ_{had} objects through the misidentification of other types of particles and constitute a non-negligible background contribution.

The main irreducible background consists of Z boson or off-shell photon decays to two tau leptons in association with jets, denoted $Z/\gamma^* \rightarrow \tau\tau + \text{jets}$. The tree-level Feynman diagrams of this production process with up to one additional parton are shown in figure 4.1. Although the final state objects are the same as for the signal process, the di-tau invariant mass provides some discrimination power as long as the resolution of the mass estimate is sufficiently high compared to the mass difference between the Z and Higgs bosons. This is mainly limited by the

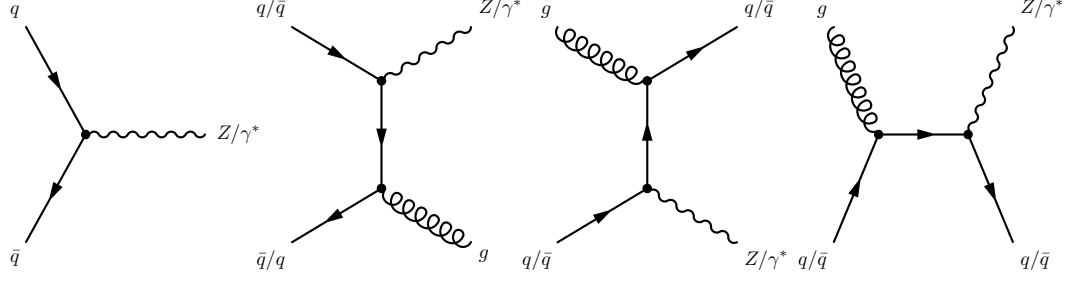


Figure 4.1: Tree-level Feynman diagrams with up to one additional parton showing processes leading to $Z/\gamma^* + \text{jets}$ events.

E_T^{miss} resolution and the choice of algorithm used in the mass reconstruction, some of which are discussed in more detail in section 4.1.3. An additional source of discrimination power can be accessed by studying the kinematics of the jets. Various selection criteria on the jets can be tweaked to further suppress this background by comparing $Z/\gamma^* \rightarrow \tau\tau + \text{jets}$ events with VBF Higgs boson events in terms of distributions of variables representing jet properties. Two example variables are shown in figure 4.2, namely the invariant mass and the p_T of the di-tau system at truth level for VBF signal and $Z/\gamma^* \rightarrow \tau\tau + \text{jets}$ background. Both variables are calculated from the full true kinematics of the two tau leptons. The distributions are shown without any selection criteria applied. Both variables can be seen to have significant discrimination power at the truth level, which will however be reduced when analysing actual data due to the invisible neutrinos.

Events with misidentified objects that pass the selection criteria are collectively called the fake background and are most prevalent in $\tau_\ell\tau_{\text{had}}$. In most events of this type a jet fakes a τ_{had} signature and has the possibility of being included in the channel if a light lepton is also reconstructed. This background is dominated by $W + \text{jets}$ and QCD multijet processes with smaller contributions from diboson and top backgrounds. Event rates for the latter two types of processes will be quoted separately in section 4.6 and are taken from calculated cross sections. In $W + \text{jets}$ events where the W boson decays leptonically a neutrino is also present creating missing transverse energy, giving such events a very similar signature to the signal. The amount of background of this type that passes the selection depends mainly on the τ_{had} misidentification rate. Discrimination against $W + \text{jets}$ events in $\tau_\ell\tau_{\text{had}}$ would ideally be done by reconstructing the invariant mass of the leptonic decay products. The presence of a neutrino prevents a full mass reconstruction, but by taking into account the missing transverse energy in the event it is possible to

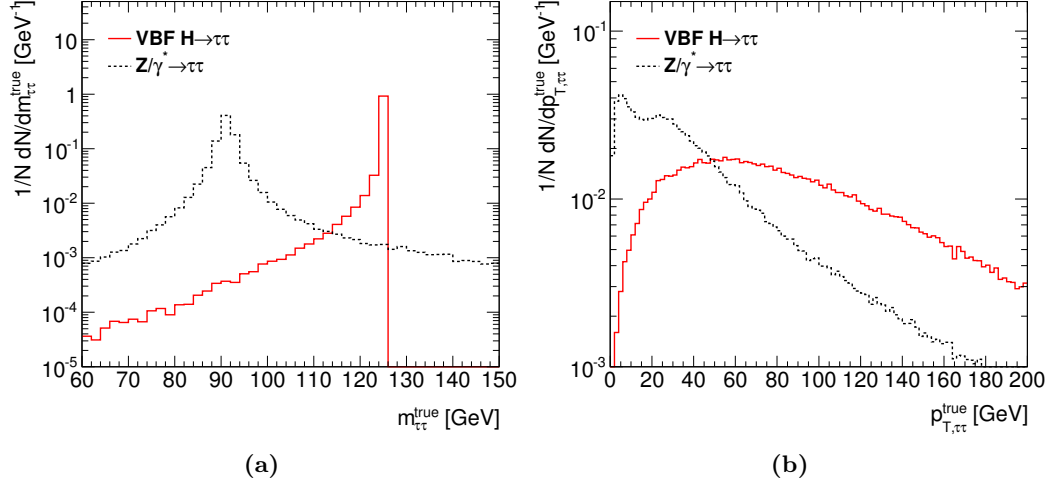


Figure 4.2: Distributions of variables with potential signal discrimination power: (a) the invariant mass and (b) the transverse momentum of the ditau system in VBF $H \rightarrow \tau\tau$ and $Z/\gamma^* \rightarrow \tau\tau$ events. The variables are calculated at truth level using the four-momenta of the tau leptons.

construct a slightly different variable called the transverse mass:

$$m_T = \sqrt{2p_T^l E_T^{\text{miss}} (1 - \cos \Delta\phi_{l, \mathbf{E}_T^{\text{miss}}})}, \quad (4.1)$$

which rises to its maximum close to the W mass in $W + \text{jets}$ events. Further suppression of these events can be extracted by studying the angles between the objects in the event. The angular separations between the reconstructed $\tau_{\text{had-vis}}$, light lepton and E_T^{miss} in a resonant Higgs decay will differ significantly from those in a background event where the objects do not originate from the same parent. Especially in signal events with one or more high- p_T jets, i.e. with a highly boosted Higgs boson, the angle between the $\tau_{\text{had-vis}}$ and light lepton will be small compared to that in a typical $W + \text{jets}$ event. The orientation of the E_T^{miss} will also differ between signal and background. In a $H \rightarrow \tau_l \tau_{\text{had}}$ event both the leptonic and hadronic tau decay will result in one or more neutrinos with a tendency to have the total E_T^{miss} pointing in a direction between the taus, while a $W + \text{jets}$ event will typically not exhibit the same behaviour. As already mentioned, the other major contribution to the fake background comes from QCD multijet events which have a large cross section in proton-proton collisions. Two jets each have to be misidentified as either a τ_{had} object or a light lepton in accordance with the studied decay channel to fake a signal event. Tight lepton selection and identification criteria will aid in

the reduction of this background. In $\tau_\ell\tau_{\text{had}}$ a data-driven background estimation method based on fake factors is used specifically for events where a jet fakes a τ_{had} signature. More details are given in section 4.5.2.

$Z/\gamma^* \rightarrow ll + \text{jets}$ events also contribute to the background. The event rate is highest in $\tau_\ell\tau_\ell$ where the basic signature of jets plus two leptons and possibly some amount of E_T^{miss} is present without any particles having to be misidentified. In $\tau_\ell\tau_{\text{had}}$ a lepton needs to be misidentified as a hadronic tau decay, or one lepton must fail reconstruction while a jet is misidentified as a τ_{had} object. Due to the very low lepton misidentification rates in ATLAS the largest contribution comes from the latter type of events. Rejection of $Z/\gamma^* \rightarrow ll + \text{jets}$ therefore mainly depends on the light lepton acceptance. Electroweak pair production of vector bosons W and Z , also called diboson events, is also a source of final states with one or more τ leptons or a τ together with an electron or muon. Diboson processes are grouped together with $Z/\gamma^* \rightarrow ll + \text{jets}$ events in the presentation of event yields later in this chapter.

Events involving $\bar{t}t$ pair or single top production have a high probability of containing top decays to a bottom quark and a W boson since this branching fraction dominates. This results in a signature containing multiple jets as well as E_T^{miss} from the subsequent W decay. Depending on the W decay, the top processes include events with both real and fake τ_{had} decays at roughly equal rates. An efficient tool for reducing these backgrounds is to reject events with identified b -jets through b -tagging.

4.1.3 Reconstruction of the Ditau Mass

Information about the Higgs boson kinematics can be extracted by adding together the four-vectors of the two reconstructed tau leptons. This can be used to distinguish resonant ditau decays from non-resonant processes. The ability to distinguish between several different resonances decaying to the same final state is highly dependent on the achievable mass resolution. The neutrinos in the event complicate the mass reconstruction and reduce the mass resolution by preventing a full Higgs boson four-momentum reconstruction. The missing transverse energy in the event can however be used to approximate the remaining momentum contributions not included in the visible decay products. Several mass reconstruction methods with different approaches to incorporating the missing transverse momentum are described below.

The Visible Mass

The visible mass m_{vis} is reconstructed as the invariant mass of only the visible decay products, thereby ignoring the E_T^{miss} produced by the neutrinos. Denoting the tau leptons as 1 and 2, it can simply be calculated from the combined four-momentum as

$$m_{\text{vis}} = \sqrt{(E_1 + E_2)^2 - (\mathbf{p}_1 + \mathbf{p}_2)^2}. \quad (4.2)$$

The momentum carried by the neutrinos is generally a significant part of the full tau lepton momenta, leading to mass spectra with a bias towards lower values compared to the mass of the reconstructed resonance. The separation between the Higgs and Z boson mass peaks is also reduced. The visible mass is not used as a discriminating variable in the analysis but is part of the validation of the $Z \rightarrow \tau\tau$ background model.

The Collinear Approximation

In the limit of a highly boosted resonant state in the lab frame, the neutrino momentum from the tau decays can be approximated as pointing in the same direction as that of the remaining decay products. With the additional assumption that all the missing transverse energy comes from the neutrinos, the collinear mass can be written as [82]

$$m_{\text{coll}} = \frac{m_{\text{vis}}}{\sqrt{x_1 x_2}}, \quad (4.3)$$

where $x_{1,2}$ are the momentum fractions carried away by the visible tau decay products and can be calculated as

$$x_{1,2} = \frac{p_{\text{vis},2}}{p_{\text{vis},2} + p_{\text{mis},2}}. \quad (4.4)$$

The total invisible momentum carried away by the neutrinos in each tau decay, $p_{\text{mis},2}$, can be found by solving two equations with two unknowns:

$$\begin{aligned} E_T^{\text{miss},x} &= p_{\text{mis}1} \sin\theta_{\text{vis}1} \cos\phi_{\text{vis}1} + p_{\text{mis}2} \sin\theta_{\text{vis}2} \cos\phi_{\text{vis}2}, \\ E_T^{\text{miss},y} &= p_{\text{mis}1} \sin\theta_{\text{vis}1} \sin\phi_{\text{vis}1} + p_{\text{mis}2} \sin\theta_{\text{vis}2} \sin\phi_{\text{vis}2}. \end{aligned} \quad (4.5)$$

A special case is when the tau decay products are back-to-back in the x, y plane, leading to equations 4.5 becoming degenerate. A large amount of $H \rightarrow \tau\tau$ events have such a back-to-back topology rendering the collinear approximation unsuitable for this type of analysis.

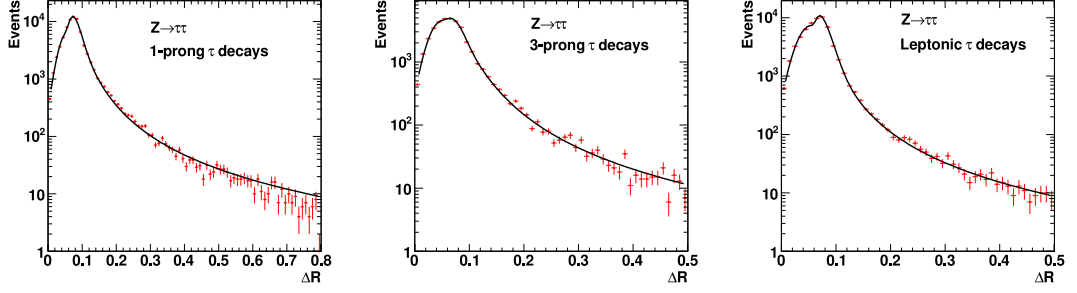


Figure 4.3: Probability distribution functions $\mathcal{P}(\Delta R, p_\tau)$ in a particular p_τ bin of the decaying tau lepton. Functions are shown for one-prong (left) and three-prong (middle) hadronic tau decays as well as leptonic tau decays. The functions are used as part of the maximum-likelihood scan to determine the optimal parameters for the kinematics of the invisible decay products [82].

The Missing Mass Calculator

Improvements to the ditau invariant mass estimate can be attained by considering the more general case of a non-negligible opening angle between the neutrinos and the visible decay products while exploiting known kinematic properties of tau decays. An algorithm called the missing mass calculator (MMC) [82] has been developed for this purpose. Depending on the decay channel of the ditau system (hadronic, semi-leptonic or leptonic) there are between six and eight unknown variables describing the spatial coordinates of the invisible momentum carried away by the neutrinos and the invariant mass of the neutrinos specifically in leptonic tau decays. However, only 4 independent equations can be constructed containing these variables:

$$\begin{aligned}
 E_T^{\text{miss},x} &= p_{\text{mis}_1} \sin\theta_{\text{mis}_1} \cos\phi_{\text{mis}_1} + p_{\text{mis}_2} \sin\theta_{\text{mis}_2} \cos\phi_{\text{mis}_2}, \\
 E_T^{\text{miss},y} &= p_{\text{mis}_1} \sin\theta_{\text{mis}_1} \sin\phi_{\text{mis}_1} + p_{\text{mis}_2} \sin\theta_{\text{mis}_2} \sin\phi_{\text{mis}_2}, \\
 M_{\tau_1}^2 &= m_{\text{mis}_1}^2 + m_{\text{vis}_1}^2 + 2\sqrt{p_{\text{vis}_1}^2 + m_{\text{vis}_1}^2} \sqrt{p_{\text{mis}_1}^2 + m_{\text{mis}_1}^2}, \\
 M_{\tau_2}^2 &= m_{\text{mis}_2}^2 + m_{\text{vis}_2}^2 + 2\sqrt{p_{\text{vis}_2}^2 + m_{\text{vis}_2}^2} \sqrt{p_{\text{mis}_2}^2 + m_{\text{mis}_2}^2}.
 \end{aligned} \tag{4.6}$$

An exact solution cannot be found since the number of unknowns exceeds the number of constraining equations. Instead, the remaining allowed phase space can be studied in terms of the likelihood of different kinematic configurations occurring. As an example, probability density functions are shown in figure 4.3 as a function of the distance between the visible and invisible decay products $\Delta R = \sqrt{(\eta_{\text{vis}} - \eta_{\text{mis}})^2 + (\phi_{\text{vis}} - \phi_{\text{mis}})^2}$ for leptonic as well as one-prong and three-prong hadronic tau decays. The distributions have been obtained from simulated samples of tau decays from $Z/\gamma^* \rightarrow \tau\tau$ events and depend on the momentum of the

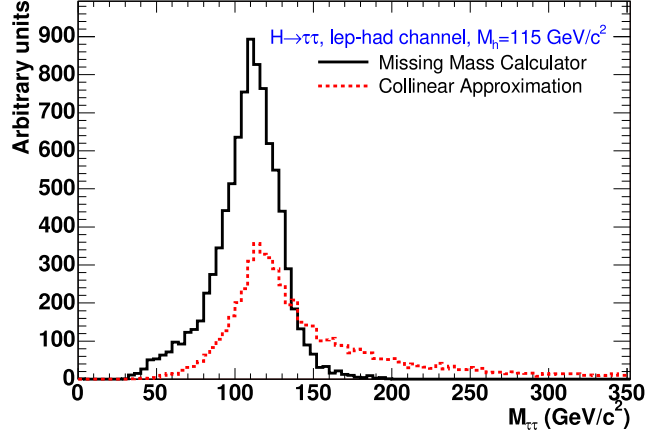


Figure 4.4: Reconstructed ditau mass $M_{\tau\tau}$ for simulated ggF $H \rightarrow \tau_l \tau_{had}$ events with a Higgs boson mass of $m_H = 115$ GeV including detector resolution effects. The missing mass calculator (solid line) is compared to the collinear approximation (dashed line). The MMC algorithm increases the mass resolution and suppresses the long upward tail observed in the distribution obtained from the collinear approximation, mostly consisting of events with an approximately back-to-back topology where the algorithm is still able to produce a valid result. The higher efficiency of the MMC algorithm also leads to an increased event yield [82].

initial tau lepton. This information can be incorporated as an additional constraint by defining a log-likelihood in terms of the probabilities of the particular configurations of the two tau decays. In the case where both taus decay hadronically there are only six unknowns and a log-likelihood function containing information about ΔR and the tau momenta is enough to constrain the system. In cases where one or both taus decay leptonically, the invariant masses of the neutrinos need to be included in the likelihood as well. The E_T^{miss} resolution is also taken into account by including the transverse components $E_T^{miss,x}$ and $E_T^{miss,y}$ in the likelihood scan by defining a Gaussian probability function according to the experimental resolution. The log-likelihood function for hadronic tau decays can then be written as

$$\mathcal{L} = -\log(\mathcal{P}(\Delta R_1, p_{\tau 1}) \times \mathcal{P}(\Delta R_2, p_{\tau 2}) \times \mathcal{P}(\Delta E_T^{miss,x}) \times \mathcal{P}(\Delta E_T^{miss,y})), \quad (4.7)$$

where the probability functions $\mathcal{P}(\Delta R_{1,2}, p_{\tau 1,2})$ depend on the decay types and $\Delta E_T^{miss,x,y}$ are the variations of the x - and y -components of the missing transverse energy compared to the experimentally determined value.

The MMC algorithm allows a solution to be found in a much higher fraction of events than using the collinear approximation, thereby significantly increasing the efficiency. Additionally, the resolution of the invariant mass is improved by

considering the full phase space of possible neutrino opening angles and E_T^{miss} values including their probabilities. Figure 4.4 shows distributions of $M_{\tau\tau}$ in $\tau_\ell\tau_{\text{had}}$ events calculated using the collinear approximation and the MMC algorithm respectively. Additional kinematic selection criteria have been applied to ensure a well-defined output of each algorithm. The analysis presented in this thesis determines the ditau mass $m_{\tau\tau}$ using the MMC algorithm.

4.2 Monte Carlo Simulated Samples

The theoretically predicted distributions of signal and background are produced through the use of numerous Monte Carlo event generators. While major background contributions such as $Z \rightarrow \tau\tau$ and fake taus are modelled using data-driven techniques, their development and validation still depend on simulation. Smaller background contributions are estimated directly from the simulated samples. The simulation of VBF signal is a hybrid procedure with Monte Carlo generation of an SM sample and re-weighting of events from this sample to describe non-zero CP-odd mixing parameter values. This section will describe the MC generation, while section 4.3 details the re-weighting technique. All Higgs samples are produced with a Higgs mass of $m_H = 125$ GeV. Table 4.1 shows the complete list of MC simulated samples used in the analysis as well as the perturbative order of the QCD calculation used for each sample.

4.2.1 Signal Samples

Although several Higgs production modes are included in the analysis, only VBF $H \rightarrow \tau\tau$ and $H \rightarrow WW$ production modes are considered to be signal, since the goal is to test CP-invariance in VBF. Samples of 2×10^6 parton-level VBF events are generated for both decay modes using POWHEG [152]. Simulation of the underlying event, parton showering and hadronisation is subsequently performed in PYTHIA 8 [102] at next-to-leading order in QCD and using the CT10 PDF set [88]. The event selection and background discrimination procedure exploit a large number of kinematic variables, making it important to have high precision modelling of the signal events. An electroweak correction of the Higgs transverse momentum p_t^H distribution is introduced by comparing the output of PYTHIA to the distribution obtained from HAWK [79–81], which takes the complete NLO electroweak contributions into account. The correction size depends on p_T^H , increasing from 1-2% at low values to approximately 20% at $p_T^H = 300$ GeV [55].

Signal	MC generator	$\sigma \times \mathcal{B}$ [pb]	Order	
VBF, $H \rightarrow \tau\tau$	POWHEG+PYTHIA8	0.100	(N)NLO	[56, 129–131]
VBF, $H \rightarrow WW$	Same as VBF $H \rightarrow \tau\tau$	0.34	(N)NLO	[56, 129–131]
Background				
ggF, $H \rightarrow \tau\tau$	MINLO+PYTHIA8	1.22	NNLO+NNLL	[56, 132–137]
ggF, $H \rightarrow WW$	POWHEG+PYTHIA8	4.16	NNLO+NNLL	[56, 132–137]
WH , $H \rightarrow \tau\tau$	PYTHIA8	0.0445	NNLO	[56, 138]
ZH , $H \rightarrow \tau\tau$	PYTHIA8	0.0262	NNLO	[56, 138]
$W(\rightarrow l\nu)$, ($l = e, \mu, \tau$)	ALPGEN+PYTHIA8	36800	NNLO	[139, 140]
$Z/\gamma^*(\rightarrow ll)$, $10 \text{ GeV} < m_{ll} < 60 \text{ GeV}$	ALPGEN+HERWIG	13000	NNLO	[139, 140]
$Z/\gamma^*(\rightarrow ll)$, $60 \text{ GeV} < m_{ll} < 2 \text{ TeV}$	ALPGEN+PYTHIA8	3910	NNLO	[139, 140]
VBF, $Z/\gamma^*(\rightarrow ll)$	SHERPA	1.1	LO	[141]
$t\bar{t}$	POWHEG+PYTHIA8	253	NNLO+NNLL	[142–146]
Single top: Wt	POWHEG+PYTHIA8	22	NNLO	[147]
Single top: s -channel	POWHEG+PYTHIA8	5.6	NNLO	[148]
Single top: t -channel	ACERMC+PYTHIA6	87.8	NNLO	[149]
$q\bar{q} \rightarrow WW$	ALPGEN+HERWIG	54	NLO	[150]
$gg \rightarrow WW$	GG2WW+HERWIG	1.4	NLO	[151]
WZ, ZZ	HERWIG	30	NLO	[150]

Table 4.1: Signal and background samples used in the analysis and the Monte Carlo generators used to model them. All Higgs samples are generated with mass $m_H = 125 \text{ GeV}$. The cross sections times branching fractions ($\sigma \times \mathcal{B}$) are quoted for $\sqrt{s} = 8 \text{ TeV}$ alongside the perturbative order of the QCD calculation. The signal processes include the SM $H \rightarrow \tau\tau$ and $H \rightarrow WW$ branching fractions, and the W and Z/γ^* backgrounds include the leptonic decay branching fractions. Inclusive cross sections are quoted for all other backgrounds.

4.2.2 Background Samples

Background samples include all Higgs production modes except VBF in addition to all non-Higgs processes. For gluon fusion Higgs production, ggF, generated with standard POWHEG (the generator for this process in the couplings analysis) a large discrepancy is observed between Bjorken x values of the initial state partons at generator level and when calculated at reconstruction level. This ggF POWHEG sample is only NLO for 0 jets, which means that only one parton comes from the hard interaction, while any additional jets will originate from the parton shower. This makes the sample unsuitable for calculating the Optimal Observable, since it makes use of both the leading and sub-leading jet in the event. In the simulated VBF signal events these jets both need to originate from the hard interaction in order to have an accurate value for the matrix element used as input to the Optimal Observable calculation. In this analysis the Higgs-plus-one-jet process is simulated at NLO accuracy in QCD with POWHEG using the MINLO feature [153]. The POWHEG

event generator is interfaced to PYTHIA8, and the CT10 PDF set is used. Associated VH production is simulated using PYTHIA8 using the CTEQ6L1 [89] PDF set. Contributions from associated $t\bar{t}H$ production were previously evaluated in the $H \rightarrow \tau\tau$ couplings analysis [2] to be negligible and are not included. As with VBF production, the background Higgs production modes are simulated with $H \rightarrow \tau\tau$ and $H \rightarrow WW$ decays since the contribution from $H \rightarrow WW$ in $\tau_\ell\tau_\ell$ is potentially non-negligible in the signal region of the analysis.

Other background samples are generated using various event generators interfaced to either PYTHIA or HERWIG [103] to simulate the underlying event, parton shower and hadronisation. For the HERWIG samples the tau lepton decays are simulated using TAUOLA [154]. Photon radiation from charged leptons in these samples is calculated by PHOTOS [155]. Samples containing $Z/\gamma^* + \text{jets}$ and $W + \text{jets}$ events are generated with ALPGEN [156] by using the LO matrix elements for W and Z production including a maximum of five additional partons in the final state. It uses a matching scheme [157] between the matrix element and the parton shower algorithm. $Z/\gamma^* + \text{jets}$ events are generated in two separate intervals of the true dilepton mass with $m_{ll}^{\text{true}} \leq 40$ GeV. This was done to avoid generating excessive events with a dilepton mass below the event selection requirements of the analysis. The low-mass samples are interfaced to HERWIG and JIMMY [158], while high-mass samples are interfaced to PYTHIA. Since $Z/\gamma^* + \text{jets}$ events constitute a considerable fraction of the background in signal enriched kinematic regions, a high statistical power is desirable when optimising methods of background rejection. Therefore, additional VBF-enriched samples were generated by applying a VBF-like kinematic filter at generator level before detector simulation that allows the simulation of large samples of events in the kinematic region sensitive to signal.

Event samples with top quarks are generated separately depending on the process and channel. A $t\bar{t}$ sample is generated using POWHEG [159,160] interfaced to PYTHIA using the CTEQ6L1 PDF set. The s-channel and Wt processes in single-top events are likewise generated with POWHEG interfaced to PYTHIA, while the t-channel processes are generated using ACERMC [161,162] interfaced to PYTHIA. In the $\tau_\ell\tau_{\text{had}}$ decay channel these samples are only used to model the part of top background where the τ_{had} object originates from a real τ_{had} decay or a misidentified light lepton. Events where jets are faking hadronic tau decays are instead modelled by the data-driven fake factor method described in section 4.5.2.

Production of diboson events is simulated with HERWIG in the case of ZZ and WZ events, whereas WW events are simulated with ALPGEN interfaced to HERWIG. The loop-induced $gg \rightarrow WW$ process is generated using the GG2WW [163] program

interfaced to HERWIG and JIMMY to model non-perturbative QCD effects. ALPGEN has a superior description of the jet topology in diboson events, but off-shell Z contributions are not included in ZZ and WZ leading instead to the use of HERWIG in these cases.

4.3 Signal Re-weighting

Methods are used to re-weight individual events in an SM signal sample to produce a sample of a given CP-odd coupling strength, instead of generating individual signal samples with the required number of events for each coupling strength. This allows convenient access to arbitrary coupling strength values without the need to decide on these in advance. In order to simulate any degree of CP-mixing, a matrix element-based re-weighting procedure is applied to the existing POWHEG+PYTHIA8 SM VBF signal sample. The re-weighting takes as input truth-level information for each event, more specifically the Bjorken x values of the incoming partons, the four-vectors of the outgoing Higgs boson¹ and of the final-state partons (before any hadronisation), as well as the flavour of the involved partons. Using these, the weights are obtained as the ratio of the matrix element squared evaluated for the CP-mixed case one wishes to reweight to, and the SM matrix element squared. All of the matrix element calculations for the re-weighting are performed using code extracted from HAWK. Version 2.0 is the only version that includes anomalous HVV couplings. For the CP-mixed scenarios the parameter \tilde{d}_B is set equal to \tilde{d} (see eqs. 1.53).

POWHEG includes matching of matrix elements and parton shower at NLO, meaning that there can be three different kinds of events (with q and \bar{q} interchangeable): $qq \rightarrow qqH$, $qg \rightarrow qq\bar{q}H$ and $qg \rightarrow qgqH$. The re-weighting uses the corresponding matrix element at LO from HAWK for the $2 \rightarrow 2+H$ or $2 \rightarrow 3+H$ process, taking into account the flavours of incoming and outgoing partons. This procedure is expected to give a very good approximation to a real and full NLO re-weighting.

4.3.1 Validation of the Re-weighting Procedure

The validation of the re-weighting procedure is achieved by comparing a sample of SM simulated events after the re-weighting, and a sample of events directly generated assuming the same amount of CP-mixing. The Monte Carlo generators considered in this study are VBFNLO and MG5_AMC@NLO [68]. Both the HAWK routines and VBFNLO implement the same parametrisation of the effective lagrangian (see

¹The Higgs boson with status 22 is used, i.e. before any radiation etc.

VBF@NLO parameter	Description
PARAMETR2 = true	Parametrisation of the L3-Collaboration
D_EVEN = 0.0	d , CP-even
DB_EVEN = 0.0	d_B
DG1Z_EVEN = 0.0	$\Delta g1_Z$
DKGAM_EVEN = 0.0	$\Delta \kappa_{\gamma}$
D_ODD = 0.1	\tilde{d} , CP-odd
DB_ODD = 0.1	\tilde{d}_B
KGAM_ODD = 0.0	κ_{γ}
HVV1 = 4	All anomalous couplings activated
TREEFACW = 1.0	SM HWW tensor factor ($\sin(\alpha\text{-}\beta)$ in MSSM)
TREEFACZ = 1.0	SM HZZ tensor factor ($\sin(\alpha\text{-}\beta)$ in MSSM)
LOOPFAC = 0.0	SM loop factor multiplying $HZ\gamma$ and $H\gamma\gamma$

Table 4.2: Parameters set in the VBFNLO input file *anom_HVV.dat* corresponding to $\tilde{d} = \tilde{d}_B = 0.1$.

VBF-like cuts
$p_T(p) > 25\text{GeV}$
$ \eta(p) < 4.5$
$ \Delta\eta(p_1, p_2) > 2.8$
$ \Delta R(p_1, p_2) > 0.4$
$M(p_1 p_2) > 500\text{GeV}$

Table 4.3: List of minimal selection requirements applied to all the comparisons discussed in this section. The cuts are applied to parton-level quantities for both LO and NLO comparisons.

equation 1.44), hence it is easier to make sure that the re-weighted sample and the one generated with VBFNLO correspond to the same BSM model. Table 4.2 shows the value of the parameter used to simulate events with VBFNLO.

VBFNLO can only generate events at leading order (and differential distributions at next-to-leading order) that are then stored in standard Les Houches files (LHE) [164]. The four-vectors of the incoming and outgoing partons and of the Higgs boson, retrieved from the VBFNLO LHE files, are used as input to the re-weighting code. The minimal set of VBF-like cuts applied to the partons is listed in table 4.3.

Additionally, the MG5_AMC@NLO program can simulate events with anomalous couplings also at NLO. This makes it possible to test how good an approximation it is to separately re-weight the $2 \rightarrow 2 + H$ and $2 \rightarrow 3 + H$ processes instead of developing a full NLO re-weighting. This is important for the analysis since the SM VBF signal sample that is re-weighted is generated at next-to-leading order by

MG5_AMC@NLO parameters numerical values
$c_\alpha = 0.6$
$K_{SM} = 1.6$
$k_{AWW} = -2.03$
$k_{AZZ} = -2.03$
$k_{A\gamma\gamma} = -155.97$
$k_{AZ\gamma} = 0$

Table 4.4: Input parameters for MG5_AMC@NLO corresponding to $\tilde{d} = \tilde{d}_B = 0.1$. All other anomalous couplings term are set to zero.

σ_{VBFNLO} [fb]	$\sigma_{\text{MG5_AMC@NLO}}$ [fb]
1451 ± 1	1447 ± 2

Table 4.5: Cross sections for VBF production of a CP-mixed state, corresponding to $\tilde{d} = \tilde{d}_B = 0.1$

POWHEG+PYTHIA. In MG5_AMC@NLO, the effective Lagrangian is expressed in terms of the couplings of the Higgs bosons to the photon A, and the W and Z bosons (see 1.48). The relation between those couplings and \tilde{d} are shown by equations 1.49. Table 4.4 shows the numerical values of the input parameters used to generate the MG5_AMC@NLO samples using the characterisation model [165] that correspond to $\tilde{d} = \tilde{d}_B = 0.1$. Good agreement is found when comparing the cross sections calculated by MG5_AMC@NLO and VBFNLO for the same CP-mixed state (see table 4.5).

The LHE files generated by MG5_AMC@NLO at NLO carry only parton level information and are therefore subsequently showered by interfacing MG5_AMC@NLO with PYTHIA8. A Rivet [166] routine is used to store the information on the partons in a ROOT file. The weights and Optimal Observable are calculated from these root files. Figure 4.5 shows comparisons between distributions of events generated by MG5_AMC@NLO directly with a CP-odd coupling strength of $\tilde{d} = \tilde{d}_B = 0.1$ and SM events re-weighted with the same CP-odd coupling strength. Neither $\Delta\phi_{jj}^{\text{sign}}$ nor the Optimal Observable show any significant disagreement in this comparison.

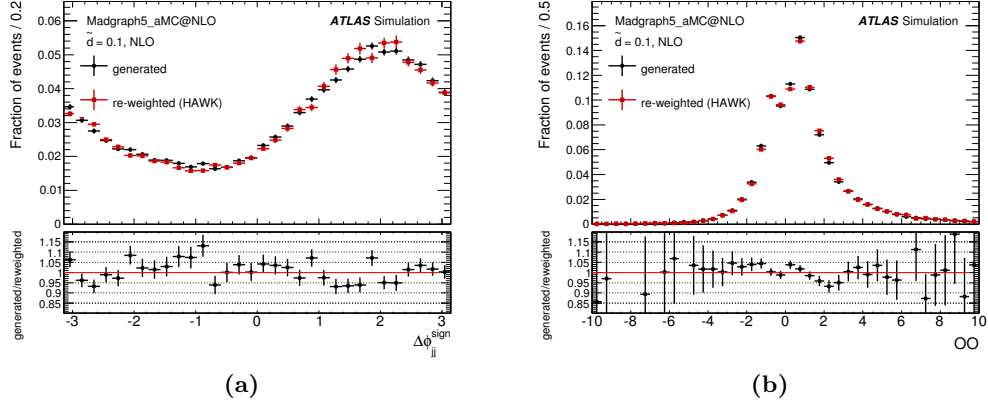


Figure 4.5: Distribution of $\Delta\phi_{jj}^{\text{sign}}$ and Optimal Observable for MG5_AMC@NLO Standard Model events after re-weighting (in red) and MG5_AMC@NLO events generated with $\tilde{d} = \tilde{d}_B = 0.1$. Both processes are generated at NLO, and parton-level information is used as input to the re-weighting.

4.4 Data Samples and Triggers

The analysis includes the full set of data collected by the ATLAS detector in proton-proton collisions in 2012 at a centre-of-mass energy of $\sqrt{s} = 8$ TeV. The 2012 dataset corresponds to 20.3 fb^{-1} after applying data quality criteria in the form of a Good Runs List (GRL). The GRL is identical to the one used in the couplings analysis for 8 TeV data. The peak instantaneous luminosity during the 2012 data taking period was $7.7 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ corresponding to a mean number of interactions per bunch crossing of 20.

Online event selection during data taking in the $\tau_\ell\tau_{\text{had}}$ decay channel is performed using single electron and muon triggers, both having a threshold of $p_T > 24$ GeV, to select $\tau_e\tau_{\text{had}}$ and $\tau_\mu\tau_{\text{had}}$ decays respectively. The lack of statistics in the anti-tau control region, defined as the region containing τ_{had} candidates that fail the identification criteria (see section 4.5.2), in lepton+ τ_{had} triggered events leads to a background distribution with high statistical fluctuations, and these triggers are therefore not used. The $\tau_\ell\tau_\ell$ channel employs different triggers to cover $\tau_e\tau_e$, $\tau_e\tau_\mu$ and $\tau_\mu\tau_\mu$ decays. The di-muon case accepts events both from a di-lepton trigger and a single electron trigger. The trigger thresholds are summarised in table 4.6. The thresholds are all kept identical to those employed in the $H \rightarrow \tau\tau$ couplings analysis. Offline p_T thresholds are subsequently applied as part of the analysis regions defined in section 4.6. Sufficiently high thresholds are chosen to reach the stable efficiency plateau in order to avoid potentially large systematic uncertainties associated with

Trigger [GeV]	$\tau_\ell\tau_\ell$	$\tau_\ell\tau_{\text{had}}$
Single electron	$p_T^e > 24$	$p_T^e > 24$
Single muon	-	$p_T^\mu > 24$
Di-electron	$p_T^{e1,2} > 12$	-
Di-muon	$p_T^{\mu1} > 18$	-
	$p_T^{\mu2} > 8$	-
Di-lepton	$p_T^e > 12$	-
	$p_T^\mu > 8$	-

Table 4.6: Online trigger p_T thresholds used in the $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ analysis channels.

the turn-on region, where the trigger efficiency rises steeply as a function of p_T . Only unprescaled triggers were utilised in this analysis.

4.4.1 Object Definitions

The objects used in the analysis include jets, hadronic tau decays, electrons and muons. Table 4.7 summarises the selection criteria of these objects in both channels. Some of them are only used in one of the channels, such as hadronic tau decays being exclusive to the $\tau_\ell\tau_{\text{had}}$ channel, and some of the electron and muon criteria differ between channels, while additional track and calorimeter isolation criteria are added individually for the two channels.

Jets are used to tag events that have a high probability of including VBF production. Local calibrated clusters are used to reconstruct jets with the anti- k_t algorithm using $r = 0.4$ (see section 3.3) and are included in the analysis if their pseudo-rapidity is within $|\eta| < 4.5$ and have $p_T > 30$ GeV. Jets within the ID acceptance of $|\eta| < 2.4$ having $p_T < 50$ GeV additionally need to have a jet vertex fraction of $|\text{JVF}| > 0.5$, which helps reject jets from pile-up. Furthermore, a veto is applied on events containing a b -jet by using the b -tagging algorithm MV1 (see section 3.3.2). These jets have a high probability of originating from $t\bar{t}$ events, and the veto therefore effectively reduces the $t\bar{t}$ background.

The reconstruction of hadronic tau decays follows the procedure described in section 3.6 at medium ID. They are required to have $p_T > 20$ GeV and need to be within $|\eta| < 2.47$. The number of charged tracks in the cone is limited to one or three with a total electric charge of 1. A medium electron veto is applied on the hadronic tau candidate in the $\tau_e\tau_{\text{had}}$ channel.

Electron and muon object selections differ slightly between $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$.

Object	Selection criterion	
	$\tau_\ell\tau_\ell$	$\tau_\ell\tau_{\text{had}}$
Muon	combined	combined
	$p_T > 10 \text{ GeV}$	$p_T > 10 \text{ GeV}$
	$ \eta < 2.5$	$ \eta < 2.5$
	ID quality criteria	ID quality criteria
	$I(p_T, 0.4) < 0.18$	$I(p_T, 0.4) < 0.06$
	$I(E_T, 0.2) < 0.09$	$I(E_T, 0.2) < 0.06$
Electron	loose identification	tight identification
	$p_T > 15 \text{ GeV}$	$p_T > 15 \text{ GeV}$
	$ \eta < 2.47$	$ \eta < 2.47$
	not within $1.37 < \eta < 1.52$	not within $1.37 < \eta < 1.52$
	$I(p_T, 0.4) < 0.17$	$I(p_T, 0.4) < 0.06$
	$I(E_T, 0.2) < 0.09$	$I(E_T, 0.2) < 0.06$
Hadronic tau decay	-	medium identification
	-	$p_T > 20 \text{ GeV}$
	-	$ \eta < 2.47$
	-	1 or 3 tracks
	-	charge = ± 1
	-	$\tau_e\tau_{\text{had}}$ electron veto
Jet	LC TopoClusters	LC TopoClusters
	$ \eta < 4.5$	$ \eta < 4.5$
	$ JVF > 0.5$ for $ \eta < 2.4$	$ JVF > 0.5$ for $ \eta < 2.4$

Table 4.7: Object selection criteria in the $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ decay channels.

In both channels electrons in the crack region $1.37 < |\eta| < 1.53$ between the barrel and end-cap calorimeters are ignored to avoid the inclusion of objects with poor performance in terms of identification and reconstruction. Electrons are also required to have $p_T > 15 \text{ GeV}$ and lie within $|\eta| < 2.47$. The equivalent kinematic requirements on muons are $p_T > 10 \text{ GeV}$ and $|\eta| < 2.5$. The $\tau_\ell\tau_\ell$ channel, in order to suppress fake backgrounds, requires tight muon identification using combined tracks (see section 3.5) with an additional requirement of $|z_0| < 10 \text{ mm}$ to avoid cosmic particle tracks from passing the selection. Electrons are required to pass the loose identification working point (see section 3.4). Events with τ_{had} candidates are vetoed. The $\tau_\ell\tau_{\text{had}}$ channel also requires combined muon tracks, while electron identification uses the tight working point.

Overlap removal is performed on objects passing the object selection criteria. It identifies objects lying within $\Delta R < 0.2$ of each other and keeps only one of them in the analysis. The ranking with the highest priority first is: muons, electrons, τ_{had} objects and lastly hadronic jets due to the relative purity of these objects in recon-

struction. However, this procedure can be improved in $\tau_\ell\tau_{\text{had}}$ by applying a looser set of selection criteria. During overlap removal both combined and segment-tagged muons are considered, while electrons are only required to pass the loose identification working point. This improves the rejection of events in which leptons are faking a τ_{had} . Overlap removal in $\tau_\ell\tau_{\text{had}}$ is followed by a dilepton veto to suppress $Z \rightarrow \ell\ell$ decays, and only events with exactly one light lepton and one τ_{had} object are retained. As an additional step, the light lepton purity is subsequently improved by applying additional isolation requirements on both tracks and calorimeter cells. In $\tau_\ell\tau_\ell$, muons must have tracks fulfilling $I(p_T, 0.4) < 0.18$, while calorimeter cells must have $I(E_T, 0.2) < 0.09$. Electron isolation requirements are $I(p_T, 0.4) < 0.17$ and $I(E_T, 0.2) < 0.09$. In $\tau_\ell\tau_{\text{had}}$ muons have track and calorimeter isolation requirements of $I(p_T, 0.4) < 0.06$ and $I(E_T, 0.2) < 0.06$ respectively. The electron isolation requirements are $I(p_T, 0.4) < 0.06$ and $I(E_T, 0.2) < 0.06$.

4.5 Data-driven background estimation

It is preferable to estimate the background contributions through data-driven modelling, which avoids any dependence on the choice of MC generator and minimises the systematic uncertainties while ensuring an accurate description of pile-up and the underlying event. The major background contributions such as Z decays to tau leptons and the fake backgrounds are all modelled using real data, with a special case being $Z/\gamma^* \rightarrow \tau\tau + \text{jets}$ that incorporates simulated tau decays into data events. In looser event categories not included in the present analysis, the couplings analysis used a method exploiting the ratio of event rates with electrically same-sign and opposite-sign light lepton and τ_{had} . It therefore uses a control region to estimate the background in the main signal region. A slightly different approach based on fake factors is needed in VBF events where instead of reversing the requirement of a particle’s electric charge, the τ_{had} identification is required to fail, creating a sample of so-called ‘anti-taus’. The following sections will describe the above mentioned methods of background modelling.

4.5.1 $Z \rightarrow \tau\tau$ Embedding Procedure

The irreducible nature of the $Z/\gamma^* \rightarrow \tau\tau + \text{jets}$ background complicates the rejection of signal events in a sample of $Z/\gamma^* \rightarrow \tau\tau$ events taken from data. Instead, the background is modelled with a hybrid data/MC approach called *embedding* [167]. The embedding is done by selecting as input a large sample of $Z \rightarrow \mu\mu$ data events. Muons are the most precisely reconstructed objects in ATLAS having negligible

momentum uncertainties compared to the uncertainties on tau decays. This clean dimuon event signature therefore provides the necessary means to accurately model $Z/\gamma^* \rightarrow \tau\tau$ events. The reconstructed energy deposits and tracks of the muons from the Z decay are then removed from each event and simulated tau decays are embedded in their place at the particle level. This approach has several advantages over using simulated events. The expected signal contribution in $Z \rightarrow \mu\mu$ events is negligible due to the very low branching ratio of $H \rightarrow \mu\mu$ decays, making it an especially pure source of Z events giving access to event kinematics directly from data, of which the additional jets are of special interest since these are an integral part of defining the VBF selection category. The underlying event and pile-up conditions are also taken directly from data with this approach and do not depend on simulation.

Events are selected using single muon triggers and requiring that all events have two combined isolated muons with opposite electric charge, a dimuon invariant mass $m_{\mu\mu} > 40$ GeV and pass p_T thresholds of 20 (15) GeV for the leading (subleading) muon in the offline selection. In events with more than one reconstructed muon pair, the one with an invariant mass closest to the Z boson mass is selected. While removal of the muon ID tracks is straightforward, subtraction of the calorimeter energy deposits requires the simulation of an equivalent muon pair. The kinematics of the two simulated muons are determined by their reconstructed counterparts in the data event, and the resulting simulated calorimeter energy deposits are subtracted from the reconstructed event. Two tau leptons are then simulated according to the reconstructed muon kinematics while correcting for the difference between the muon and tau masses. Simulation of the tau decays is performed in TAUOLA [154] with subsequent final state radiation handled by PHOTOS [155]. The spin correlation and polarisations of the two taus are handled by the TAUSPINNER [168] program. The result is then run through the ATLAS simulation and the tracks and calorimeter deposits of the simulated taus are added to the data event. The embedding sample only models the differential distribution of kinematic variables, while the initial normalisation of the sample is determined through a fit to the observed data. The fit is performed at pre-selection level with visible mass constrained to $40\text{GeV} < m_{\text{vis}} < 70\text{GeV}$. Although the selected dimuon sample has a high purity of $Z \rightarrow \mu\mu$ events there are still small numbers of $t\bar{t}$ and diboson events included. An overlap removal is performed by removing events that pass the embedding selection criteria in the simulated $t\bar{t}$ and diboson samples, which has a negligible effect on the analysis.

In order to validate the embedding procedure both the muon energy sub-

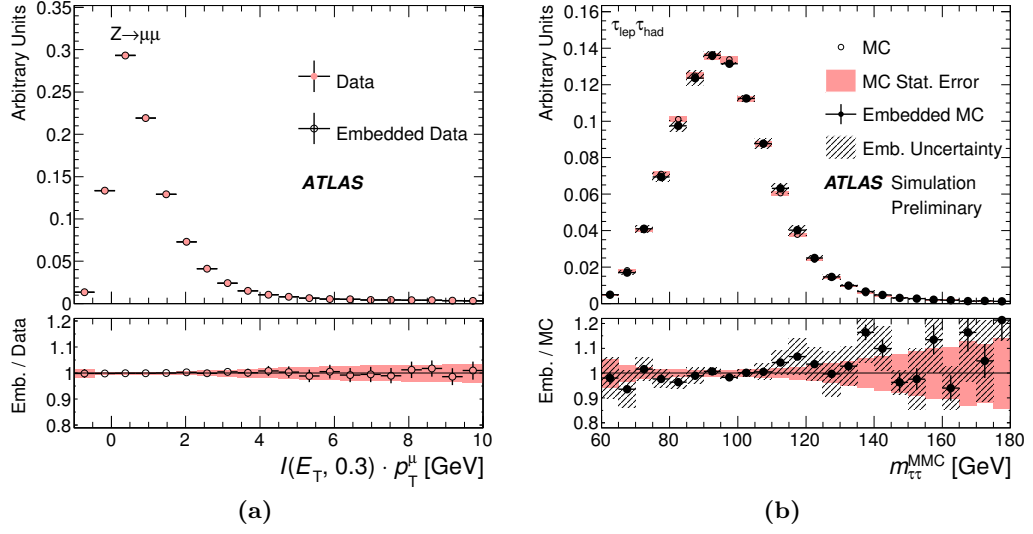


Figure 4.6: Variables used in the validation of the $Z \rightarrow \tau\tau$ embedding procedure. (a) shows the calorimeter isolation energy within a cone of radius $\Delta R < 0.3$ around the muons in $Z \rightarrow \mu\mu$ events comparing the results before and after the embedding of simulated muons. (b) shows the reconstructed invariant ditau mass in the $\tau\ell\tau_{\text{had}}$ final state comparing simulated $Z \rightarrow \tau\tau$ events with simulated $Z \rightarrow \mu\mu$ events after the embedding of simulated tau decays. Statistical uncertainties are given by the error bars on the data points, while systematic uncertainties are indicated by the hatched bands [2].

traction and the addition of simulated tau decays must be investigated. A separate embedding sample was created by replacing the muons from data with simulated muons instead of taus. Figure 4.6 (a) shows a comparison of the calorimeter energy deposits within a cone of $\Delta R < 0.3$ around the muons from data and the embedded simulated muons. The results are in good agreement with no indication of energy biases. In order to validate the embedding of simulated tau decays, samples of $Z \rightarrow \mu\mu$ and $Z \rightarrow \tau\tau$ events were generated by the same MC generator using identical settings followed by the embedding of taus into the $Z \rightarrow \mu\mu$ sample. The validation was performed by comparing kinematic distributions, one of which is the invariant ditau mass seen in figure 4.6 (b). The variables agree within uncertainties showing that the embedding procedure gives a valid description of $Z \rightarrow \tau\tau$ events.

4.5.2 Fake Tau Background Estimation

Simulation of events with jets faking hadronic tau decays is complicated by the calorimeter shower shapes of the jets. The suppression of fake hadronic tau de-

cays relies on algorithms that depend on shower shape variables that are difficult to accurately reproduce in MC simulation. Reliable simulation also depends on the modelling of the quark and gluon fractions within the jet sample. Gluon jets generally have a wider shape and contain more final state particles compared to quark jets due to their larger colour charge, which leads to differences in the misidentification rate. MC modelling of the quark and gluon fractions in samples containing multiple jets can be difficult and can easily lead to large systematic uncertainties. The production of large multijet event samples also requires immense amounts of processing power that further lowers the feasibility of simulating this kind of events. A data-driven estimation of these backgrounds is therefore highly preferred.

The Fake Factor Method

The ‘Fake Factor’ method [2] is a data-driven procedure based on measuring the ratio between the number of events containing τ_{had} candidates that pass and fail the identification criteria respectively. This is done in control regions that are enriched in the types of background faking τ_{had} objects, thereby providing control regions with much higher statistical power than the signal region.

The basic idea relies on exploiting the knowledge of the number of τ_{had} candidates that fail and pass the identification criteria. A τ_{had} candidate that fails the identification procedure is called an anti- τ_{had} in the following. The fake factor (FF) is then defined as the relative fraction of anti- τ_{had} to identified τ_{had} events, from which the background that enters the analysis can be estimated:

$$\text{FF} = \frac{N_{\text{CR}}^{\text{identified-}\tau}}{N_{\text{CR}}^{\text{anti-}\tau}}, \quad (4.8)$$

$$N_{\text{bkg}} = \left(N_{\text{data,SR}}^{\text{anti-}\tau} - N_{\text{others,SR}}^{\text{anti-}\tau} \right) \cdot \text{FF}, \quad (4.9)$$

where $N_{\text{data,SR}}^{\text{anti-}\tau}$ is the number of τ_{had} candidates in data in the signal region that fail the identification criteria, while $N_{\text{others,SR}}^{\text{anti-}\tau}$ is the equivalent event yield for real τ_{had} decays or light leptons reconstructed as τ_{had} candidates in processes that are modelled separately, including $Z \rightarrow \tau\tau$, $Z \rightarrow \ell\ell$, top and diboson. The terms included in ‘others’ are found in simulated event samples and have an effect on the final event yield of less than 10%. The number $N_{\text{CR}}^{\text{identified-}\tau}$ includes events with exactly one identified τ_{had} object in a given CR, conforming to the signal region selection criterion, while $N^{\text{anti-}\tau}$ in both CR and SR can include one or more anti- τ_{had} objects. The background entering the VBF category contains jets originating

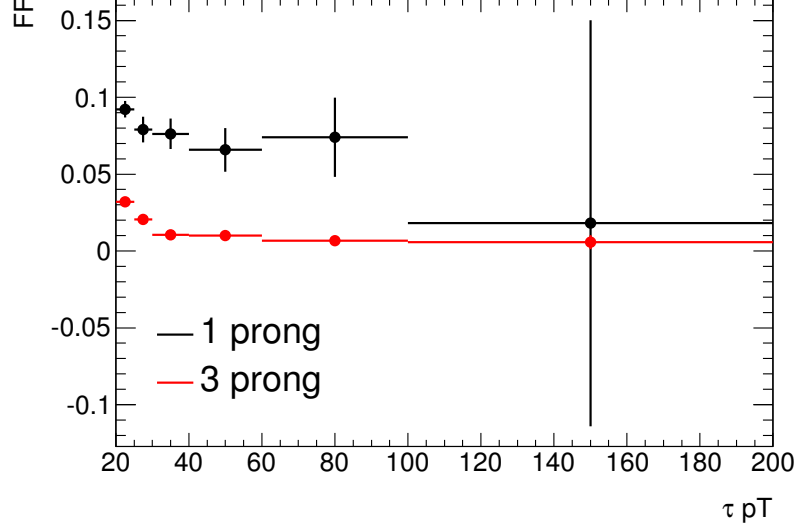


Figure 4.7: Fake factors extracted from the first six months of the 2012 data set. A clear p_T dependence is visible as well as differing values between one-prong and three-prong events.

from both gluons and quarks, and since the fake factor is highly dependent on the type of production mechanism, separate estimations of FF are performed in CRs dominated by jets originating from each type of process. Estimates are also needed from CRs enriched in $Z \rightarrow \ell\ell$ and top events. The fake factors are determined separately for one- and three-prong taus and in bins of the τ_{had} object p_T since they have been found to vary significantly depending on these variables as seen in figure 4.7. The final fake factor in a given n_{prong} and p_T bin is then given by

$$FF(n_{\text{prong}}, p_T) = \sum_i R_i FF_i(n_{\text{prong}}, p_T), \quad (4.10)$$

where i runs over the included processes, and R_i is the fraction of events contributed by process i . The R_i are determined using simulated event samples for $W + \text{jets}$ (0.46), $Z \rightarrow \ell\ell$ (0.11) and top (0.03), after which the contribution from QCD multijet events is calculable as $R_{\text{QCD}} = 1 - R_W - R_{\text{top}} - R_{Z \rightarrow \ell\ell} = 0.40$. The CRs used to determine the fake factors are listed in table 4.8 and are defined by modifying one or more selection criteria from the SR defined in table 4.12.

The reconstruction of E_T^{miss} is performed prior to the analysis, which means that anti- τ_{had} objects are assumed to be hadronic jets and are calibrated at the

Process	CR difference from SR
$W + \text{jets}$	Inverted m_T selection
QCD	Inverted track isolation and loosened calorimeter isolation
Top ($j \rightarrow \tau_{\text{had}}$)	At least one b -jet and $m_T < 70$ GeV
$Z \rightarrow \ell\ell$ ($j \rightarrow \tau_{\text{had}}$)	Requires two same-flavour leptons with invariant mass $61 \text{ GeV} < m_{\ell\ell} < 121 \text{ GeV}$

Table 4.8: CRs used to determine fake factors for $W + \text{jets}$ and QCD multijet events as well as events from $Z \rightarrow \ell\ell$ and top processes where a jet is misidentified as a τ_{had} object.

jet energy scale in the E_T^{miss} calculation. This would introduce a bias in the fake factor calculation and is avoided by re-running the E_T^{miss} reconstruction on the data samples used to determine the fake factors in order to ensure that all τ_{had} object are treated using the tau energy scale.

4.6 Event Selection

This section describes the event selection performed in the $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ decay channels to construct the phase space regions of interest to the analysis. Most of the selection criteria were developed in the couplings analysis, apart from an additional multivariate discriminator requirement (see section 4.6.1). As a first step after passing the data quality requirements and trigger thresholds described in section 4.4, events with the desired final state objects from the signal process are selected using a set of pre-selection requirements. These are followed by the definition of an event category enriched in VBF signal events and finally a requirement on the output of the multivariate discriminator to obtain the final signal region (SR). Several control regions (CRs) are defined in order to constrain individual background normalisations by modifying certain kinematic requirements in the VBF category to obtain regions enriched in those background processes.

4.6.1 Background Suppression

The current analysis performs a measurement of CP violation in the VBF production of $H \rightarrow \tau\tau$ decays. Variables describing the characteristics of the ditau final states and the two leading jets in VBF events are integral to studying the differences between $H \rightarrow \tau\tau$ signal and background events. The discriminating variables employed in the $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ decay channels include various kinematic constructs

incorporating momenta, invariant masses and angular separations of the jets and visible tau decay products. These include the opening angle ΔR in the η - ϕ plane between the visible tau decay products, the invariant mass of the two leading jets m_{j_1, j_2} , the MMC ditau mass $m_{\tau\tau}^{\text{MMC}}$ as well as various centrality measures that describe the relative position of one object with respect to two other objects in the event. One such type of centrality is called the η centrality of an object relative to the two leading jets in an event. It is defined as

$$C_{\eta_1, \eta_2}(\eta) = \exp \left[\frac{-4}{(\eta_1 - \eta_2)^2} \left(\eta - \frac{\eta_1 + \eta_2}{2} \right)^2 \right], \quad (4.11)$$

where η , η_1 and η_2 are the pseudo-rapidities of the object and the two leading jets respectively. The variable has a value of 1 when the object is centred in η between the jets and drops below $1/e$ if it lies outside the range of η values between the jets. The $\tau_\ell \tau_\ell$ channel makes use of the η centrality of a third jet in the event, $C_{\eta_1, \eta_2}(\eta_{j_3})$, and the product of the η centralities of the two leptons from the tau decays. The $\tau_\ell \tau_{\text{had}}$ channel uses the η centrality of the lepton from the leptonic tau decay $C_{\eta_1, \eta_2}(\eta_l)$. VBF events typically have higher η centrality values compared to other processes as the quarks that radiate the vector bosons continue on to become jets in the forward regions of the detector, while the Higgs boson is generally found in the central detector.

Another type of centrality utilised only in the $\tau_\ell \tau_{\text{had}}$ channel is the E_T^{miss} ϕ centrality, which describes the relative angular orientation of the missing transverse momentum compared to the visible tau decay products in the transverse plane. The transverse plane is transformed in such a way that the direction of the tau decay products are orthogonal and the ϕ angle between the tau decay products defines the positive quadrant of the transformed plane. The E_T^{miss} ϕ centrality is defined as the sum of the x - and y -components of the E_T^{miss} unit vector in this transformed plane. It is mathematically defined as

$$E_T^{\text{miss}} \phi \text{ centrality} = \frac{A + B}{\sqrt{A^2 + B^2}}, \quad (4.12)$$

where

$$A = \frac{\sin(\phi_{E_T^{\text{miss}}} - \phi_{\tau_{\text{had}}})}{\sin(\phi_l - \phi_{\tau_{\text{had}}})}, \quad B = \frac{\sin(\phi_l - \phi_{E_T^{\text{miss}}})}{\sin(\phi_l - \phi_{\tau_{\text{had}}})}. \quad (4.13)$$

It will have a value of $\sqrt{2}$ in the case where the E_T^{miss} vector is precisely centered between the lepton and the τ_{had} object, while a value of $-\sqrt{2}$ indicates that the E_T^{miss} vector points in the opposite direction. In resonant decays to ditau states such as

$H \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$ it is expected that the missing transverse energy points in a direction lying between the tau decay products since both tau decays contribute to the E_T^{miss} vector. A value of the E_T^{miss} ϕ centrality close to $\sqrt{2}$ therefore indicates a decay of this type.

Suppression of $W + \text{jets}$ events in $\tau_\ell\tau_{\text{had}}$ can be achieved through the transverse mass m_T of the lepton and E_T^{miss} (see equation 4.1). Additionally, the two high- p_T jets present in VBF events can be used to construct a range of observables to aid in the discrimination between $H \rightarrow \tau\tau$ and $Z/\gamma^* \rightarrow \tau\tau$ events. These all rely on the distinct properties of the jets that are typically well-separated in η , have high transverse momenta and occur with relatively low amounts of additional hadronic activity in the event due to the absence of colour exchange in the VBF process. Apart from the η centralities already mentioned above, the following observables are used in the analysis:

- $\Delta\eta_{j_1,j_2}$: Pseudo-rapidity separation between the two leading jets, which is typically large in VBF events. This variable is directly used in the definition of the VBF category.
- $\eta_{j_1} \times \eta_{j_2}$: Product of the pseudo-rapidities of the two leading jets. This is another measure of their separation indicating whether the jets occupy the same or the opposite detector hemisphere according to the sign of the variable.
- m_{j_1,j_2} : Invariant dijet mass, containing information about their momentum and angular separation. The distribution is expected to tend towards higher values in $H \rightarrow \tau\tau$ signal events compared to $Z/\gamma^* \rightarrow \tau\tau$ and fake background events.
- p_T^{tot} : Length of the vector sum of the visible tau decay products, the two leading jets and the missing transverse energy. This variable will tend towards small values in events with little or no additional activity, which is expected in VBF events.

Many of the variables mentioned above will be correlated to some degree. Figure 4.8 shows the linear correlation coefficients between a subset of the discriminating variables used by the BDT in both a signal and background sample. The highest level of correlation is seen between the opening angle of the visible decay products ΔR and the ditau mass $m_{\tau\tau}$. This is expected since a heavier resonance produces a ditau system of higher momentum, which is therefore boosted to a higher degree. If the variable correlations differ between signal and background, this can be exploited to extract further discrimination power.

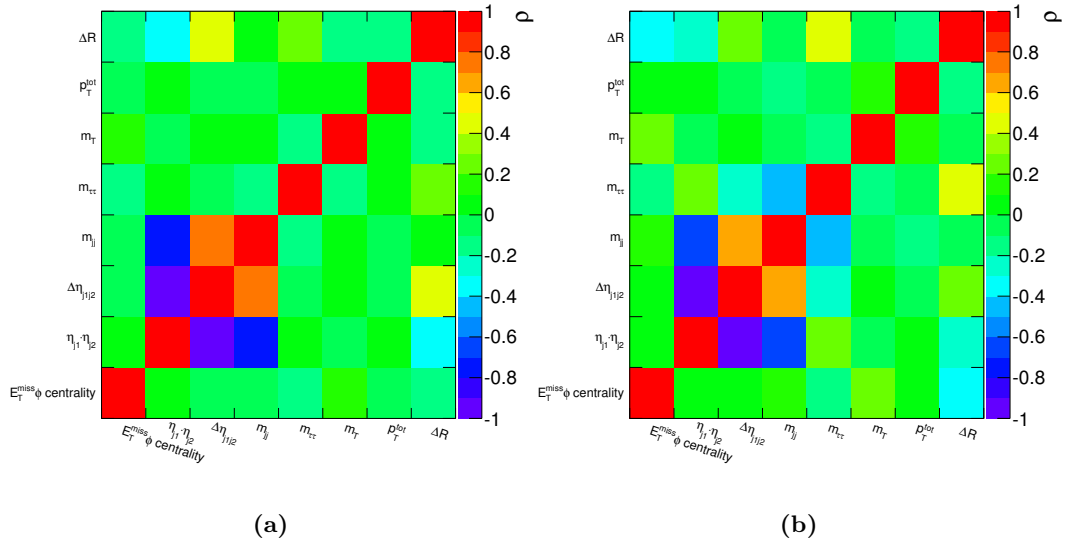


Figure 4.8: Linear correlation factors between input variables to the BDT discriminator for VBF $H \rightarrow \tau\tau$ signal (a) and $Z \rightarrow \tau\tau$ embedding background (b) in the VBF event category.

Boosted Decision Trees

A boosted decision tree (BDT) [169] is a predictive multivariate classifier used in machine learning. In contrast to binary pass/fail event classification based on a set of variables with corresponding selection criteria, the output of a BDT is a continuous variable that describes how signal- or background-like an event is. In order to optimise the classifier to distinguish between specific categories of events, it is *trained* on a set of simulated events where the categories are known in advance. The input variable space typically has a high number of dimensions with potentially large and complex correlations between the variables. Multivariate classifiers such as BDTs are able to exploit features and correlations between variables that would be difficult to access in a more conventional cut-based analysis.

A BDT is related to a cut-based analysis with event categories and a set of selection criteria as the cut-based approach can also be presented in the form of a decision tree. The selection criteria can be seen as a sequential series of pass/fail decisions or *splits*, each leading to a pass node and a fail node and finally ending in a set of leaf nodes having some mixture of true signal and background events that determines the signal-likeness of each leaf node. Figure 4.9 illustrates the structure of such a decision tree. The order of the variables in the tree and the value of each variable where the split occurs can be optimised using a metric such as the Gini

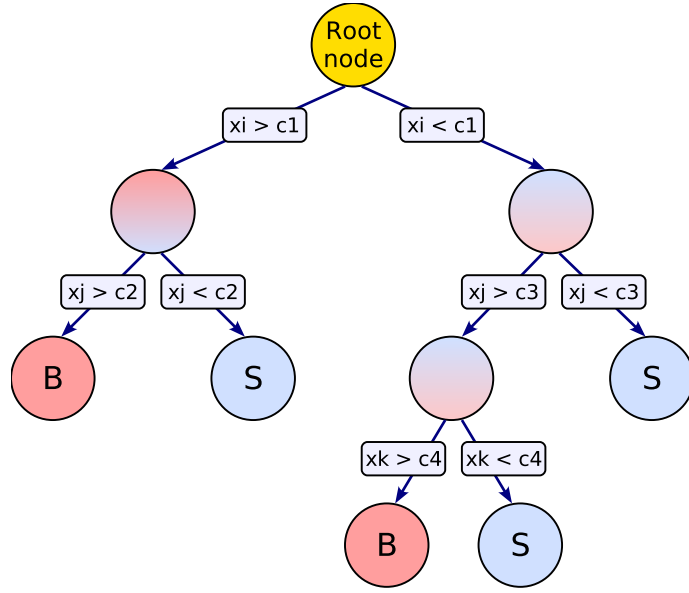


Figure 4.9: Illustration of a decision tree with a sequence of splits according to selection criteria on the discriminating variables x_i . The final leaf nodes indicate how signal- or background-like the events in that node are according to the mixture of training events in each node [170].

index $G = p \cdot (1 - p)$, where p is the signal purity at each node. At every node the goal is then to find the variable and selection cut value that maximises the increase of the Gini index when comparing the index value of the parent node with the sum of the index values of the two daughter nodes weighted according to their relative event fractions. Such a training procedure can be taken too far however. With no restrictions on the number of splits, also called the tree's depth, a situation can occur where each leaf node is so specialised that it contains only one training event. While this would provide perfect background rejection on the training sample, any other test sample that is run through this decision tree will not perform this well. Decision trees with an excessive depth tend to identify statistical fluctuations as genuine features of the training sample, leading to *overtraining*. Therefore, criteria are often added that limit the number of allowed splits and require a minimum number of events in a leaf node.

A single decision tree is a relatively weak classifier. The misclassification rate can be significantly reduced by combining many decision trees, each with different event weights. By analysing the result of the training of the initial decision tree, events that were difficult to categorise correctly can be given a higher relative weight to emphasise the problematic features, whereafter the re-weighted event

sample is used to train a new decision tree. Such a combination of multiple decision trees results in a classification that is no longer a binary outcome, but a function of the responses of the individual trees. This analysis implements BDTs through the Toolkit for MultiVariate Data Analysis (TMVA) [170] and uses the Gradient Boosting algorithm [171]. The output of the BDT algorithm will be referred to as the $\text{BDT}_{\text{score}}$.

In the couplings analysis BDTs were trained separately on simulated ggF and VBF events to optimise the sensitivity to both event topologies. This analysis utilises only the VBF-trained BDTs, with the $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ decay channel each having their own BDT with individually picked lists of input variables. It is preferred to include only a limited number of input variables since the inclusion of a variable with low discrimination power potentially introduces an unfavorable amount of complexity compared to its contribution to the classifier. Extra systematic uncertainties can arise, and it could complicate the successful modelling of the BDT output distribution, which depends on the correct description of the variable correlations. The list of potential input variables is therefore reduced by first discarding any variables that are not well-modelled and by requiring only weakly correlated variables or variables with well-understood correlations that differ between signal and background. The importance of each variable in the subsequent BDT training can be determined by studying how many times each variable is used to define a node splitting and weighting each occurrence by its increase in the Gini index and the number of events in the node. Variables with a low ranking according to this procedure can be considered for removal from the input list. Table 4.9 lists the final choice of discriminating variables used as input when training the BDT in each channel.

The BDT training event samples should accurately describe the signal and background distributions of the training variables in the event category where the BDT is to be applied. While one could use the exact same background model as the final analysis for this purpose, some of the background samples used for the BDT training are generated differently. The size of systematic uncertainties in the samples are the deciding factor in the choice of background model, whereas the number of available events is more important in the BDT training procedure. Therefore, simulated samples of $Z/\gamma^* \rightarrow \tau\tau$ events are substituted, offering a much larger number of events than the data-driven fake factor method. These events are generated using ALPGEN including a VBF filter as described in section 4.2.2. The OS-SS method is used on these samples to model events with misidentified hadronic tau decays.

Variable	Description	$\tau_\ell\tau_\ell$	$\tau_\ell\tau_{\text{had}}$
$m_{\tau\tau}$	Ditau mass calculated by the MMC algorithm	•	•
$\Delta R_{\tau_1, \tau_2}$	η - ϕ separation of the two tau leptons	•	•
$\Delta\eta_{j_1, j_2}$	η separation of the two leading jets	•	•
m_{j_1, j_2}	Invariant mass of the two leading jets	•	•
$\eta_{j_1} \times \eta_{j_2}$	Product of η values of the two leading jets		•
p_T^{tot}	$ \mathbf{p}_T^{\tau_1} + \mathbf{p}_T^{\tau_2} + \mathbf{p}_T^{j_1} + \mathbf{p}_T^{j_2} + \mathbf{E}_T^{\text{miss}} $		•
m_T	Transverse mass of $l + E_T^{\text{miss}}$		•
E_T^{miss} ϕ centrality	E_T^{miss} direction relative to visible tau decays		•
$C_{\eta_1, \eta_2}(\eta_l)$	η centrality of lepton		•
$\min(\Delta\eta_{l_1 l_2, \text{jets}})$	Min. $\Delta\eta$ of dilepton system and either jet	•	
$C_{\eta_1, \eta_2}(\eta_{l_1}) \times C_{\eta_1, \eta_2}(\eta_{l_2})$	Product of η centralities of leptons	•	
$C_{\eta_1, \eta_2}(\eta_{j_3})$	η centrality of third jet	•	

Table 4.9: Discriminating variables used in the training of the BDT for each decay channel. The bullets indicate which variables are used in a particular decay channel. Some of the variables are described in more detail in the text.

Events used to train a BDT should not be included in any test samples that are subsequently processed by it. This requires a splitting of the available events into two mutually exclusive samples that can only be used for either training or applying the BDT respectively. A technique called cross-evaluation is utilised to maximise the statistical power available to the BDT training by opening up the possibility of using the full set of events. This is done by randomly separating the background samples into two groups, each used to train an independent BDT. Each BDT is then applied to the other half in order to test its performance. Since the event splitting is random (picking alternate events in the sequence they reside in the data samples) the two BDTs will exhibit the same traits in terms of output score and shape within statistical uncertainties, and their scores can simply be added. Applying the BDTs to data then entails splitting the events into two random samples of equal size and applying each of the two cross-evaluated BDTs to one of the halves.

4.6.2 Signal and Control Regions

Following the object selection, the decay channels have individual event categorisations designed to be sensitive to VBF signal events. The defining characteristic of a VBF event is the detection of two high- p_T jets that are well-separated in pseudo-rapidity. The VBF region is therefore defined by applying selection criteria to the p_T and $|\Delta\eta|$ of these jets in order to reduce the contamination from other production

modes such as ggF and VH. However, requiring the presence of a VBF signature does not reduce the non-VBF contributions to a negligible amount. The selection criteria summarised in table 4.12 result in the event yields in table 4.14, which show that in the VBF category the total event yield from Higgs production has an expected non-VBF contribution of 36% in $\tau_\ell\tau_{\text{had}}$ and 46% in $\tau_\ell\tau_\ell$. A high-purity VBF sample for the purposes of the CP measurement is achieved by selecting events with a high $\text{BDT}_{\text{score}}$. This selection criterion is optimised for each channel individually as discussed in the sections below detailing the channel selections.

The $\tau_\ell\tau_{\text{had}}$ and $\tau_\ell\tau_\ell$ decay channels apply their respective event selections developed in the couplings analysis [2], including the the preselection criteria, the BDT training and the definition of the VBF event category in each decay channel. This analysis adds an additional selection cut on the $\text{BDT}_{\text{score}}$ discriminant that is optimised and applied on top of the VBF category in order to define a signal region with very signal-like events, thereby increasing the signal over background ratio considerably. This is a necessary step to ensure sensitivity to the differences in the Optimal Observable distribution of CP-mixing models compared to SM signal (see section 1.3.2). It is possible to estimate an optimal range for this cut in each channel individually by calculating the statistical significance for a counting experiment as a function of the cut on the $\text{BDT}_{\text{score}}$. The significance is defined as [172]

$$Z_0 = \sqrt{2((s+b)\ln(1+s/b) - s)}. \quad (4.14)$$

However, the $\text{BDT}_{\text{score}}$ selection value having the highest sensitivity for a counting experiment in both channels approaches 1.0 and therefore contains an insufficient number of events for the likelihood fit to be trusted. In the low-statistics limit, error estimation may become inaccurate and the fit may become increasingly biased due to asymmetries in the distribution. A compromise must therefore be reached between the estimated sensitivity and the amount of data present in each bin of the Optimal Observable distribution in the signal region. In order to ensure stability of the binned maximum-likelihood fit, the $\text{BDT}_{\text{score}}$ selection cut and histogram binning are chosen such that at least one expected total background event is present in each bin as given by MC. This necessitates the merging of high $|O_1|$ bins, since statistics are quickly depleted in these regions of the distribution. The final choice of binning is found by varying the number of bins as well as the placement of the bin edges of the two outermost bins, and in each case performing a likelihood fit between a simulated dataset including SM signal and datasets including signal generated in a range of \tilde{d} values. More details are provided for the $\tau_\ell\tau_{\text{had}}$ decay channel in

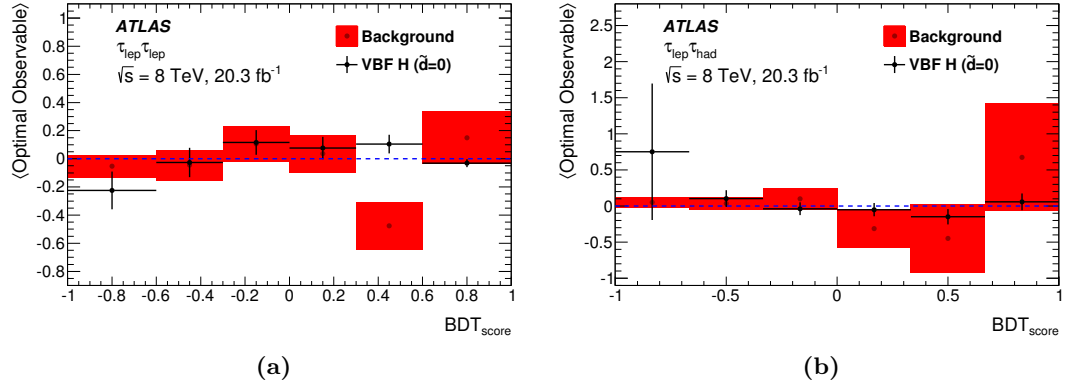


Figure 4.10: Dependence of the mean value of O_1 on BDT_{score} shown for total background as well as VBF $H \rightarrow \tau\tau$ SM signal in the $\tau_\ell\tau_\ell$ channel (a) and the $\tau_\ell\tau_{\text{had}}$ channel (b).

section 4.6.2.

A check was performed in both decay channels in order to ensure that the Optimal Observable distribution is not biased by the BDT_{score} selection cut used for the signal region. Figure 4.10 shows the mean value of the Optimal Observable in each decay channel as a function of the BDT_{score} value for SM signal in addition to the full background estimate. The mean value is seen to be compatible with zero within the statistical uncertainties as expected for a purely SM coupling.

The $\tau_\ell\tau_{\text{had}}$ channel

The $\tau_\ell\tau_{\text{had}}$ analysis channel is expected to have the largest impact on the CP-odd coupling sensitivity, having a branching fraction of approximately 4 times that of $\tau_\ell\tau_\ell$. It contains the final states having exactly one isolated lepton (e or μ) and one hadronic tau candidate with opposite charge.

The event selection in this analysis is based on the selection developed in the couplings analysis with an additional tighter selection criterion placed on the BDT_{score} to select a purer sample of VBF events. In order to define a preliminary search region that is relatively clean, a set of pre-selection cuts are applied. These are designed to remove e.g. events that may have been contaminated by cosmic particles that can result in fake E_T^{miss} , as well as events containing inaccurately measured energy of electrons in the EM calorimeter. Following the definitions in the couplings analysis, events are selected by applying p_T requirements on the lepton and hadronic tau, a di-lepton veto removing all events with two or more high- p_T leptons to reduce the $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ backgrounds, and requiring that the

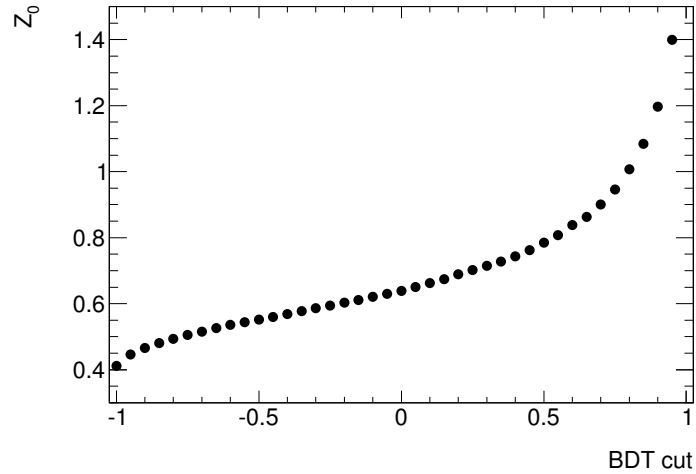


Figure 4.11: Statistical significance Z_0 of a counting experiment including signal and all backgrounds in $\tau_\ell\tau_{had}$ as a function of the chosen BDT_{score} requirement.

lepton and hadronic tau have opposite electric charge. The VBF event category, again identical to that in the couplings analysis, is designed to select events with VBF kinematics and requires $E_T^{miss} > 20$ GeV, a τ candidate with $p_T > 30$ GeV and at least two jets with leading $p_T > 40$ GeV and subleading $p_T > 30$ GeV. In addition, the two leading jets must be in the forward and backward halves of the detector, requiring $\eta_{j1} \times \eta_{j2} < 0$, and have an invariant mass $m_{j1j2} > 500$ GeV. The pre-selection and signal regions are summarised in table 4.12.

A sensitivity optimisation of the BDT_{score} cut gives the rejection of BDT_{score} values below 0.95 as a compromise between the number of events and the sensitivity. Figure 4.11 shows the functional dependence of the statistical significance on the value of the chosen BDT_{score} selection cut in the VBF category. The O_1 binning in the signal region is chosen such that the sensitivity is maximized while all bins still contain at least one unweighted event from every background sample. The choice of cutting at a BDT_{score} of 0.95 was made in order to allow the requirement of one unweighted event per bin to be met. A looser cut would lower the sensitivity, while a tighter cut would deplete the periferal bins of the Optimal Observable distribution leaving an insufficient amount of events to satisfy the requirement. Several binning choices have been tested and for each choice the full fit procedure has been repeated. Due to the complexity of the fit when including systematic uncertainties, these are not included when performing the optimisation. The final choice of signal region binning is listed in table 4.10. A single empty bin is observed in the top background histogram, and a fix is applied to this bin in order to avoid any potential fit problems.

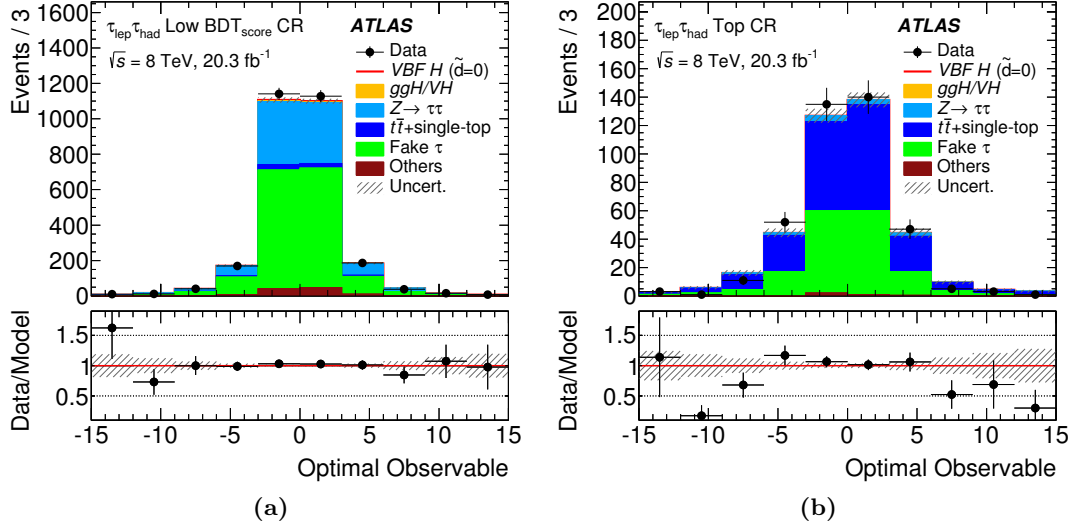


Figure 4.12: Distributions of O_1 in the low BDT control region (a) and top control region (b) in the $\tau_\ell \tau_{had}$ channel. The error bands indicate only statistical uncertainties.

This fix fills the empty bin with the mean event weight calculated from the set of top background events in the signal region.

All expected event yields for individual backgrounds and the signal are listed in table 4.14. In addition to the final signal region, the fit includes a top-enriched control region using O_1 and a low BDT_{score} control region using the BDT_{score} distribution below 0.95. The regions are defined in table 4.13, and figure 4.13 illustrates the fit model. The control regions help constrain the normalisation of the $Z \rightarrow \tau\tau$ and top related backgrounds, which are free parameters in the fit.

Figure 4.14 shows O_1 for pure VBF signal with various CP-odd coupling hypotheses. The distributions are normalised to unity to remove any additional discrepancy arising from an increased cross section when adding CP-odd couplings. Furthermore, two control region plots of O_1 for the low BDT control region and the top control region are shown in figure 4.12. The top control region is used both as

Variable	Region	Bin edges
Optimal observable O_1	Signal region	-15,-5,-2.5,0,2.5,5,15
	Top CR	Only one bin
BDT score	Low BDT CR	-1, -0.95, -0.35, 0.35, 0.7, 0.851, 0.904, 0.936, 0.95

Table 4.10: Histogram binnings in the final log likelihood fit in $\tau_\ell \tau_{had}$

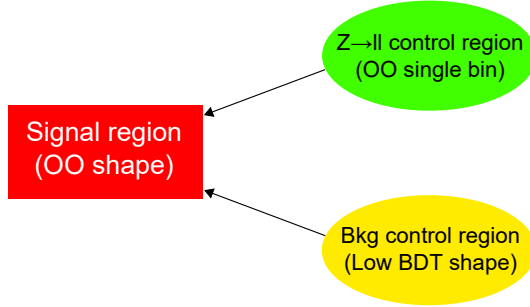


Figure 4.13: Illustration of the kinematic regions included in the $\tau_\ell \tau_{\text{had}}$ fit model. The normalisations of the $Z \rightarrow \tau\tau$ and top backgrounds are free in the fit and are constrained by the low $\text{BDT}_{\text{score}}$ CR and top CR respectively.

a check of the modelling of O_1 , and as a means to constrain the normalisation of the top background in the final likelihood fit. In the final fit only the total yield is used, corresponding to a histogram with a single bin. Both control regions show good agreement between the expected yields and data.

A separate event category containing boosted Higgs events, mutually exclusive from the VBF category, was included in the couplings analysis. The choice to not include the boosted category as a control region in this analysis has the potential to change the optimal fit model in the remaining regions, which in the case of $\tau_\ell \tau_{\text{had}}$ is limited to the low $\text{BDT}_{\text{score}}$ control region, since this is the only other control region that uses the shape of the distribution. A complete fit including systematic uncertainties and with pseudo-data containing the estimated background plus signal for different values of \tilde{d} , also called Asimov data, in the signal region was performed for a set of binning variations in the low $\text{BDT}_{\text{score}}$ control region. Real data was used in the low $\text{BDT}_{\text{score}}$ control region. For more on this ‘hybrid’ fit including Asimov and real data depending on the region, see section 6.2. The binning variations included changing the number of bins in the remaining region below 0.95 as well as limiting the control region to lower values only. It was decided to retain the same non-equidistant binning used in the couplings analysis, meaning all bins below a $\text{BDT}_{\text{score}}$ of 0.95 have identical bin edges to those optimised in the couplings analysis (see table 4.10). This decision was informed by taking into account the discrepancy between the signal strength μ included in the Asimov data and the best estimate from the fit as well as the compatibility of the data in control regions with the post-fit background distributions.

The Optimal Observable is constructed using the theoretical matrix element of the process (eq. 1.60), which means that the best performance would have been obtained if one had access to the true values of the kinematic variables used to cal-

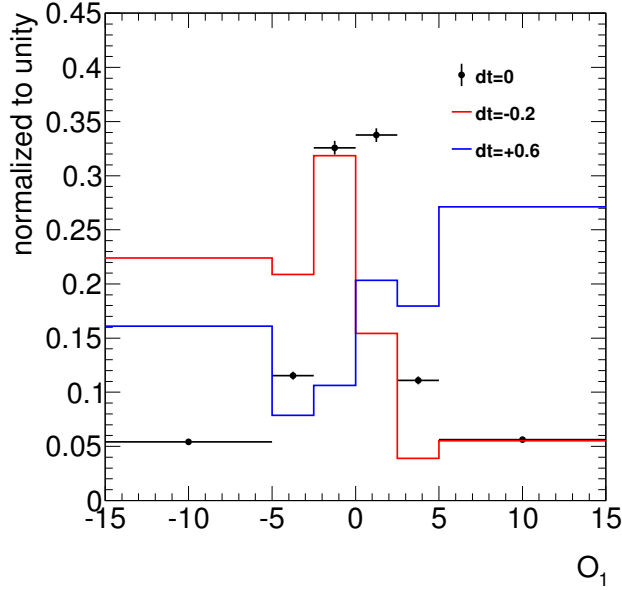


Figure 4.14: Comparison of VBF signal distributions in $\tau_\ell\tau_{\text{had}}$ when assuming $\tilde{d} = 0.0, -0.2, +0.6$. The distributions have all been normalised to unity in order to compare only the shape difference when varying the CP-odd coupling strength.

culate it. In order to gauge the impact on the O_1 distribution when transitioning from truth to reconstruction level variables, figure 4.15 shows distributions of O_1 calculated with all input variables at truth level, reconstruction level and various mixtures of both. Although some bin migration of individual measurements is observed, the difference in the input values is seen to have a very minimal impact on the total distribution of O_1 .

A binning optimisation was performed on the input histograms of the Optimal Observable in the $\tau_\ell\tau_{\text{had}}$ signal region. Only statistical errors were included for simplicity. It was found that the shape of the resulting ΔNLL curve is approximately independent of the chosen selection cut on the $\text{BDT}_{\text{score}}$ variable within the investigated range of $0.9 < \text{BDT}_{\text{score}} < 0.98$. The binning can therefore be optimised independently, followed by a choice of $\text{BDT}_{\text{score}}$ cut that ensures at least one total expected background event in each bin. This condition, in conjunction with the requirement that there be at least one unweighted event included from each background sample, is in place to prevent instabilities in the final fit. The optimisation is performed with respect to the resulting one sigma confidence interval. An upper limit on the numerical value of the Optimal Observable $|O_1| < 15$ is introduced to exclude the region where an insufficient number of events is expected and fluctuations can have a large impact on the results. This is also a safety measure to

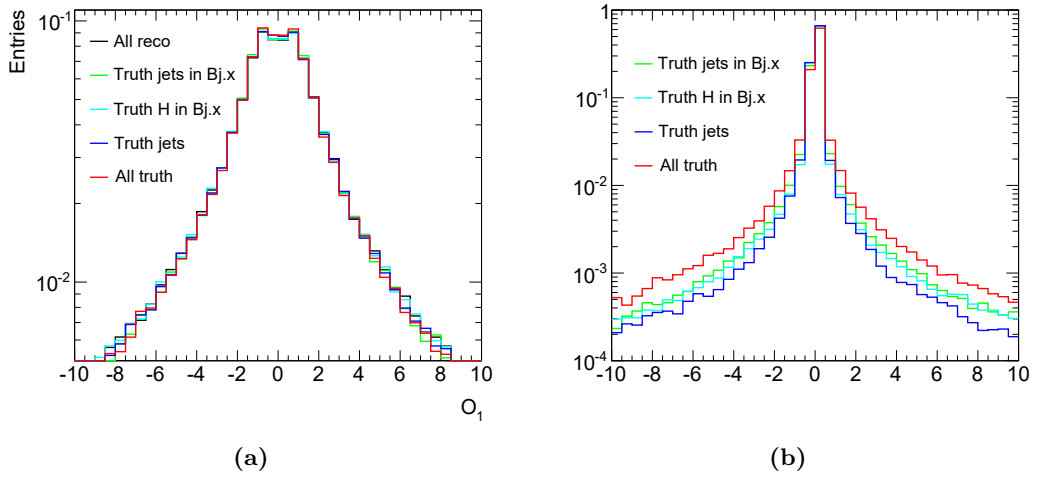


Figure 4.15: *Distribution of O_1 using VBF signal events in $\tau_\ell\tau_{had}$ replacing individual reco level variables with truth reconstructed values in order to gauge the impact of going from truth to reco level. O_1 itself is shown in (a), while the residual between the full reco calculation and each of the calculations with a truth value included is shown in (b).*

avoid large event weights in the region of low statistics. Re-weighting signal events at high values of \tilde{d} can potentially lead to excessive event weights, especially at large O_1 values. The number of histogram bins is varied from 4 to 12 in the case of equidistant bins. This leads to bin configurations both with a bin centered on $O_1 = 0$ and with a bin edge at this position. It was found that configurations with a bin edge at $O_1 = 0$ have slightly superior exclusion power. The optimisation also considers binning choices where the outermost two or three bins are merged in order to reduce the effects of low statistics in the tails of the distribution. A selection of these binning choices are shown in figures 4.16 to 4.18. Their corresponding ΔNLL curves in the case of $\text{BDT}_{\text{score}} > 0.95$ are plotted in figure 4.19. The final binning choice (J) was found given the criterion that the one standard deviation confidence interval should be minimised. It consists of four central equidistant bins of width 2.5 surrounded by two peripheral bins of width 10.

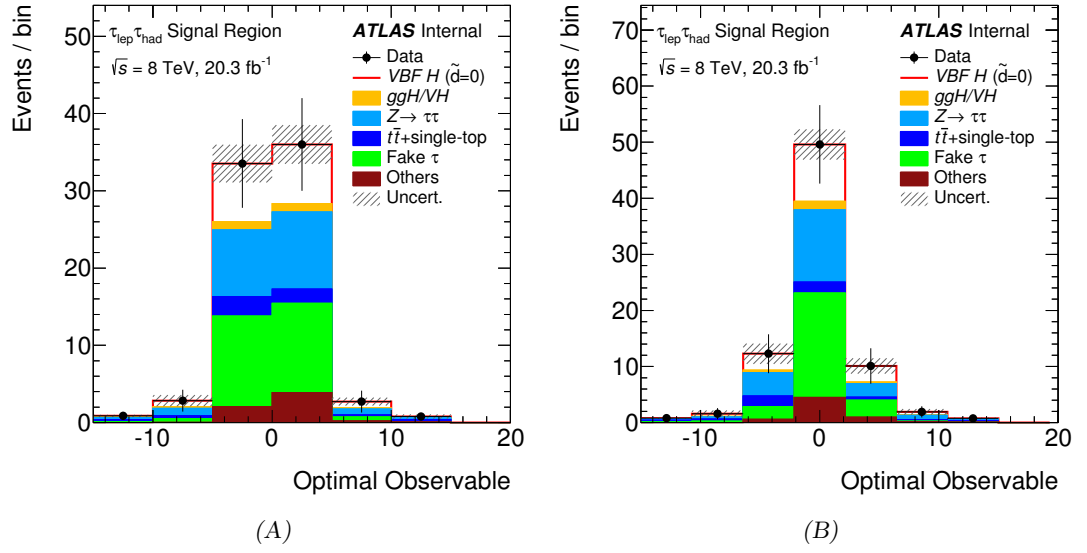


Figure 4.16: Optimal Observable in the final signal region plotted using a range of binning choices to optimise the fit sensitivity. The data points are Asimov data with an injected signal strength of $\mu = 1.55$.

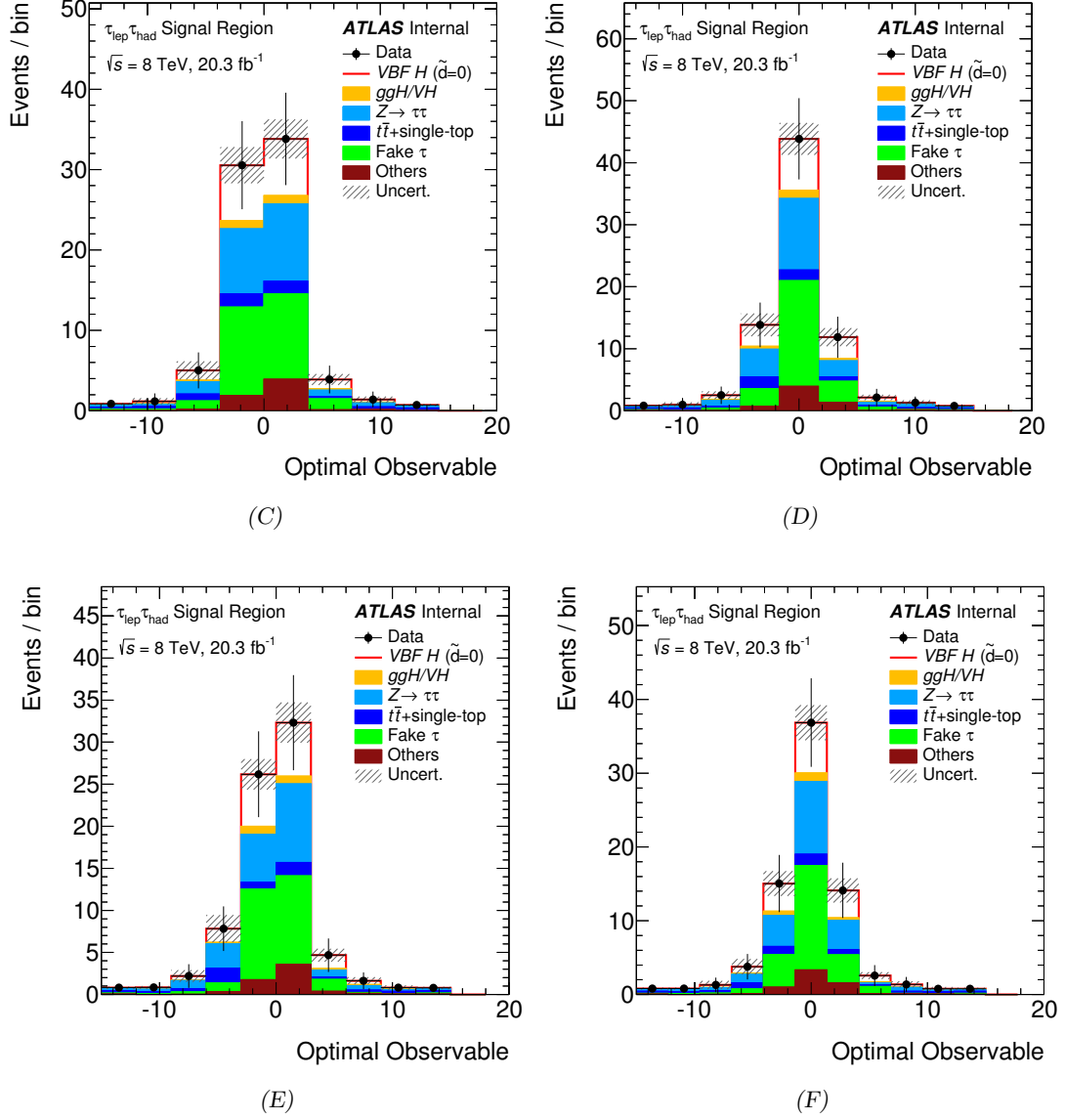


Figure 4.17: Optimal Observable in the final signal region plotted using a range of binning choices to optimise the fit sensitivity. The data points are Asimov data with an injected signal strength of $\mu = 1.55$.

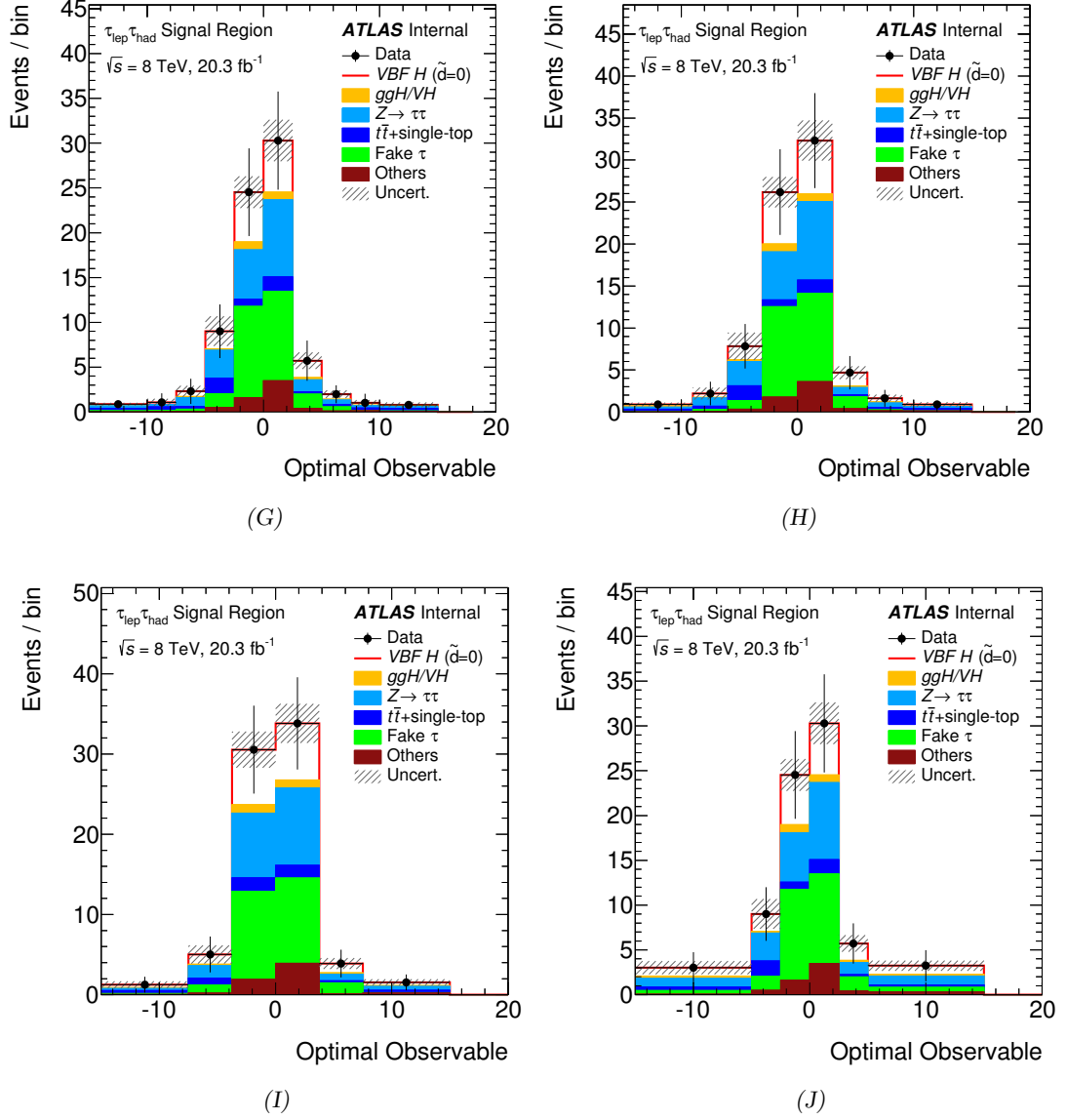


Figure 4.18: Optimal Observable in the final signal region plotted using a range of binning choices to optimise the fit sensitivity. The data points are Asimov data with an injected signal strength of $\mu = 1.55$.

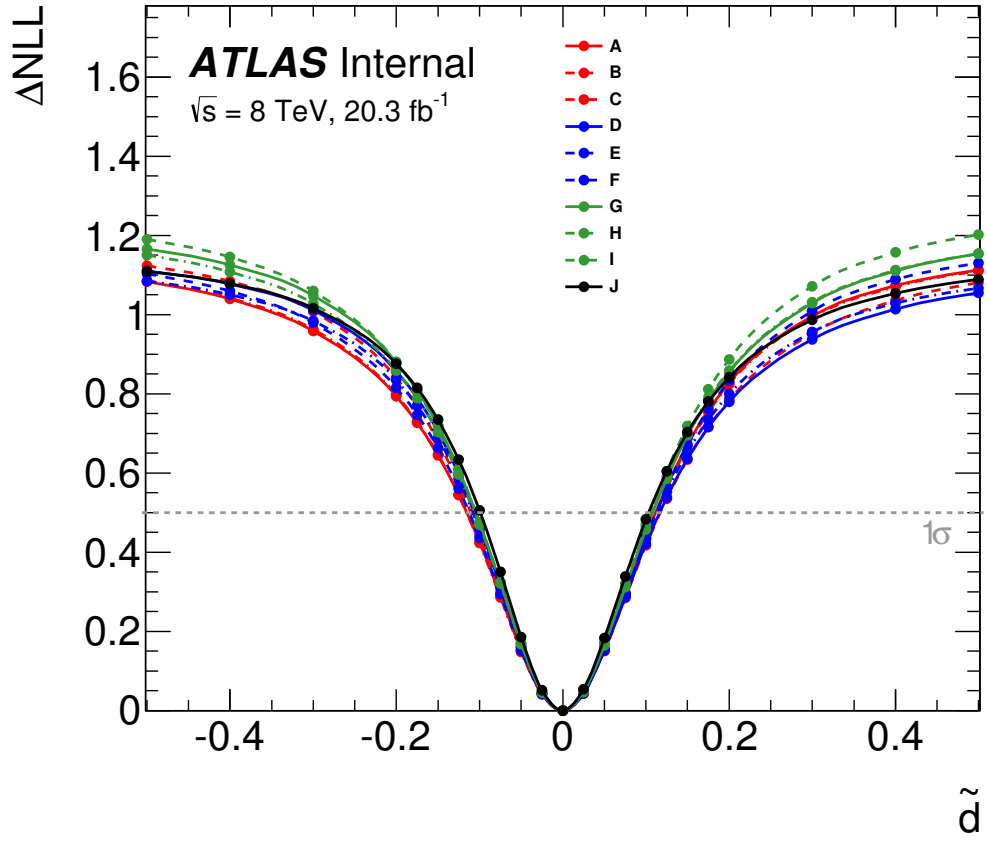


Figure 4.19: ΔNLL curves produced for $BDT_{score} > 0.95$ for the binning choices shown in figures 4.16 to 4.18. Binning option J corresponds to the binning used in the final analysis.

Variable	Region	Bin edges
Optimal observable O_1	Signal region	-15,-5,-2.5,0,2.5,5,15
	Top CR	Only one bin
	$Z \rightarrow \ell\ell$ CR	Only one bin
BDT score	Low BDT CR	-1, -0.79, -0.58, -0.37, -0.16, 0.05

Table 4.11: *Histogram binnings in the final log likelihood fit in $\tau_\ell\tau_\ell$*

The $\tau_\ell\tau_\ell$ channel

The $\tau_\ell\tau_\ell$ channel targets all final states with exactly two isolated leptons ($ee, \mu\mu$ or $e\mu$) with opposite charge. The definitions of the pre-selection and signal regions are listed in table 4.12. A selection criterion of $\text{BDT}_{\text{score}} > 0.68$ is added according to the same considerations as in the $\tau_\ell\tau_{\text{had}}$ channel, and an identical limit on the Optimal Observable of $|O_1| < 15$ is set. The chosen signal region corresponds to the three most signal-like bins of the $\text{BDT}_{\text{score}}$ distribution, which is a reasonable choice in terms of maximising the significance while avoiding too little statistics in the background samples. The lower value of the $\text{BDT}_{\text{score}}$ selection cut in $\tau_\ell\tau_\ell$ compared to $\tau_\ell\tau_{\text{had}}$ is mainly due to the lower expected signal and background yields in this channel. A significance-based bin optimisation of the Optimal Observable distribution in the signal region yields results very similar to those found in $\tau_\ell\tau_{\text{had}}$. Based on the similarities, the decision was made to use the same binning in both channels. Table 4.11 summarises the histogram binning choices in the $\tau_\ell\tau_\ell$ channel.

The expected event yields in the VBF category and in the high BDT signal region for each background and the signal component are listed in table 4.14. The signal component contains 5.21 ± 0.04 events from VBF $H \rightarrow \tau\tau$ and 1.04 ± 0.04 events from VBF $H \rightarrow WW$ in the signal region. CP-mixing VBF $H \rightarrow WW$ signal samples are generated by re-weighting the pure SM sample equivalently to VBF $H \rightarrow \tau\tau$. Anomalous couplings in the decay of VBF $H \rightarrow WW$ are expected to have a negligible impact on this production mode sensitive analysis as shown in appendix C. The modelling of the Optimal Observable in comparison to data can be reviewed in several control regions. The low $\text{BDT}_{\text{score}}$ region is defined by a cut on $\text{BDT}_{\text{score}} < 0.05$, which are the 10 most background-like bins of the $\text{BDT}_{\text{score}}$ distribution, in order to reduce the signal contribution and exclude bins in an intermediate range with low statistics. Furthermore, the original $\text{BDT}_{\text{score}}$ is re-binned by a factor of two to minimize the influence of statistical fluctuations. The low $\text{BDT}_{\text{score}}$ control region is dominated by events from $Z \rightarrow \tau\tau$ decays. As in the couplings analysis, further control regions for processes including top quarks and $Z \rightarrow \ell\ell$ decays are defined by inverting the b -veto and excluding events outside

the Z mass window in the VBF category respectively. The O_1 distributions in these control regions are shown in figure 4.20. A summary of the control region definitions can be found in table 4.13, and an illustration of the fit model can be found in figure 4.21. The background model is in good agreement with the data in all control regions.

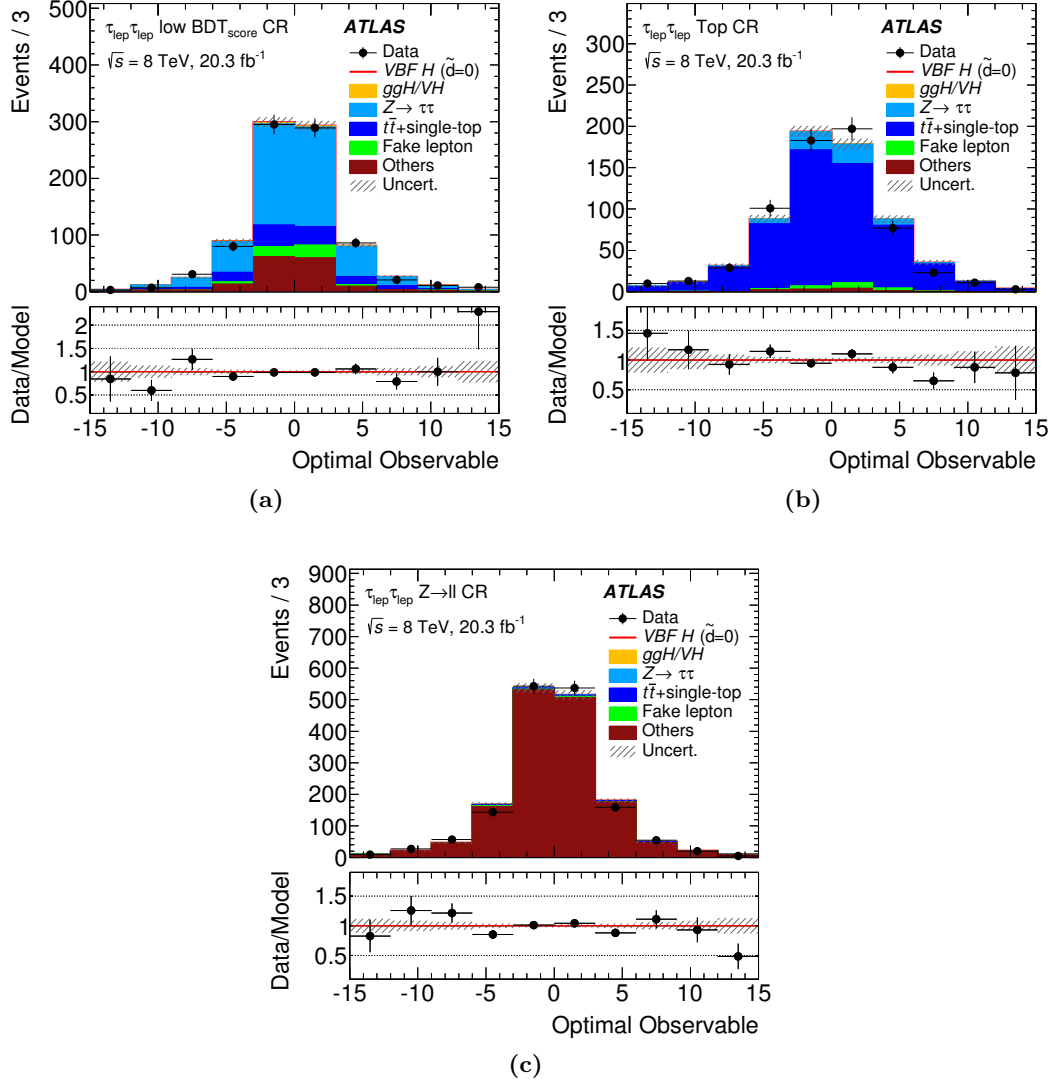


Figure 4.20: Distributions of O_1 in the low BDT control region (a), top control region (b) and $Z \rightarrow \ell\ell$ control region (c) in the $\tau_\ell\tau_\ell$ channel. The error bands indicate only statistical uncertainties.

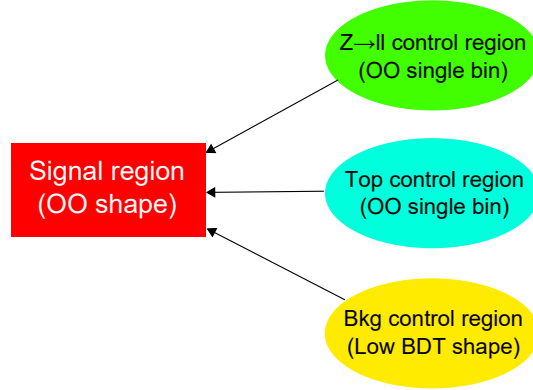


Figure 4.21: Illustration of the kinematic regions included in the $\tau_\ell\tau_\ell$ fit model. The normalisations of the $Z \rightarrow \tau\tau$, $Z \rightarrow ll$ and top backgrounds are free in the fit and are constrained by the low BDT_{score} CR, $Z \rightarrow ll$ CR and top CR respectively.

Channel	Pre-selection
$\tau_\ell\tau_\ell$	<p>Exactly two isolated opposite-sign leptons</p> <p>Events with τ_{had} candidates are rejected</p> <p>$30 \text{ GeV} < m_{\tau\tau}^{\text{vis}} < 100 \text{ (75) GeV}$ for DF (SF) events</p> <p>$\Delta\phi_{\ell\ell} < 2.5$</p> <p>$E_T^{\text{miss}} > 20 \text{ (40) GeV}$ for DF (SF) events</p> <p>$E_T^{\text{miss,HPTO}} > 40 \text{ GeV}$ for SF events</p> <p>$p_T^{\ell_1} + p_T^{\ell_2} > 35 \text{ GeV}$</p> <p>Events with a b-tagged jet with $p_T > 25 \text{ GeV}$ are rejected</p> <p>$0.1 < x_{\tau_1}, x_{\tau_2} < 1$</p> <p>$m_{\tau\tau}^{\text{coll}} > m_Z - 25 \text{ GeV}$</p>
$\tau_\ell\tau_{\text{had}}$	<p>Exactly one isolated lepton + one medium τ_{had} candidate with opposite charges</p> <p>$m_T < 70 \text{ GeV}$</p> <p>Events with a b-tagged jet with $p_T > 30 \text{ GeV}$ are rejected</p>
Channel	Signal region selection cuts
$\tau_\ell\tau_\ell$	<p>At least two jets with $p_T^{j_1} > 40 \text{ GeV}$ and $p_T^{j_2} > 30 \text{ GeV}$</p> <p>$\Delta\eta(j_1, j_2) > 2.2$</p> <p>$BDT_{\text{score}} > 0.68$</p> <p>$O_1 < 15$</p>
$\tau_\ell\tau_{\text{had}}$	<p>At least two jets with $p_T(j_1) > 50 \text{ GeV}$ and $p_T(j_2) > 30 \text{ GeV}$</p> <p>$\Delta\eta(j_1, j_2) > 3.0$</p> <p>$m_{\tau\tau}^{\text{vis}} > 40 \text{ GeV}$</p> <p>$BDT_{\text{score}} > 0.95$</p> <p>$O_1 < 15$</p>

Table 4.12: Summary of the pre-selection and the signal region selection used in the analysis, per channel. For definitions of the various quantities, see sections 2.3.1 and 1.3.2.

Region	$\tau_\ell \tau_\ell$	$\tau_\ell \tau_{\text{had}}$
$Z \rightarrow \ell\ell$	$80 < m_{\tau\tau}^{\text{vis}} < 100 \text{ GeV}$ same-flavour events	
Top	Invert b -jet veto	Invert b -jet veto $m_T > 40 \text{ GeV}$
Low $\text{BDT}_{\text{score}}$	$\text{BDT}_{\text{score}} < 0.05$	$\text{BDT}_{\text{score}} < 0.95$

Table 4.13: Definitions of the control regions used in the analysis per channel. The CRs are given as changes with respect to the VBF category.

Process	VBF region		Signal region	
	$\tau_\ell \tau_\ell$	$\tau_\ell \tau_{\text{had}}$	$\tau_\ell \tau_\ell$	$\tau_\ell \tau_{\text{had}}$
Data	1014 ± 32	2830 ± 53	54 ± 7	68 ± 8
$Z \rightarrow \tau\tau$	601.2 ± 7.9	900 ± 17	19.0 ± 1.5	20.8 ± 2.7
Fake bkg	60.8 ± 4.1	1637 ± 13	2.2 ± 0.7	24.3 ± 1.6
Top	142.7 ± 6.4	84.5 ± 4.8	3.1 ± 0.8	4.7 ± 1.4
$Z \rightarrow \ell\ell + \text{Diboson}$	199.5 ± 10.0	119.1 ± 5.6	10.5 ± 1.6	6.3 ± 1.1
Non-VBF $H \rightarrow \tau\tau/WW$	10.7 ± 0.3	12.4 ± 0.6	1.6 ± 0.1	2.6 ± 0.3
Total bkg	1015 ± 15	2753 ± 23	36.4 ± 2.4	58.8 ± 3.6
VBF $H \rightarrow \tau\tau + \text{VBF } H \rightarrow WW$	12.8 ± 0.1	21.7 ± 0.2	6.3 ± 0.1	11.1 ± 0.1
S/B	≈ 0.013	≈ 0.008	≈ 0.17	≈ 0.19

Table 4.14: Expected pre-fit event yields in the VBF region and in the final signal region including the cut on high BDT values per process and per channel.

Summary

In this chapter the relevant signal and background processes and their modelling as well as the kinematic regions of interest have been established. Signal re-weighting has been introduced to produce VBF signal kinematic distributions reflecting CP-odd admixtures to the HVV couplings through the parameter \tilde{d} . The dataset collected by ATLAS in 2012 has also been described, including the chosen triggers and object quality criteria. Background suppression using multivariate BDT discriminators has been discussed, and the specific implementation in terms of the choice of training variables and samples has been specified. Additionally, the expected signal and background yields have been listed alongside plots of the distributions of the Optimal Observable in the regions of interest. The next chapter will introduce the fit procedure used to extract limits on \tilde{d} followed by the treatment of systematic uncertainties and how they enter the fit.

Chapter 5

Fit Model and Systematic Uncertainties

The fit model for the VBF $H \rightarrow \tau\tau$ CP analysis takes as input the variable distributions in the regions described in chapter 4. It is performed by comparing the collected data to the combined signal and background model by maximising a log-likelihood expression formed from the input distributions including their statistical and systematic errors. By performing multiple fits using signal hypotheses with different CP-odd coupling strengths in the HVV vertex determined by the parameter \tilde{d} , a confidence interval on \tilde{d} can be obtained. Section 5.1 describes the construction of the log-likelihood expression and explains how limits on \tilde{d} can be extracted. A pruning and smoothing procedure described in section 5.1.1 is applied to the uncertainties with the goal of improving fit stability and removing negligible contributions.

The sources of systematic uncertainties that affect the analysis arise from imperfect knowledge of various parameters and properties of the ATLAS detector as well as lack of knowledge of the underlying physics. These include uncertainties on the nominal integrated luminosity, trigger efficiencies and E_T^{miss} as well as energy scales and resolutions that affect the object reconstruction. The dominant systematics include uncertainties on the derived jet energy scale and tau energy scale. These experimental systematics are listed and described in section 5.2. Additionally, a number of theoretical uncertainties apply to the modelling of simulated samples used to estimate the expected signal and background contributions. These are treated in section 5.4. Each systematic can result in uncertainties on the normalisation and/or shape of the sample distributions, and these two effects are treated separately in the fit. The estimation of systematics is based on the work performed in the couplings analysis [2] while modifying and adding extra uncertainties where necessary, i.e. when differences in the choice of variables and signal regions influence how the sources of uncertainties affect the analysis.

5.1 Description of the Log-likelihood Fit

In order to extract competitive limits on the CP-odd mixing parameter \tilde{d} it is necessary to estimate the sensitivity to CP-odd couplings using a maximum-likelihood fit to the Optimal Observable distribution instead of only its mean. A prescription for how to estimate confidence intervals on the CP-odd mixing strength \tilde{d} is described below. Further details can also be found in [173]. The fitting procedure relies on HistFactory [174] for the construction of a likelihood function from input histograms, and RooStats [175] for the PDF creation and fit optimisation.

The statistical analysis employs a likelihood function $\mathcal{L}(\mathbf{x}; \epsilon, \boldsymbol{\theta})$, where \mathbf{x} denotes the set of measured data, ϵ is the set of parameters of interest, and $\boldsymbol{\theta}$ is the set of nuisance parameters. The nuisance parameters encompass statistical and systematic uncertainties that are able to vary in the fit. Nuisance parameters are treated using a log-normal Gaussian description [176]. The fit aims to accurately estimate the values of the parameters of interest and their uncertainties. This analysis defines only one parameter of interest describing the signal strength μ , where a value of $\mu = 1$ corresponds to the SM signal expectation. The aim of the analysis is to perform a test of CP invariance on the data, which means that only CP-odd observables are of interest. The cross section, i.e. the signal event rate, is a CP-even observable and does therefore not contribute to the sensitivity of the CP test. The rate information can be used to probe the more general tensor structure of the HVV coupling, but this is outside scope of a CP test. Hence, only the shape information from the Optimal Observable distribution is relevant when estimating the sensitivity of this analysis. It is not possible to perform the log-likelihood fit directly to \tilde{d} since there is no smooth prediction of the shape of the Optimal Observable distribution in relation to this variable, only distinct re-weighted shapes at discrete values.

The approach of a binned likelihood function with an underlying model of signal plus background is chosen, which is a product of Poisson probability terms for each bin in the histograms from all input regions

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{i \in \text{bins}} \frac{\lambda_i(\mu, \boldsymbol{\theta})^{n_i}}{n_i!} e^{-\lambda_i(\mu, \boldsymbol{\theta})}, \quad (5.1)$$

where n_i is the observed number of events in bin i and $\lambda_i(\mu, \boldsymbol{\theta})$ is the expected number of events in bin i depending on the parameter of interest and the set of nuisance parameters.

A set of signal samples corresponding to different CP-odd mixing strengths \tilde{d} are created by re-weighting the purely CP-even (VBF) $H \rightarrow \tau\tau$ signal sample produced through ATLAS simulation, as described in section 4.3. The likelihood

function can then be calculated in each point of \tilde{d} for the corresponding CP-mixing model to produce a likelihood curve. In the large sample limit one can utilise the fact that the negative logarithm of the likelihood function NLL allows for directly reading off the central confidence interval $[\hat{\tilde{d}} - \sigma_{\hat{\tilde{d}}}, \hat{\tilde{d}} + \sigma_{\hat{\tilde{d}}}]$ using the approximate correspondence

$$-\log\mathcal{L}(\hat{\tilde{d}} \pm \sigma_{\hat{\tilde{d}}}) = -\log\mathcal{L}_{\max} + \frac{1}{2}. \quad (5.2)$$

After constructing the NLL curve by calculating the NLL value for each \tilde{d} hypothesis and a specific dataset \mathbf{x} , the 68.3% central confidence interval can be determined from the best estimator $\hat{\tilde{d}}$, at which the NLL curve is minimal, by reading off the $\Delta\text{NLL} = \text{NLL} - \text{NLL}_{\min}$ at 0.5. This interval should contain the true value of \tilde{d} in 68.3% of all cases. The expected confidence interval can be determined by constructing a pseudo-dataset \mathbf{x} containing the background plus signal for a particular value of \tilde{d} , also called an Asimov dataset. In the case of simulated pure CP-even SM signal, the ΔNLL curve is expected to have its minimum at $\tilde{d} = 0$.

The normalisation of the CP-mixed or CP-even (VBF) $H \rightarrow \tau\tau$ signal sample in each fit is always described by the same free-floating parameter μ , i.e. this analysis does not take into account any information about the relative cross sections of CP-mixing scenarios predicted by the signal re-weighting, as mentioned in section 1.3.1. Only the shape of the Optimal Observable is different in each scenario. Higgs production through other processes (like gluon fusion, or associated with a vector boson) is normalised to the SM prediction with the corresponding uncertainties, as is the small amount of contamination from WW decays of the Higgs boson in $\tau_\ell\tau_\ell$.

The full fit model per decay channel consists of a signal region in addition to several control regions that are included in order to constrain background normalisations and nuisance parameters. In both the $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ channels the signal region is defined as the high $\text{BDT}_{\text{score}}$ region in the respective channel. Both channels use a low $\text{BDT}_{\text{score}}$ region as well as a top quark dominated control region in order to better constrain the background normalisations. Furthermore, $\tau_\ell\tau_\ell$ includes an additional control region enriched with events from $Z \rightarrow \ell\ell$ decays, since its contribution to the background composition in this channel is non-negligible. The signal and control regions are defined for each decay channel in section 4.6.2.

5.1.1 Pruning and Smoothing of Nuisance Parameters

The impact of every systematic on the fit is split into two parts, normalisation and shape, each with its own nuisance parameter (NP). The term nuisance parameter

refers to the fact that it is not of immediate interest, but must be accounted for in the analysis. The normalisation NP moves the total distribution up or down and hence has the same effect on all bins. Systematics relating to e.g. TES, JES or E_T^{miss} also act as shape NPs in the fit that change the individual bin contents while preserving the normalisation.

Several of the processes in the fit are relatively low in statistics, especially within the signal region. Therefore, small upward and downward variations are potentially dominated by statistical fluctuations. Including these noisy variations can potentially cause fit instabilities or an over-estimation of the impact stemming from those particular NPs. In each individual process the normalisation uncertainty is included provided that it is larger than 0.5%. Lower normalisation uncertainties are deemed below the noise threshold and are excluded from the fit. An algorithmic pruning and smoothing procedure is introduced in order to minimise unwanted effects from shape variations. The following treatment is applied to shape variations to a process in every region before adding them to the fit:

- *χ^2 pruning*: For each shape systematic, the upward and downward fluctuations of the shape are compared to the nominal histogram by performing a χ^2 test yielding a probability of the nominal and varied histogram being compatible. For each bin, only the largest of the statistical uncertainties of the nominal or varied histogram enters the χ^2 value, since they are usually highly correlated. The shape systematic is accepted into the fit if the χ^2 test of either the upward or downward variation returns a probability less than a threshold value that differs between the decay channels. For $\tau_\ell\tau_{\text{had}}$ this threshold is $p < 0.98$ and was determined in the couplings analysis through a careful study of how it impacts fit stability while requiring that it has a negligible impact on the sensitivity.
- *Smoothing*: The ratio of variation to nominal (separately for upwards and downwards variations) is smoothed using the TH1::SMOOTH(1) method of ROOT [3]. The smoothed varied shape is then obtained by multiplying the nominal with the smoothed ratio. The reason for smoothing the ratio rather than the varied shape directly is that the BDT distribution can be genuinely (i.e. not due to statistical noise) strongly varying, which may cause the smoothing method to over-smooth it, while the ratio should be a relatively smooth and non-rapidly changing function in the absence of noise.
- *Reflection*: The up and down variations are normalised to the nominal histogram to account for changes made in the previous steps. In bins where both

variations are above or below the nominal value the smaller variation is reflected such that it lies the same distance away from nominal in the opposite direction. Finally, the variation histograms are again normalised to nominal.

5.2 Experimental Uncertainties

Accurate reconstruction of physics objects depends on the performance of the ATLAS detector subsystems. Luminosity estimation, momentum and energy calibration, E_T^{miss} reconstruction and the efficiencies of the trigger simulation as well as particle identification and reconstruction are all associated with uncertainties. Their effects are propagated to the analysis by varying the corresponding parameters in the MC event reconstruction, which results in a new event yield with altered reconstructed particle dynamics and energy deposits that can subsequently be compared to the nominal case. In some cases data-driven background modelling can reduce the severity of such uncertainties by extracting event parameters directly from data, but this introduces its own set of uncertainties. The uncertainties associated with the data-driven modelling of $Z \rightarrow \tau\tau$ and fakes are discussed in section 5.3. The following lists the major sources of systematic uncertainties:

Luminosity

The uncertainty on the integrated luminosity of the full 2012 ATLAS dataset has been derived following the methodology outlined in [177]. Several subdetectors in ATLAS can be used to measure the visible interaction rate per bunch crossing $\mu_{\text{vis}} = \epsilon\mu$, where ϵ is the efficiency of the subdetector. Using the corresponding visible cross-section $\sigma_{\text{vis}} = \epsilon\sigma$ the luminosity can be expressed as

$$\mathcal{L} = \frac{\mu_{\text{vis}} n_b f_r}{\sigma_{\text{vis}}}, \quad (5.3)$$

where f_r is the revolution frequency in the accelerator ring and n_b is the number of bunch pairs colliding per revolution. The visible cross-section can be estimated using dedicated beam-separation, or van der Meer (vdM), scans [177] that give an estimate of the total luminosity as

$$\mathcal{L} = \frac{n_b f_r n_1 n_2}{2\pi \Sigma_x \Sigma_y}, \quad (5.4)$$

where n_1 and n_2 denote the number of protons per bunch in beam 1 and 2 respectively, and Σ_x and Σ_y represent the horizontal and vertical beam widths. By

combining equations 5.3 and 5.4 the visible cross-section can be obtained for a particular subdetector using the measured visible interaction rate during the vdM scan $\mu_{\text{vis}}^{\text{vdM}}$ as

$$\sigma_{\text{vis}} = \mu_{\text{vis}}^{\text{vdM}} \frac{2\pi \Sigma_x \Sigma_y}{n_1 n_2}. \quad (5.5)$$

Equation 5.3 can then be used to monitor the luminosity in physics runs. Systematic uncertainties on both the vdM scans and individual luminosity detectors are taken into account, resulting in a total estimated uncertainty of $\pm 2.8\%$.

Tau Energy Scale and Resolution

Energy scale uncertainties arise due to incomplete knowledge of the calorimeter response to jets and particles. There can also be components from MC mis-modelling. While electromagnetic showers are a result of cascading pair production and bremsstrahlung, hadronic showers are complicated by large per-event fluctuations of their width and depth as well as decays to neutrinos and muons that reduce the energy deposited in the calorimeters. These complexities reduce the energy resolution of reconstructed hadronic energy deposits compared to electromagnetic showers. The resulting variation in the energy from such uncertainties can lead to changes to the acceptance of a selection category, which in turn changes the shape of variable distributions. Secondary variables such as E_T^{miss} also need to be recalculated for each variation.

The tau energy scale (TES) refers to the calorimeter response to hadronically decaying taus (see also section 3.6). The TES uncertainty has several components including the calorimeter energy response to the τ_{had} decay products, the detector model used in the ATLAS simulation, the underlying event model included in the event generator and uncertainties inherent to the energy calibration itself [122]. The method of TES uncertainty estimation depends on the tau p_T . At $p_T < 50$ GeV an in-situ correction is employed by fitting the reconstructed visible mass from $Z \rightarrow \tau\tau$ events in data. For higher tau p_T (outside the Z peak) where the statistics are too low to employ this method, a decomposition procedure is needed. In this case each tau lepton is decomposed into its decay products and the calorimeter responses are combined according to the tau branching ratios to different final states. The total TES uncertainty is estimated to be of the order $\pm 2\text{-}4\%$ [122]. The tau energy resolution as a function of the true visible tau momentum ranges from 20% at low tau momenta to below 10% at energies above a few hundred GeV. Comparisons of the tau energy resolutions in simulations with two different hadronic shower models and

underlying event tunes show a difference of less than 1% [126], and this uncertainty is propagated to the $H \rightarrow \tau\tau$ CP analysis by applying a Gaussian smearing.

Jet Energy Scale and Resolution

The jet energy scale (JES) uncertainty impacts both normalisation and shape and has a number of sub-components [178]. Each component was evaluated in terms of its relative impact, and only components with a contribution larger than 10% of the leading component and being at least twice as large as the statistical uncertainty on the shift between nominal and systematically varied yields were included in the final fit. This component reduction was performed independently of and preceding the pruning procedure described in section 5.1.1. Apart from avoiding the inclusion of statistical noise it also assisted in minimising the run-time of the analysis since each JES component necessitates a separate event selection step. The two largest background contributions coming from $Z \rightarrow \tau\tau$ and QCD fakes are modeled using data and are therefore not affected by JES uncertainties. The important sources of JES uncertainty are:

- *In-situ jet energy correction:* A collection of sub-components accounting for bin-to-bin correlations in the in-situ calibration and corrections of jets alongside components describing the detector response and MC jet modelling.
- *η intercalibration:* Jets in the forward detector regions are intercalibrated with jets from a central reference region using di-jet events to account for effects not included in MC calibration. Modelling and statistical uncertainty components have been evaluated.
- *Flavour composition and response:* Quark-initiated and gluon-initiated jets have different calorimeter responses, and knowledge of the quark-gluon composition of jets is limited. Two channel-dependent nuisance parameters are added for light quark flavours since the jets can be more gluon-dominated or quark-dominated depending on the production channel. Jets containing a truth b -quark are treated separately with another nuisance parameter that acts only as a normalisation uncertainty.
- *Pile-up:* Uncertainties on the JES correction due to in-time and out-of-time pileup, parametrised by the number of primary vertices and the average number of interactions per bunch crossing respectively.

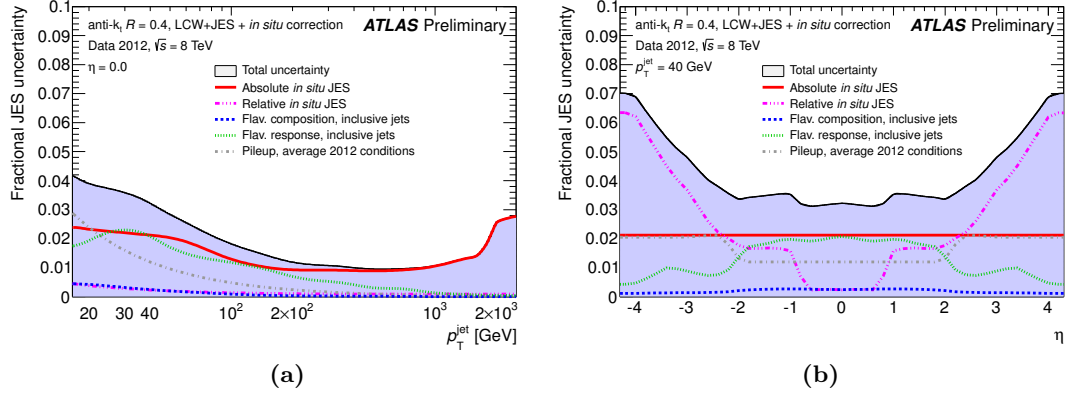


Figure 5.1: Components of the fractional JES uncertainty estimated for data taken in 2012 as a function of (a) transverse momentum and (b) pseudorapidity [180].

The combined JES uncertainty and its components in 2012 data are shown in figure 5.1 as a function of the transverse momentum of the jets and their pseudorapidity. The jet energy resolution (JER) was measured in dijet events and is well-modelled in MC with relative uncertainties below 10% [179].

Lepton Energy Scale and Resolution

Analogously to taus and jets, electrons are subject to energy scale uncertainty and resolution. However, these are determined to a much higher precision. The electron energy scale uncertainty is typically well below $\pm 1\%$ for electrons with a transverse energy $E_T > 20$ GeV in the barrel region and up to $\pm 1\%$ in the endcap region [118]. The electron energy resolution has a relative inaccuracy of less than 10% for electrons up to 60 GeV transverse energy [181]. The muon momentum scale has an estimated uncertainty below 0.2%, while the momentum resolution is 1.7% at central rapidities and $p_T \sim 10$ GeV, growing to 4% at large rapidities and $p_T \sim 100$ GeV [182].

Missing Transverse Energy

The E_T^{miss} is calculated using inputs from all physics objects in an event, and its uncertainty is therefore dependent on the energy scales and resolutions of these objects. Each systematic variation on the inputs results in a recalculated E_T^{miss} . A separate systematic uncertainty is estimated on the soft term in the E_T^{miss} that takes into account low p_T objects that are not part of the nominal terms in the E_T^{miss} calculation. This uncertainty is studied in $Z \rightarrow \mu\mu$ events without jets where only the muons and the soft term contribute to the E_T^{miss} , providing an optimal environment

for studying the MC modelling of the soft term. Systematic uncertainties on the soft term are estimated to be $\pm 8\%$ for the scale and $\pm 5\%$ for the resolution [127]. However, the impact on the total E_T^{miss} is below 1%.

Trigger Efficiencies

The electron and muon trigger efficiencies have associated uncertainties. MC scale factors are derived through tag-and-probe studies to correct possible mismodellings of the trigger efficiencies, and uncertainties are estimated for these scale factors [120, 183]. The scale factor uncertainties depend on η and ϕ in the detector and are included in the analysis by propagating them to the trigger efficiencies. For electrons the size of the scale factor uncertainties are relatively small and range from $\pm 2\%$ for electrons at large pseudo-rapidities and low transverse momenta to $\pm 1\%$ for electrons in the barrel region above $E_T > 28$ GeV. Uncertainties on the muon scale factors are of the order $\pm 1\%$.

Lepton Reconstruction and Identification

Uncertainties are estimated on the scale factors related to the algorithms identifying and reconstructing leptons. The signal samples have lepton uncertainties of approximately $\pm 2\%$ [115, 182].

Hadronic Tau Identification

Hadronically decaying taus have uncertainties originating from identification and reconstruction [122] but are not used as trigger objects. Scale factors and uncertainties are estimated through tag-and-probe studies on $Z \rightarrow \tau\tau$ events and are only applied to tau candidates matched to a true tau within $\Delta R < 0.2$. Electrons faking taus also have associated corrections with uncertainties that are only applied to taus truth matched to an electron. Uncertainties due to hadronic tau identification and reconstruction are $\pm 3\%$ on the signal samples and $\pm 2\%$ on backgrounds.

b -tagging

The uncertainties due to b -tagging depend on the choice of hadronisation model, the PDFs and experimental uncertainties. They are derived as systematic variations of the scale factors used to correct simulated samples to match with experimental data, see figure 3.4. Uncertainties range from $\pm 6\%$ at low transverse momenta to $\pm 2\%$ at transverse momenta in the range of 60 to 140 GeV. The mistag rate of light flavour

jets has a larger systematic uncertainty, depending on p_T and η , between $\pm 15\%$ and $\pm 25\%$ [184].

5.3 Data-driven Background Modelling Uncertainties

The treatment of systematics in the fake tau background, containing events where the tau lepton is faked by a jet, is greatly simplified using the fake factor method (see section 4.5.2). All components of the events are taken from data, removing the need to simulate them. However, the data-driven fake background estimation introduces its own set of systematics. The three significant components are:

- Limited statistics in the control regions used to estimate the fake factors leading to a statistical uncertainty on the fake factors themselves. This is propagated and applied alongside the fake factors to the anti tau data events. The impact on the number of expected fake events is of the order of 5%.
- The background composition given by the fraction R_i of each contributing process extracted from the simulation. The impact R_i bias was estimated using two methods giving compatible results. The first method performs an ad hoc variation of each R_i between $R_i/2$ and $2R_i$ giving an uncertainty of roughly 5%. The second method looks at differences in R_W between W -rich (high M_T and high p_T^l) and W -depleted (low M_T and low p_T^l) subsets of the anti-tau signal region. The impact on the fake tau background is again estimated to be approximately 5%.
- A closure test performed on MC and a separate closure test on data in the same-sign control region. Here a closure test means deriving and applying the fake factors in the same dataset, e.g. the simulation. If plotting a variable included in the FF dependence, namely p_T^τ , the agreement will be perfect by construction. By looking at other observables, it is possible to probe the effect of only deriving the FF as a function of a few variables. The closure test looks at distributions of e.g. the MMC ditau mass and the BDT distribution. Additionally, a final closure test is performed on data in the same sign control region. No systematic bias is observed in either test.

The embedding technique used to model the $Z \rightarrow \tau\tau$ background extracts everything except the tau lepton decays from data. The uncertainties on the leptons, jets and E_T^{miss} are therefore already included, and only uncertainties related to the $Z \rightarrow \mu\mu$ event selection and muon energy subtraction need to be estimated. The

isolation criteria on the $Z \rightarrow \mu\mu$ event selection were tightened and loosened to obtain a variation on the efficiency. The calorimeter energy subtraction was varied by $\pm 20\%$ for the embedding samples. This resulted in an uncertainty of $\pm 1.5\%$.

5.4 Theoretical Uncertainties

Theoretical normalisation and shape uncertainties have already been estimated in the couplings analysis for the VBF region defined in table 4.12. However, the addition of a selection criterion on the $\text{BDT}_{\text{score}}$ potentially introduces another uncertainty component due to event migration in the new signal region. Also, the use of a different observable (Optimal Observable or $\Delta\phi_{jj}^{\text{sign}}$ instead of $\text{BDT}_{\text{score}}$) requires the shape uncertainties to be derived again. In some cases the uncertainty was derived on the full signal region (VBF region + $\text{BDT}_{\text{score}}$ cut), while in other cases the VBF region uncertainty was re-confirmed and an additional contribution from the $\text{BDT}_{\text{score}}$ cut added. Each theoretical systematic in the $\tau_\ell\tau_{\text{had}}$ channel will be described below. The treatment and addition of extra systematics were performed identically in $\tau_\ell\tau_\ell$.

QCD Scale Uncertainties

Higgs production channels have uncertainties related to higher order QCD corrections that are not included in the cross sections in MC. These QCD scale uncertainties are estimated by varying the renormalisation and factorisation scales by a factor of two around the nominal scale in the generator with the constraint

$$\frac{1}{2} \leq \frac{\mu_F}{\mu_R} \leq 2. \quad (5.6)$$

Appropriate parton level kinematics cuts on the jets were applied to the samples beforehand. The normalisation uncertainty due to the QCD scale for the VBF region was evaluated to be 2.1% in VBF production and 24% in gluon fusion production.

In VBF production, possible shape and normalisation uncertainties on the Optimal Observable/ $\Delta\phi_{jj}^{\text{sign}}$ distribution when including the $\text{BDT}_{\text{score}}$ selection criterion were evaluated at truth level using MG5_AMC@NLO and varying the renormalisation and factorisation scales. No noticeable effect was observed beyond statistical fluctuations, so no further uncertainty was assigned. In gluon fusion, to evaluate whether an additional normalisation uncertainty needed to be applied to account for the cut on the $\text{BDT}_{\text{score}}$, a comparison was made between the default H+1jet MINLO sample, and an alternative H+2jets MINLO sample. The acceptance

of the cut on the $\text{BDT}_{\text{score}}$ was compared, relative to the VBF region selection (i.e. $N_{\text{SR}}/N_{\text{VBF region}}$). The difference between the samples represents the uncertainty on the LO description of the second jet in the default sample. This acceptance was the same for both samples within statistical uncertainties. Therefore no additional uncertainty due to this was applied beyond the QCD scale uncertainty already evaluated in the context of the couplings analysis, for the acceptance of the VBF region selection. To account for the uncertainty in the shape of the Optimal Observable/ $\Delta\phi_{jj}^{\text{sign}}$, the distribution for the default H+1jet sample was compared to that from the alternative H+2jets MINLO sample. The difference was symmetrised and taken as a systematic uncertainty. The resulting shape variations are shown in figure 5.2.

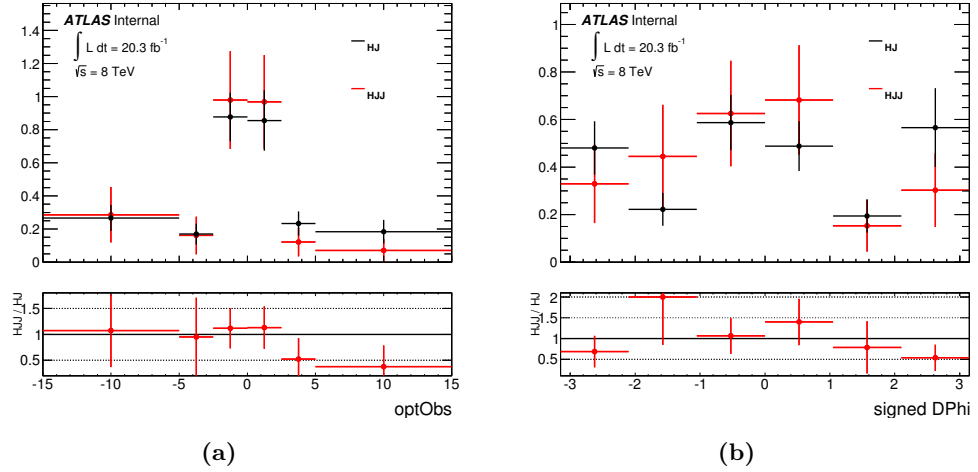


Figure 5.2: QCD scale shape uncertainty on the gluon fusion ggF events in the final signal region introduced by the $\text{BDT}_{\text{score}}$ selection cut for the Optimal Observable (a) and $\Delta\phi_{jj}^{\text{sign}}$ (b). The uncertainty was estimated by comparing distributions of MINLO samples containing $H + 2j$ and $H + 3j$ events respectively.

Modelling of Underlying Event/Parton Shower

The chosen underlying event and parton showering model affects the acceptance in simulated VBF and ggF samples, defined as the ratio of the number of reconstructed events in a given region and the number of events with the corresponding truth-level definition in the same region. To estimate the size of the variation, the acceptance was compared between POWHEG+PYTHIA and POWHEG+HERWIG. The parton shower simulations employ different models for multi-parton interactions and can be used to estimate the associated uncertainty. PYTHIA incorporates the Lund string

fragmentation model [185] that treats the gluon field between quarks as a massless relativistic string with a linear potential that may split if it has enough energy to produce new quark pairs. HERWIG uses cluster hadronisation [186] describing colourless clusters characterised by their mass and flavour content that subsequently decay to the final state particles.

In order to properly compare POWHEG+PYTHIA and POWHEG+HERWIG, some corrections need to be turned off. The Higgs p_T re-weighting for ggF has been calculated by comparing POWHEG+PYTHIA to the result at NNLO+NNLL using HRES. Therefore it is incorrect to apply this correction to POWHEG+HERWIG. The electroweak correction on the event weight in VBF production has likewise been calculated using POWHEG+PYTHIA, and also has to be turned off.

The acceptance for the $\text{BDT}_{\text{score}}$ cut relative to the VBF region selection was conservatively added in quadrature with the corresponding uncertainty on the acceptance for the VBF region selection, which was evaluated (with the same samples) in the context of the couplings analysis. This procedure was followed for both the VBF and ggF processes. The effect on the shape of the Optimal Observable/ $\Delta\phi_{jj}^{\text{sign}}$ was also evaluated, and the relative bin-by-bin difference was symmetrised and applied as a shape uncertainty. The shapes can be seen in figures 5.3 and 5.4.

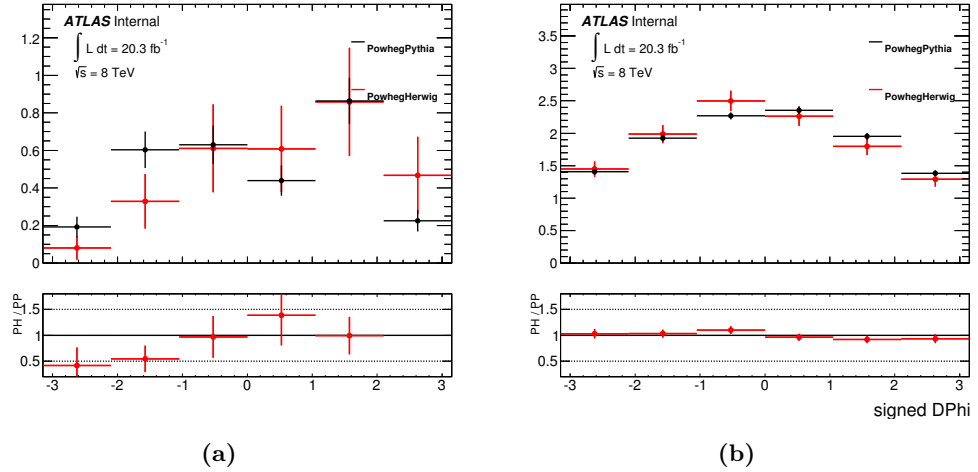


Figure 5.3: Underlying Event/Parton Shower shape uncertainty on the $\Delta\phi_{jj}^{\text{sign}}$ distribution of (a) ggF and (b) VBF signal events in the final signal region introduced by the $\text{BDT}_{\text{score}}$ selection cut. The uncertainty was estimated by comparing distributions of samples generated with POWHEG+PYTHIA and POWHEG+HERWIG respectively.

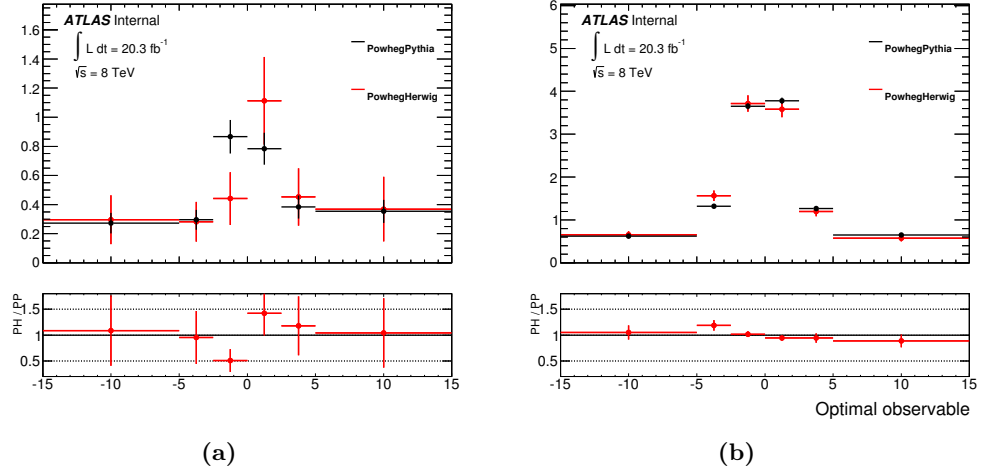


Figure 5.4: Underlying Event/Parton Shower shape uncertainty on the O_1 distribution of (a) ggF and (b) VBF signal events in the final signal region introduced by the BDT_{score} selection cut. The uncertainty was estimated by comparing distributions of samples generated with POWHEG+PYTHIA and POWHEG+HERWIG respectively.

PDF Uncertainties

Uncertainties on the PDFs lead to an acceptance uncertainty. The normalisation uncertainty was evaluated by re-weighting the PDF using the eigentunes of the default CT10 PDF, as well as two alternative PDFs (MSTW and NNPDF). Eigentunes are a collection of deviation tunes that represent the uncertainties compared to the best tune and are created from the covariance matrix of the tuned parameters to find the independent directions in which the parameters should be varied. The acceptance of the signal region selection (including the cut on the BDT_{score}) was evaluated for each re-weighting, and the largest of any of these variations was used as the uncertainty on the normalisation due to the PDF choice. For the VBF case the POWHEG+PYTHIA sample was re-weighted, whereas for the gluon-fusion case an MC@NLO sample was used. Using the same re-weighting, the largest bin-by-bin variation between the alternative PDF sets, compared to the default CT10, was assumed to be the shape uncertainty. However in both channels the effect on the shape of the discriminating variable was found to be negligible. Therefore no shape uncertainty was assigned.

Signal Re-weighting Procedure

An uncertainty was assigned on the shape of the discriminating variable to account for the re-weighting procedure described in section 4.3. From the validation of this procedure, described in section 4.3.1, a very good agreement is found at truth-level between a sample directly generated with anomalous couplings and one produced using the re-weighting procedure. The remaining small difference between generated and re-weighted distributions is assigned as a systematic uncertainty. This is done by calculating a bin-by-bin ratio (using 30 bins). Applying these ratios as weights, the reconstruction-level sample is re-weighted as a function of each event's truth level Optimal Observable/ $\Delta\phi_{jj}^{\text{sign}}$ value. The difference between this (doubly) reweighted sample and the default is then used as a shape uncertainty. The shape impact of this procedure is shown in figure 5.5.

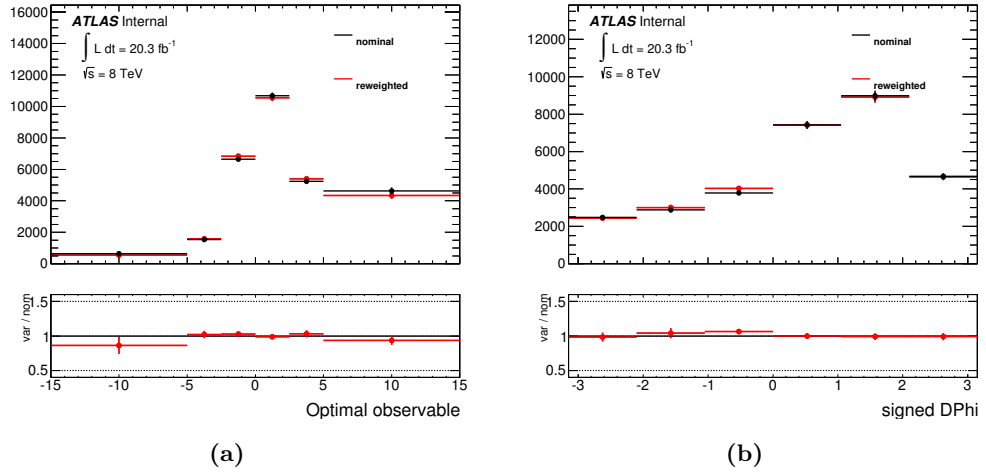


Figure 5.5: Shape uncertainty on the Optimal Observable (a) and $\Delta\phi_{jj}^{\text{sign}}$ (b) distributions of VBF signal events in the final signal region due to the signal re-weighting procedure.

Summary

The preceding sections have introduced the fit model incorporating a binned log-likelihood estimator with the VBF Higgs signal strength as its parameter of interest, used to fit data to a range of CP-odd coupling strength hypotheses in order to obtain a confidence interval on the strength parameter \tilde{d} . Also introduced were the sources of systematic uncertainties, their estimated size and how they enter the fit as nuisance parameters describing normalisation and shape variations. With this

in place it is now possible to perform the fits and extract expected and observed results.

Chapter 6

Results

This chapter presents the results of the search for CP-violating vector boson couplings in VBF Higgs boson production in the ditau decay channel. A set of preliminary studies with a simplified background model using only the mean value of the Optimal Observable has been performed to produce a rough estimate of the sensitivity and are described in section 6.1. The expected results using a log-likelihood fit to the full distribution of the Optimal Observable are presented in section 6.2, while the corresponding results using real data follow in section 6.3. The complexity of the fit model warrants extra care in its validation, and section 6.4 describes several steps taken to ensure that the fit result is sensible. Section 6.5 concludes the chapter with a discussion of the findings.

6.1 Preliminary Studies

The fit procedure utilised in the final analysis performs a binned log-likelihood fit to the distribution of the Optimal Observable O_1 as described in section 5.1. The Optimal Observable is a matrix element construction incorporating the kinematics of all final state objects from the hard interaction into a single variable. During the development of the analysis, two preliminary studies were performed on simulated events in order to substantiate the claim that the Optimal Observable is capable of providing a confidence interval on the coupling strength \tilde{d} of the BSM CP-odd admixture of Higgs boson couplings to weak bosons in the Lagrangian using the statistics available in the ATLAS Run 1 data. These studies implement a simplified fit model taking only the mean value $\langle O_1 \rangle$ into account. The first study relies on a Neyman construction [187] to calculate the mean and standard deviation of O_1 for a set of CP-odd coupling strengths \tilde{d} , which forms a ‘gauge curve’ when plotting $\langle O_1 \rangle$ and its statistical error versus \tilde{d} , from which an estimate of the \tilde{d} sensitivity can be obtained as described in section 6.1.1. The second study takes into account the mean values of both the first and second order Optimal Observables by constructing

a χ^2 measure to estimate the statistical compatibility between the purely CP-even hypothesis (SM) and each CP-odd hypothesis. Neither of these studies includes systematic uncertainties as they are both considered to be proof-of-concept studies.

6.1.1 Gauge Curve from Neyman Construction

It has been demonstrated in [67] that measuring the mean of the distribution of the first order Optimal Observable provides a sufficiently accurate measure to be sensitive to small CP-odd contributions. A first estimate of the discrimination power was therefore performed through the production of a so-called gauge curve showing the mean value of the Optimal Observable and its one sigma uncertainty band as a function of the CP-odd coupling strength parameter \tilde{d} , which can be used to ‘gauge’ the sensitivity. A total of 10000 pseudo-experiments were run for each \tilde{d} value by varying the number of events in signal and background according to a Poisson distribution with the number of events in the sample as its mean. For each pseudo-experiment the mean of the Optimal Observable distribution is found and plotted. The gauge curve is then constructed by extracting the median and 68% quantiles of the distribution of means for each value of \tilde{d} . The result is seen in figure 6.1 for pure signal and when including backgrounds. This employs a simplified background model taking into account only the gluon fusion $H \rightarrow \tau\tau$ as well as the $Z \rightarrow \tau\tau$ samples as it was performed before the full background treatment was implemented. The $Z \rightarrow \tau\tau$ process is a dominant background process constituting approximately 35% of the total background in the VBF event category. The 68% confidence interval for \tilde{d} is identified by reading off the values where $\langle\mathcal{O}\rangle$ is no longer compatible with zero within the uncertainty band. The gauge curve that does not include backgrounds is able to produce a confidence interval of roughly $-0.05 < \tilde{d} < 0.05$, while inclusion of the backgrounds produces a much flatter curve that is shifted slightly towards negative $\langle\mathcal{O}\rangle$, which is not an issue since the value of $\langle\mathcal{O}\rangle$ is still compatible with zero within the statistical uncertainty in the SM case of $\tilde{d} = 0$ as required. The uncertainty bands are too wide to extract a double-sided limit when backgrounds are added. Using the mean of the first order Optimal Observable will therefore need higher statistics in order to be sensitive to CP-odd couplings.

6.1.2 Construction of χ^2 measure

As demonstrated above, the signal yield relative to background is too low to extract a useful limit on the value of \tilde{d} using the mean of the first order Optimal Observable.

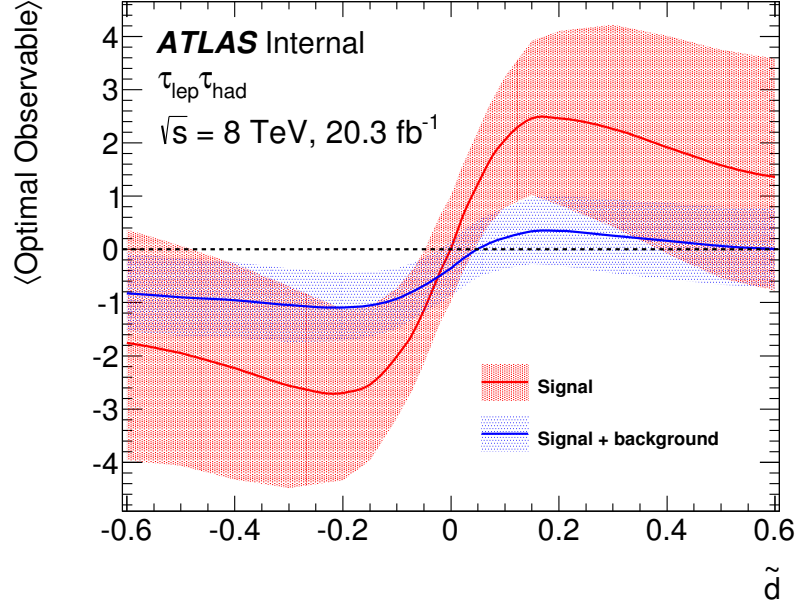


Figure 6.1: Gauge curves calculated using the first order Optimal Observable for signal only (red) and backgrounds included (blue). Only statistical errors are taken into account. There are insufficient statistics to extract a limit when including backgrounds.

In order to utilise both first and second order Optimal Observables it is necessary to construct a measure of significance that takes into account both variables and their covariance. This is easily obtained using a standard χ^2 measure:

$$\chi^2(\tilde{d}) = \sum_{i,j=1,2} \left(\langle O_i \rangle(\tilde{d}) - \langle O_i \rangle_{\text{SM}} \right) \text{Cov}[O_i, O_j](\tilde{d}) \left(\langle O_j \rangle(\tilde{d}) - \langle O_j \rangle_{\text{SM}} \right), \quad (6.1)$$

where $\langle O_i \rangle$ is the mean of the i th order Optimal Observable, and $\text{Cov}[a, b]$ is the covariance of a and b . Plots of the combined χ^2 as well as individual χ^2 measures for the first and second order Optimal Observables for signal and signal+backgrounds are shown in figure 6.2. $\langle O_i \rangle$ are truncated means due to the observation of long tails in the distribution of O_2 . The first order Optimal Observable can be seen to provide an estimate very close to the sum of O_1 and O_2 at low CP-odd coupling strengths, in this case $|\tilde{d}| \lesssim 0.15$, as expected according to the considerations outlined in [67] showing that the first order Optimal Observable is sufficiently accurate when the admixture of CP-odd couplings is significantly lower than the SM CP-even couplings. At higher \tilde{d} values the contribution from O_2 grows quadratically as it is sensitive

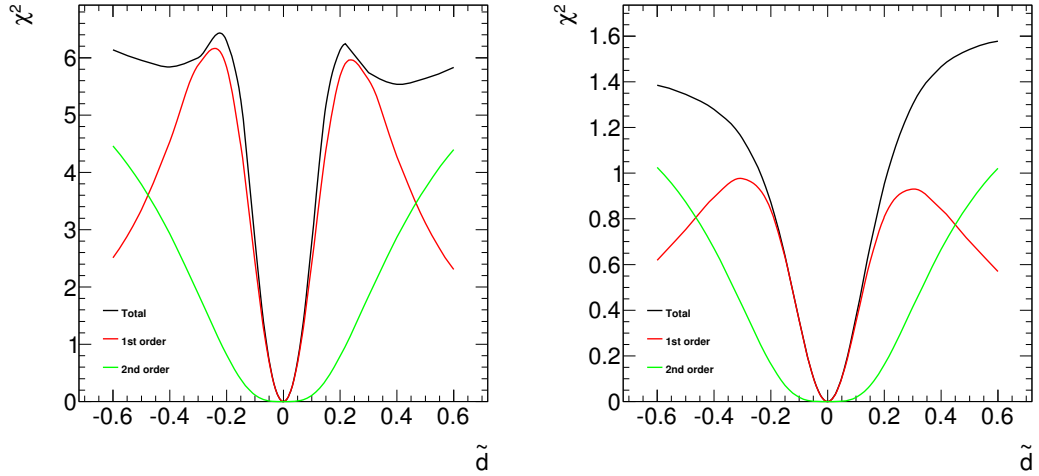


Figure 6.2: χ^2 curves for first and second order Optimal Observable and both orders combined for signal only (left) and with backgrounds included (right). The background model includes the full set of backgrounds. Only statistical errors are taken into account.

to the total CP-odd cross section of the VBF process. The total χ^2 including both orders is noticeably different from a simple sum of the two individual curves due to correlations between O_1 and O_2 . At $|\tilde{d}| \gtrsim 0.2$, a turnover is observed in the curves for O_1 . Since for a purely CP-odd VBF coupling the O_1 distribution will return to a symmetric but shape-wise different distribution it is expected that the χ^2 value will drop at higher CP-odd coupling strength \tilde{d} and approach a constant nonzero value when \tilde{d} becomes the dominant term.

The accuracy of the χ^2 measure was estimated by performing pseudo-experiments where the event yield for each SM sample was again varied according to a Poisson distribution. If all errors (in this case the entries in the covariance) are estimated correctly, the statistical probability of two random samples of SM events being compatible ($\chi^2 \leq 1$) is 68%. Performing 10000 pseudo-experiments showed that the 68% quantile of the resulting χ^2 distribution was located at $\chi^2 = 1.19$. Limits on \tilde{d} therefore correspond to the values of \tilde{d} where $\chi^2 = 1.19$. The resulting stats-only limits are $-0.32 < \tilde{d} < 0.25$.

6.2 Expected sensitivity

A preliminary sensitivity estimate was obtained by keeping the data in the signal region blinded. Data in all control regions were studied in detail during the develop-

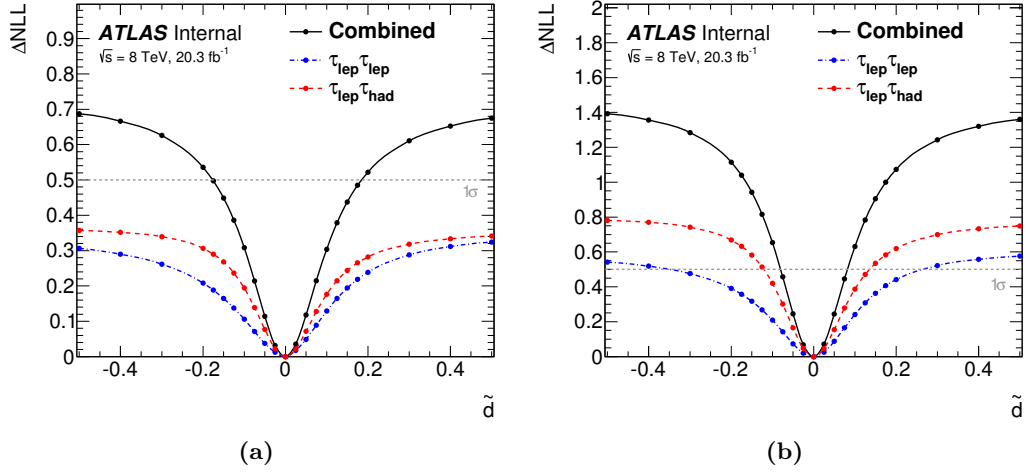


Figure 6.3: Expected ΔNLL at each \tilde{d} hypothesis using the Optimal Observable for $\tau_\ell\tau_\ell$ (blue), $\tau_\ell\tau_{\text{had}}$ (red) and their combination (black). An Asimov dataset with SM backgrounds plus pure CP-even SM signal ($\tilde{d}=0$), scaled with a signal strength of 1.0 (a) and 1.55 (b), has been used.

ment of the couplings analysis [2] where it was shown to be in good agreement with the background estimate. Inclusion of actual data in the control regions of this CP analysis was therefore assessed to pose no serious risk of analyst bias. The expected central confidence interval is thus the result of a ‘hybrid’ fit using actual data in control regions while constructing an Asimov dataset in the signal region consisting of the expected background and the pure CP-even SM VBF $H \rightarrow \tau\tau$ signal. In order to gauge the influence of the fitted signal strength on the resulting confidence interval the fit was performed twice with Asimov datasets having injected VBF $H \rightarrow \tau\tau$ signal strengths of $\mu = 1.0$ and $\mu = 1.55$ times the SM prediction respectively. The signal strength of $\mu = 1.55$ is the result of the final unblinded fit (see section 6.3) and was chosen to directly compare expected and observed results. Figure 6.3 shows the ΔNLL curves for the combination of the $\tau_\ell\tau_{\text{had}}$ and $\tau_\ell\tau_\ell$ channels in addition to each individual channel for the two hybrid-Asimov datasets. The fit includes all the systematics described in chapter 5. As ΔNLL refers to the deviation from the point where the log-likelihood is minimal, the curves will always have a minimum of $\Delta\text{NLL} = 0$ at the value of \tilde{d} where the data and signal plus background model are most compatible, surrounded by higher values rising at a rate determined by how quickly the signal shape changes with the value of \tilde{d} . The rise will level off at larger values both due to the choice of coarse bins for high $|O_1|$ values and the fact that the signal distribution will return to a symmetric configuration in the limit of large CP-odd

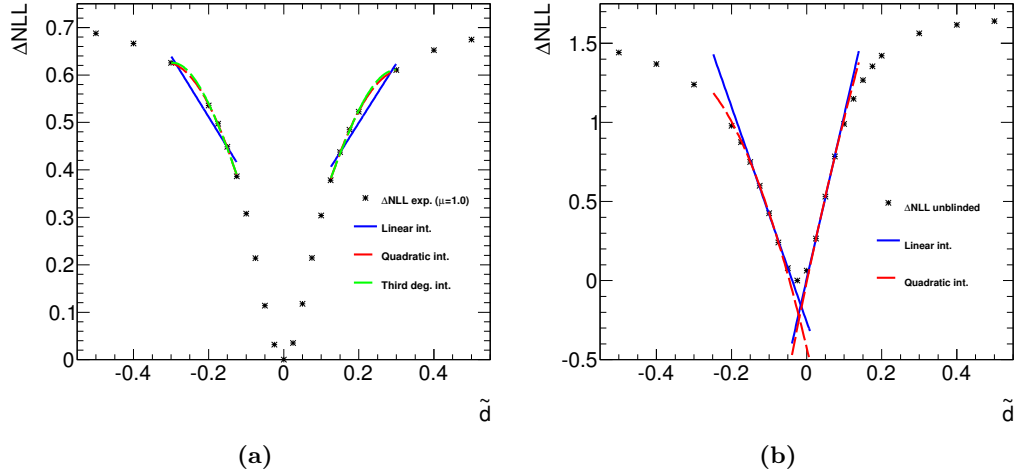


Figure 6.4: Polynomials of various orders fitted to the likelihood points surrounding the one sigma ($\Delta\text{NLL} = 0.5$) limit in the case of a hybrid fit with Asimov signal strength $\mu = 1.0$ (a) and the final unblinded fit (b).

contributions. Each curve is produced from discrete values indicated by markers where a fit has been performed, while the lines connecting the markers are graphical interpolations between the fit results. The intersection with $\Delta\text{NLL} = 0.5$ therefore has to be determined by numerical interpolation. The interpolation is performed by fitting a polynomial to the ΔNLL points immediately surrounding this value and inverting the fitted function to find the \tilde{d} value at $\Delta\text{NLL} = 0.5$. The curvature of the ΔNLL curve close to the one sigma level potentially requires a higher order polynomial to correctly reproduce the functional dependence. Figure 6.4 illustrates the interpolation method used to estimate the confidence interval for both hybrid (Asimov) and unblinded fits. Polynomials of increasing degree were added until the two highest orders were in agreement within the quoted precision. All confidence intervals are estimated using this interpolation procedure.

Figure 6.5 shows the best-fit signal strength of each signal hypothesis using the hybrid-Asimov datasets, i.e. datasets with real data in the CRs. The fit constrains the signal strength close to the injected value in the corresponding Asimov dataset at $\tilde{d} = 0$ but decreases it for less compatible CP-mixing models, which are normalised to the SM cross section at pre-fit level. This is expected since a variation of \tilde{d} will shift the modelled signal distribution compared to the injected SM signal in the Asimov data and force the fitted signal strength to decrease in order to maximise compatibility with the Asimov distribution. The remaining small difference between the injected signal strength in the Asimov data and the best-fit value is

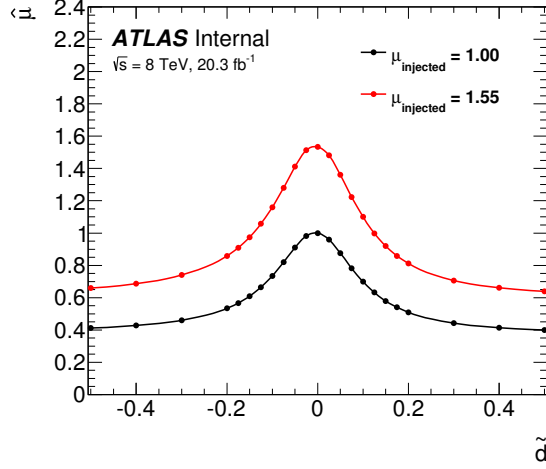


Figure 6.5: Expected best-fit signal strength $\hat{\mu}$ at each \tilde{d} hypothesis for the combined fit of both channels. An Asimov dataset with SM backgrounds plus pure CP-even SM signal ($\tilde{d}=0$), scaled with a signal strength of 1.0 (black) and 1.55 (red), has been used.

due to constraints introduced by nuisance parameters when including real data in control regions, which can have slightly different shapes compared to the estimated background due to e.g. statistical variations and therefore pull some nuisance parameters away from their nominal values and affect the best-fit signal strength value in the signal region.

The expected confidence interval on \tilde{d} at the 68% (one standard deviation) level, with all systematic uncertainties included, is determined to be $[-0.18, 0.18]$ ($[-0.08, 0.08]$) assuming $\mu = 1.0$ ($\mu = 1.55$).

6.3 Observed results

The post-fit $\text{BDT}_{\text{score}}$ distributions in $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ are shown in figure 6.6. The non-equidistant bins used in the fit make them difficult to inspect if plotted directly. Therefore the bins have been transformed to appear equidistant while keeping the contents the same. This changes the BDT values on the axis, and the $\text{BDT}_{\text{score}}$ selection criteria used to define the final signal region correspond to $\text{BDT}_{\text{score}} > 0.5$ for $\tau_\ell\tau_\ell$ and $\text{BDT}_{\text{score}} > 0.3$ for $\tau_\ell\tau_{\text{had}}$ in these plots.

The observed ΔNLL curves as a function of \tilde{d} for the combination of both channels and each channel individually are shown in figure 6.7. ΔNLL curves resulting from independent fits to the individual channels cannot be used to estimate their individual contribution to the combined result since the best-fit values of the

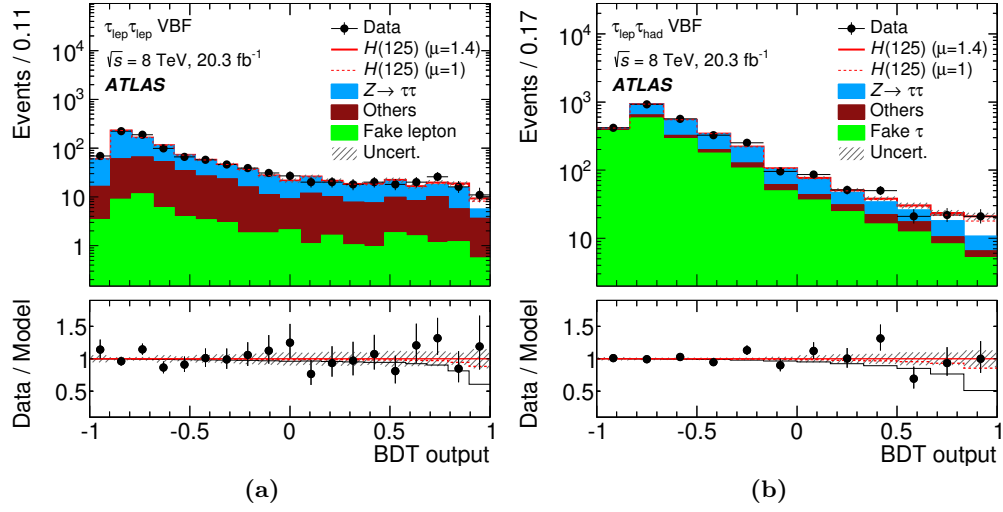


Figure 6.6: Post-fit BDT_{score} distributions in $\tau_\ell\tau_\ell$ (a) and $\tau_\ell\tau_{had}$ (b) in the VBF region. The statistical and systematic uncertainties are given by the hashed band. The BDT_{score} bins have been transformed to appear equidistant for easier inspection. The signal region selection criteria of $BDT_{score} > 0.68$ in $\tau_\ell\tau_\ell$ and $BDT_{score} > 0.95$ in $\tau_\ell\tau_{had}$ correspond to values of 0.5 and 0.3 in these plots respectively [2].

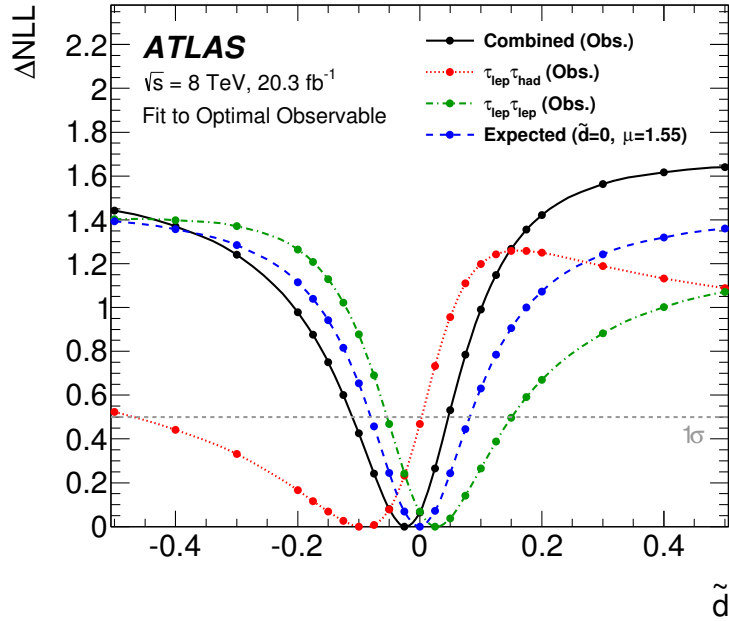


Figure 6.7: Observed ΔNLL as a function of the \tilde{d} values defining the underlying signal hypothesis for $\tau_\ell\tau_\ell$ (green), $\tau_\ell\tau_{had}$ (red) and their combination (black) as well as the expected sensitivity using an Asimov dataset injected with the observed best-fit signal strength of $\mu = 1.55$. The curves for the individual channels are calculated with the best-fit values of all nuisance parameters from the combined fit such that their contribution to the combined fit is more apparent.

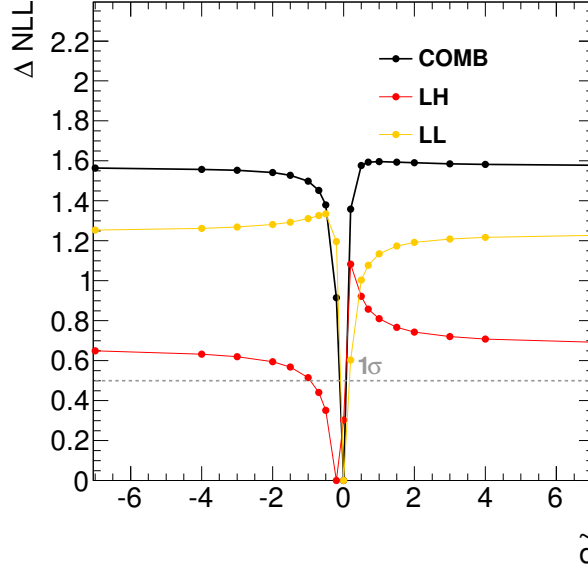


Figure 6.8: Observed ΔNLL as a function of the \tilde{d} values defining the underlying signal hypothesis for $\tau_\ell\tau_\ell$ (yellow, denoted ‘LL’), $\tau_\ell\tau_{had}$ (red, denoted ‘LH’) and their combination (black). An extended \tilde{d} range shows the curves approaching a constant non-zero value for large $|\tilde{d}|$.

nuisance parameters differ between the individual channel fits and the combination. This directly affects the sensitivity as the signal strength parameter μ is included in this list of parameters and varies by a non-negligible amount. A more accurate estimate of the individual channel contributions can be produced by utilising the best-fit result for each nuisance parameter from the combined fit at each \tilde{d} point, and then calculating the ΔNLL for the individual channels using these values. The individual channel curves in figure 6.7 have been produced in this way.

The Optimal Observable of first order is CP-odd, and thus for a purely CP-odd Higgs boson it will be symmetric again, as in the pure CP-even SM case. However, the shape of the expected distribution will deviate from the CP-even case, and hence at very high \tilde{d} values the ΔNLL curve is expected to approach a constant nonzero value. The actual value is determined by the observed shape difference between the pure CP-odd and CP-even cases. A fit extending to large positive and negative values of \tilde{d} is shown in figure 6.8 and confirms that the ΔNLL curve does not drop below the one sigma confidence level once it has been exceeded.

High levels of mixing (large values of $|\tilde{d}|$), have already been excluded by previous analyses [188]. The strength of the present analysis therefore lies with small absolute values of \tilde{d} . In the range considered, \tilde{d} values of $[-0.11, 0.05]$ are

Process	High BDT signal region	
	$\tau_\ell\tau_\ell$	$\tau_\ell\tau_{\text{had}}$
Data	54	68
$Z \rightarrow \tau\tau$	19.6 ± 1.0	19.1 ± 2.2
Fake bkg	2.3 ± 0.3	24.1 ± 1.5
Top	3.8 ± 1.0	4.8 ± 0.7
$Z \rightarrow \ell\ell$ +Diboson	11.5 ± 1.7	5.3 ± 1.6
(non-VBF) $H \rightarrow \tau\tau/WW$	1.6 ± 0.2	2.5 ± 0.7
(VBF) $H \rightarrow \tau\tau/WW$	9.8 ± 2.1	16.7 ± 4.1

Table 6.1: Observed post-fit event yields in the signal region. The errors include systematic uncertainties.

found to be consistent with the data at the one sigma level. This interval has been extracted from the negative log-likelihood curve where $\Delta\text{NLL} = 0.5$ as explained in section 5.1.

Post-fit yields for signal and all backgrounds are listed for $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ individually in table 6.1. Post-fit distributions of the first order Optimal Observable for SM signal are shown in figure 6.9. The best-fit signal strengths obtained in the individual channel fits and the combined fit to the Optimal Observable are

$$\mu_{\tau_\ell\tau_\ell} = 2.37^{+1.48}_{-1.34} \quad (6.2)$$

$$\mu_{\tau_\ell\tau_{\text{had}}} = 1.09^{+1.02}_{-0.91} \quad (6.3)$$

$$\mu_{\text{comb}} = 1.55^{+0.88}_{-0.76} \quad (6.4)$$

The numbers are all in good agreement with the values obtained in the couplings analysis considering both the relatively large uncertainties and the fact that the boosted category is not included in the present analysis. The post-fit distributions of $\Delta\phi_{jj}^{\text{sign}}$ are shown in figure 6.10 and the corresponding numbers for the signal strength are

$$\mu_{\tau_\ell\tau_\ell} = 2.89^{+1.47}_{-1.29} \quad (6.5)$$

$$\mu_{\tau_\ell\tau_{\text{had}}} = 1.54^{+1.03}_{-0.94} \quad (6.6)$$

$$\mu_{\text{comb}} = 2.02^{+0.87}_{-0.77} \quad (6.7)$$

The mean value of the Optimal Observable is expected to be zero in the case of CP-even signal, while non-zero values are expected when including CP-violating couplings. It was already shown that the total background is expected to have a

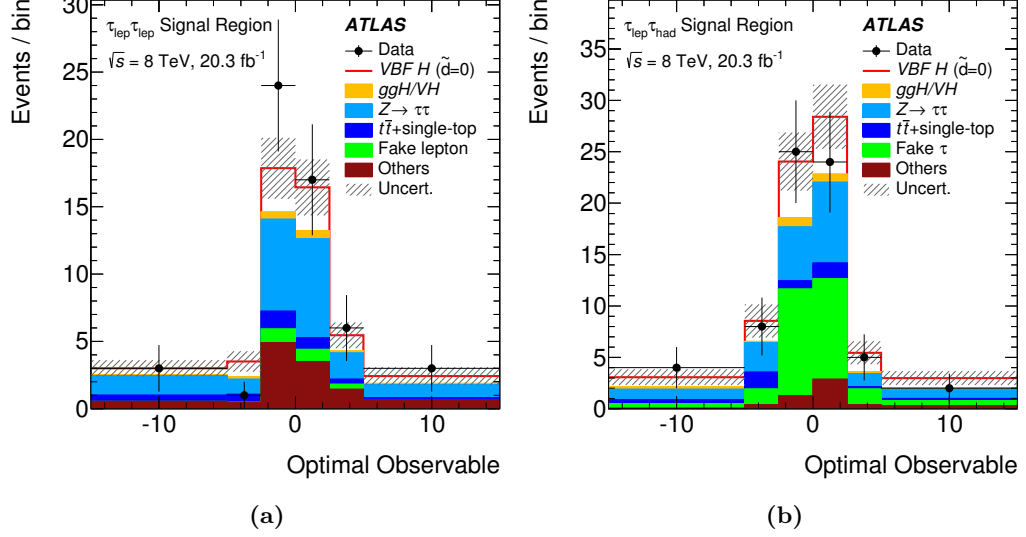


Figure 6.9: Distributions of O_1 in the signal region of the $\tau_\ell\tau_\ell$ (a) and $\tau_\ell\tau_{had}$ (b) channel after performing the combined fit for the $\tilde{d} = 0$ hypothesis. The best-fit signal strength is found to be $\mu = 1.55$. The error bands include both statistical and systematic uncertainties.

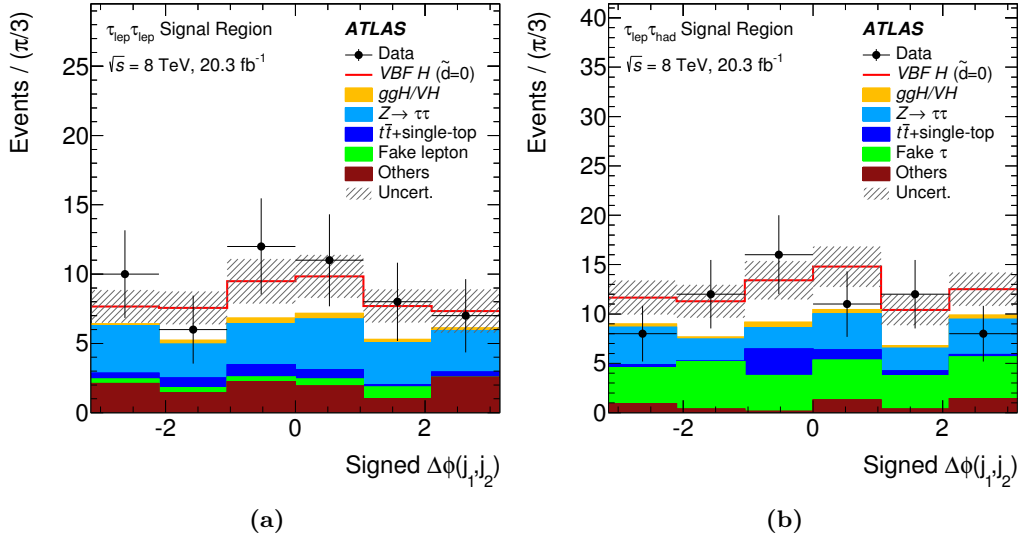


Figure 6.10: Distributions of $\Delta\phi_{jj}^{sign}$ in the signal region of the $\tau_\ell\tau_\ell$ (a) and $\tau_\ell\tau_{had}$ (b) channel after performing the combined fit for the $\tilde{d} = 0$ hypothesis. The best-fit signal strength is found to be $\mu = 2.02$. The error bands include both statistical and systematic uncertainties.

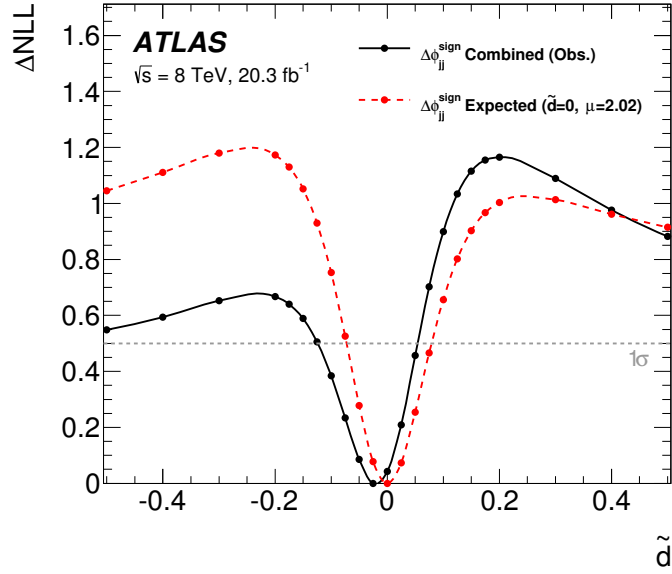


Figure 6.11: Observed (black) and expected (red) ΔNLL as a function of the \tilde{d} values defining the signal hypothesis using $\Delta\phi_{jj}^{\text{sign}}$ as the discriminating variable in the combined fit of both channels. The expected result uses an Asimov dataset injected with the observed best-fit signal strength of $\mu = 2.02$.

mean value of zero since no CP-odd processes are involved, which means that the mean value in the observed data should also be consistent with zero in the case of no CP-violating effects within the precision of the measurement. In the signal regions of $\tau_\ell\tau_\ell$ ($\tau_\ell\tau_{\text{had}}$) the observed mean value of the Optimal Observable in data is 0.3 ± 0.5 (-0.3 ± 0.4). This is fully consistent with zero and hence with the full statistical analysis, showing no sign of CP-violation.

Another goal of the analysis was to investigate whether the matrix element approach used in the Optimal Observable can lead to superior exclusion limits compared to the classical variable $\Delta\phi_{jj}^{\text{sign}}$. The ΔNLL curves produced by the combined fit of $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ when utilising the $\Delta\phi_{jj}^{\text{sign}}$ variable with a hybrid dataset as well as unblinded data are included in figure 6.11. The hybrid dataset has the best-fit signal strength from the unblinded fit ($\mu = 2.02$) injected into the Asimov data in the signal region in order to compare the expected and observed curves at the same signal strength. A small shift towards negative \tilde{d} is observed which is comparable to the shift seen in the Optimal Observable fit. The sensitivities of the Optimal Observable and $\Delta\phi_{jj}^{\text{sign}}$ are compared in figure 6.12 showing a clear advantage in terms of the expected confidence interval when the fit includes the Optimal Observable.

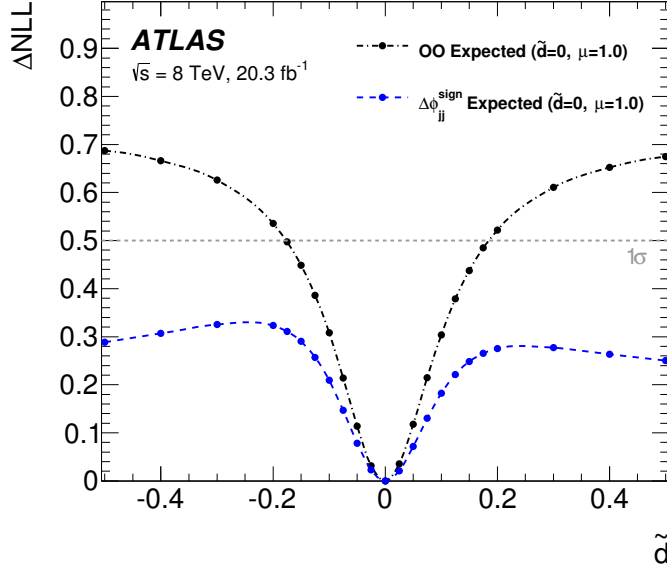


Figure 6.12: Expected ΔNLL curves when using the Optimal Observable (black) and $\Delta\phi_{jj}^{\text{sign}}$ (blue) as the final discriminating variable for the combination of both channels as a function of the \tilde{d} values defining the signal hypothesis. The curves include an Asimov dataset with SM backgrounds plus pure CP-even VBF signal ($\tilde{d} = 0$) scaled to the predicted SM signal strength $\mu = 1.0$.

6.4 Validation of the Maximum-Likelihood Fit

The full likelihood function used in the final fit includes hundreds of parameters such as histogram bin contents and their statistical uncertainties, free-floating background normalisation factors and nuisance parameters affecting both the normalisations and shapes of the input samples. This section will discuss several steps that were taken in order to validate the fit model.

The impact of individual NPs on the uncertainty of the estimated signal strength $\hat{\mu}$ was investigated by separately fixing a single NP to its $\pm 1\sigma$ uncertainties and performing a new likelihood minimisation over all remaining model parameters. The change in the fitted signal strength $\Delta\hat{\mu}$ with respect to the nominal estimate is the criterion used to rank the NPs in terms of their relative impact on the fit. This study was performed using two different signal hypotheses, namely pure SM signal (figure 6.14) and $\tilde{d} = 0.1$ (figure 6.15) in order to investigate whether the individual NPs impact the analysis differently between signal hypotheses. The yellow bands together with the upper x-axis indicate the post-fit impact on $\hat{\mu}$ compared to its nominal pre-fit value, normalised to its total post-fit uncertainty $\Delta\hat{\mu}_{\text{tot}}$ when all NPs are varied. The black markers and the lower x-axis indicate the post-fit central

value and the black error bars are the post-fit uncertainties, both normalised to one pre-fit standard deviation. The estimated pre-fit uncertainties are included as grey error bars making it easier to notice asymmetries in the post-fit errors. An exception is made in the case of normalisation parameters whose uncertainties are given as absolute values since they have no pre-fit estimates and therefore have no meaningful pre-fit uncertainty. The highest ranking parameters include components of the energy scale uncertainties of jets, electrons and τ leptons in addition to theoretical uncertainties on the Higgs boson branching ratio to two τ -leptons and the renormalisation and factorisation scales included in the cross-section calculated for the gluon-fusion production channel. No significant deviations from the pre-fit values are observed in the NP pulls as the pre- and post-fit values are all compatible within their one sigma uncertainties. Changes in parameter $\Delta\hat{\mu}$ ranking between the $\tilde{d} = 0.0$ and $\tilde{d} = 0.1$ cases are minor, both having the same top 4 parameters.

The electron energy scale NP warrants extra scrutiny. Even with the implementation of the smoothing and pruning mechanisms described in section 5.1.1 it is still possible, in the limit of low statistics, that fluctuations might overestimate an uncertainty. In the analysis the four components of the electron energy scale are combined in quadrature and enter the fit through the parameter EL_SCALE. Appendix A shows the variations related to each of the four uncertainties in the VBF region and in the final signal region including the $\text{BDT}_{\text{score}}$ selection criterion in the $\tau_\ell\tau_{\text{had}}$ channel (see table 4.12 for region definitions). In the VBF region all four variations are sufficiently small to be pruned away, both in terms of normalisation and shape. However, the final signal region has normalisation variations of several percent in three of the components. A possible solution is to derive the normalisation variations in a looser event category with a higher number of events and apply this in the signal region to mitigate any effects solely stemming from low statistics. It has not been feasible to demonstrate whether the increase in the size of the variations is due to an actual effect of applying a selection criterion on the $\text{BDT}_{\text{score}}$ or if the effect is purely of a statistical nature. Deriving the variations in a looser region could therefore potentially provide an inaccurate description. Another proposed solution is to compensate by subtracting the statistical error on the variation in each bin [189]. This will indeed reduce the size of the variation in cases with a low number of events but will introduce the possibility of negative variations in the case where the statistical error is larger than the variation itself. The choice was therefore made to proceed with the standard uncertainty estimation also in the case of the electron energy scale uncertainty.

The covariance matrix of the fit model is estimated during the minimisation

and can be studied to identify any highly correlated parameters. High correlation coefficients between NPs would reveal possible degeneracies in the fit model, hinting at a suboptimal choice in the parametrisation of the parameters. Figure 6.13 plots the correlation coefficients of the top 25 NPs from figure 6.14, i.e. those parameters having the greatest influence on the fitted signal strength $\hat{\mu}$. The highest correlation coefficients are approximately -0.3 and appear between free-floating background normalisation factors and dominant systematic uncertainties, e.g. between the normalisation of the top background and one of the TES parameters. The NPs can indeed have a non-negligible effect on the acceptance of various backgrounds, so this behaviour is expected.

Multiple minima in the likelihood function must be avoided to ensure fit stability and remove the possibility of the fit converging on a false local minimum. The likelihood function was therefore profiled along all parameters and the resulting curves were studied. A sufficiently parabolic shape close to the global minimum was observed for all parameters. Figure 6.16 contains examples of one-dimensional likelihood profile curves for several NPs as well as the signal strength parameter μ .

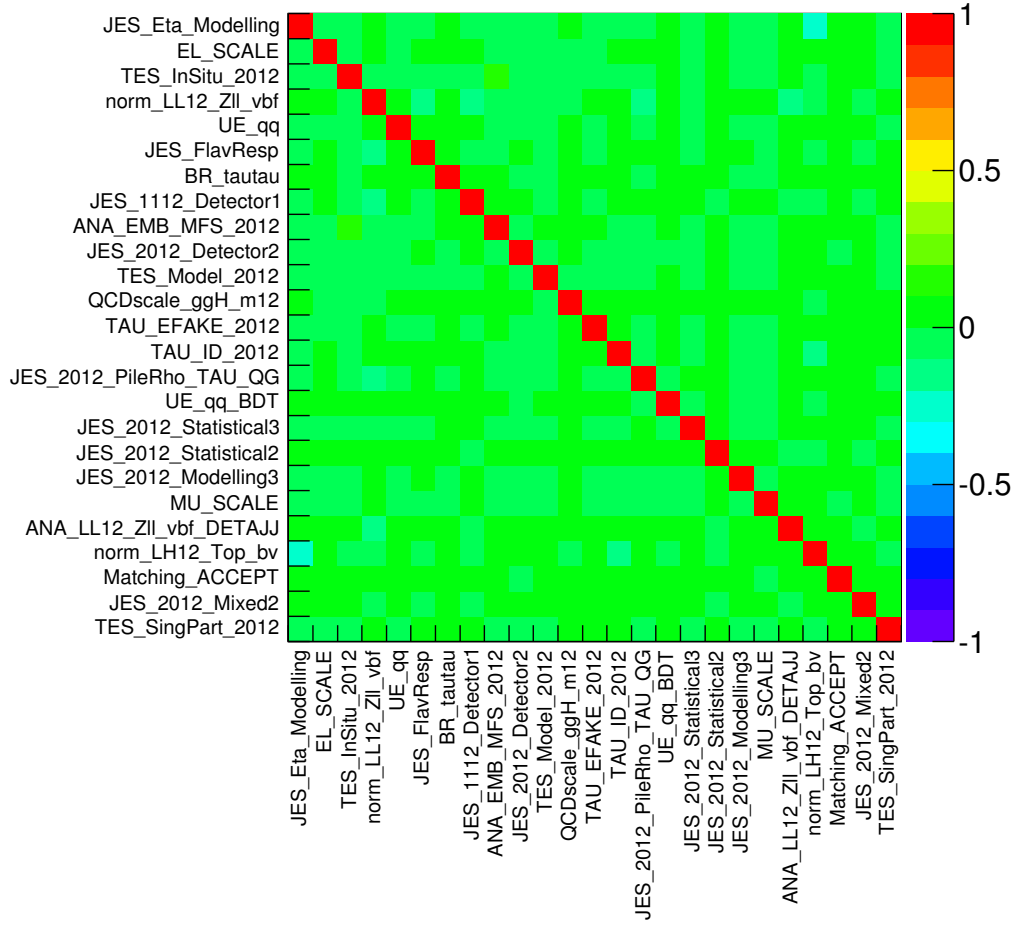


Figure 6.13: Correlation coefficients of the 25 NPs having the greatest influence on the fitted signal strength $\hat{\mu}$. A short description of the dominant parameters is given in appendix B.

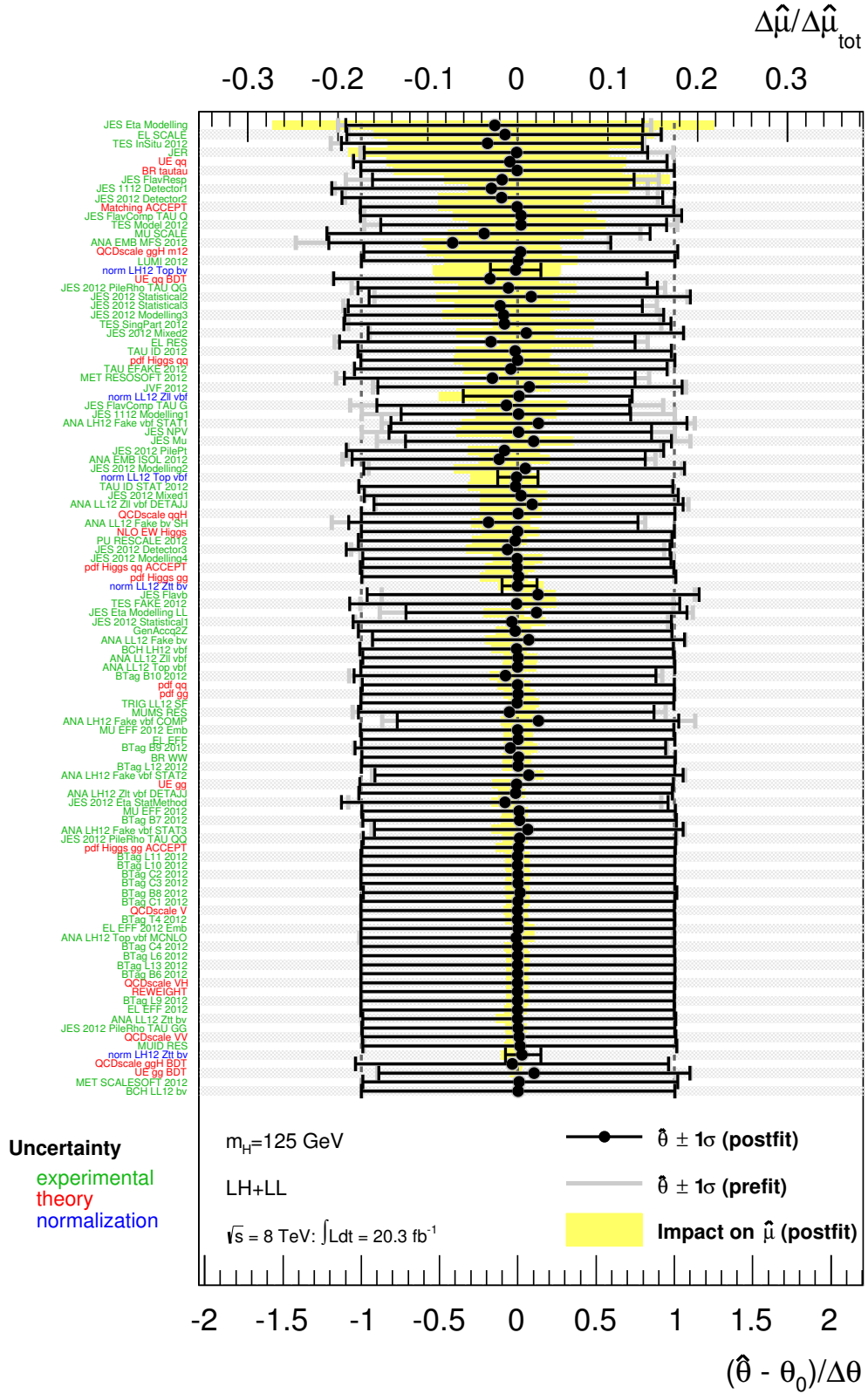


Figure 6.14: Observed NP ranking (y-axis), pulls (lower x-axis) and the relative impact on $\hat{\mu}$ with respect to its total uncertainty (upper x-axis) for the combined $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{had}$ fit using the pure SM signal hypothesis ($\tilde{d} = 0.0$). Post-fit normalisation uncertainties are given as absolute values.

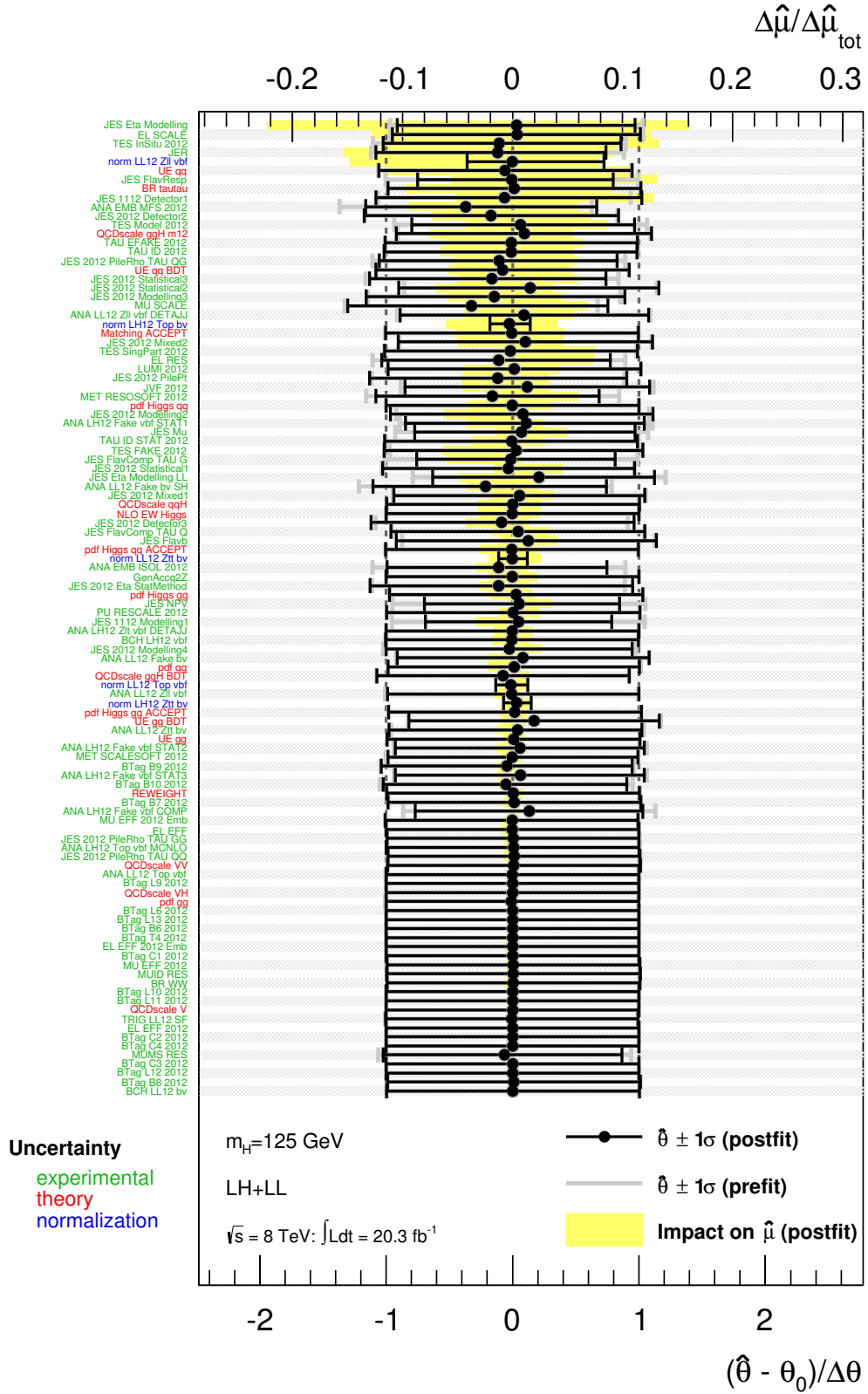


Figure 6.15: Observed NP ranking (y-axis), pulls (lower x-axis) and the relative impact on $\hat{\mu}$ with respect to its total uncertainty (upper x-axis) for the combined $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ fit using the BSM signal hypothesis $\hat{d} = 0.1$. Post-fit normalisation uncertainties are given as absolute values.

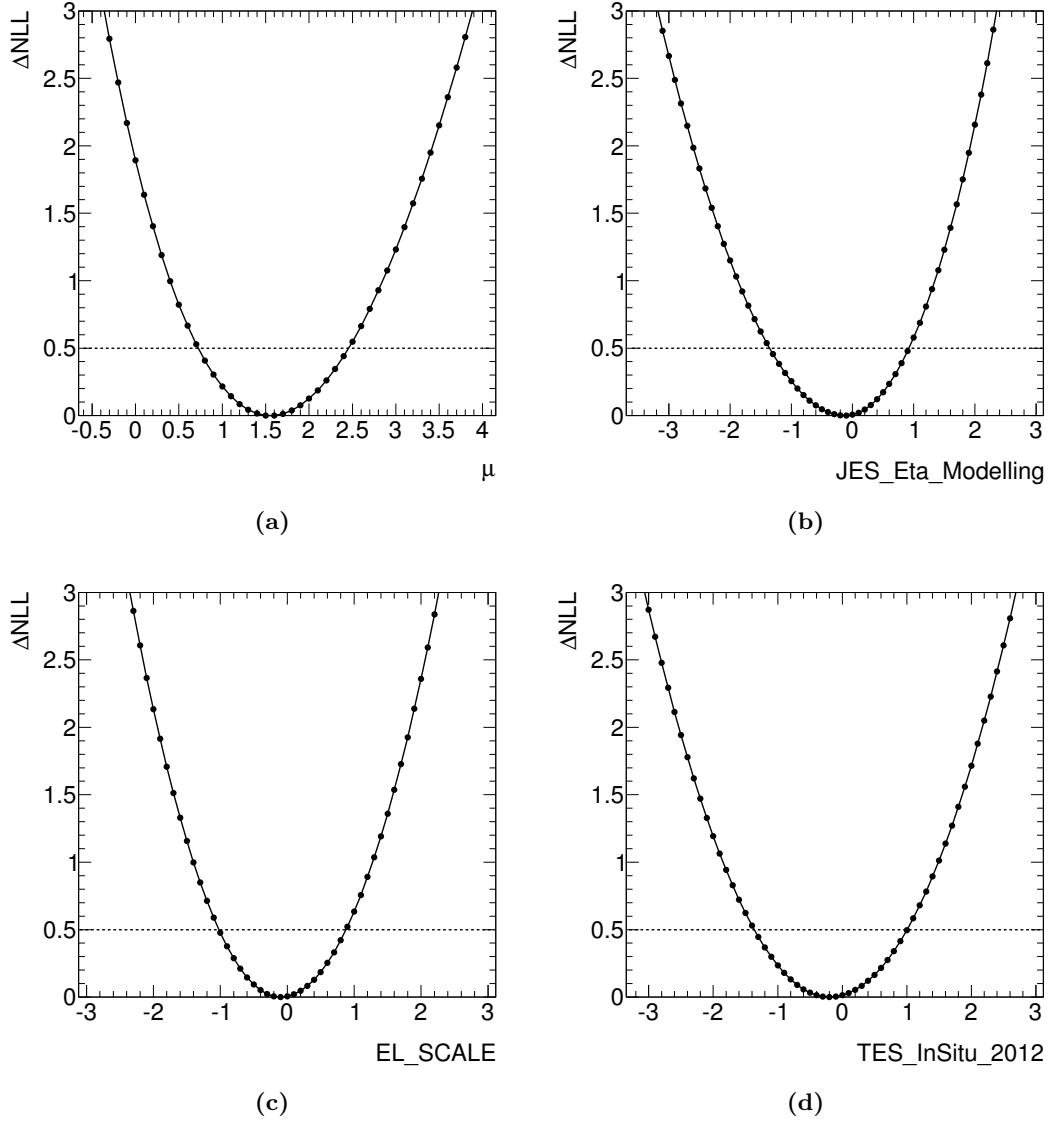


Figure 6.16: Likelihood profiles given in terms of ΔNLL as a function of four individual fit parameters. (a) shows the profile when varying the signal strength μ , while (b) to (d) give the profiles for the three NPs having the greatest influence on the best-fit value $\hat{\mu}$ as seen in figure 6.14. The one sigma deviation is indicated by the dashed line at $\Delta\text{NLL} = 0.5$. All profiles exhibit parabolic behaviour close to the global minimum.

6.5 Conclusion

Using the full 20.3 fb^{-1} of $\sqrt{s} = 8 \text{ TeV}$ proton-proton collisions recorded by the ATLAS experiment at the LHC, a test of CP invariance in the Higgs boson coupling to vector bosons has been performed using the Vector Boson Fusion production mode with decays to $\tau^+\tau^-$. A detailed account of the analysis in the $\tau_\ell\tau_{\text{had}}$ decay channel has been presented alongside a combined result with the $\tau_\ell\tau_\ell$ decay channel. The analysis makes use of the event selection, background estimation and evaluation of systematic uncertainties developed in the ATLAS analysis that provided evidence for the $H \rightarrow \tau\tau$ decay. An optimised selection cut has been added on the $\text{BDT}_{\text{score}}$ in each decay channel individually in order to achieve a higher signal purity, and additional systematic uncertainties have been estimated to account for the effects of the additional selection and the use of a new discriminating variable, an Optimal Observable, constructed directly from the matrix element of the process. This variable is shown to achieve sensitivity superior to that provided by the variable traditionally proposed for CP-violation studies, $\Delta\phi_{jj}^{\text{sign}}$. No evidence for CP violation is observed. Using the $\tau_\ell\tau_\ell$ and $\tau_\ell\tau_{\text{had}}$ decay channels, values of \tilde{d} outside the interval $[-0.11, 0.05]$ are excluded at the 68% confidence level (CL).

Confidence intervals are normally quoted at the 95% level. The present analysis was unable to produce a 95% CL, which is a consequence of the adopted approach that ignores the increase in the total signal cross section when including CP-odd coupling terms in the chosen effective field theory. However, in terms of the width of the remaining allowed range of \tilde{d} the obtained 68% CL interval is approximately an order of magnitude more powerful in constraining the strength of CP-odd weak boson couplings to the Higgs than that found in a similar recent analysis of $H \rightarrow WW$ and $H \rightarrow ZZ$ [65], which excludes values of \tilde{d} outside the interval $[-1.3, 0.0]$ at the 68% CL. With more data already underway from ATLAS in Run 2 of the LHC, the methodology of a statistical fit to an Optimal Observable for the process is expected to be highly competitive in the exploration of the remaining allowed low-coupling-strength phase space for CP violation. More data will also enable the inclusion of the second order Optimal Observable that can be added by performing a two-dimensional fit to both orders.

Appendix A

Electron Energy Scale

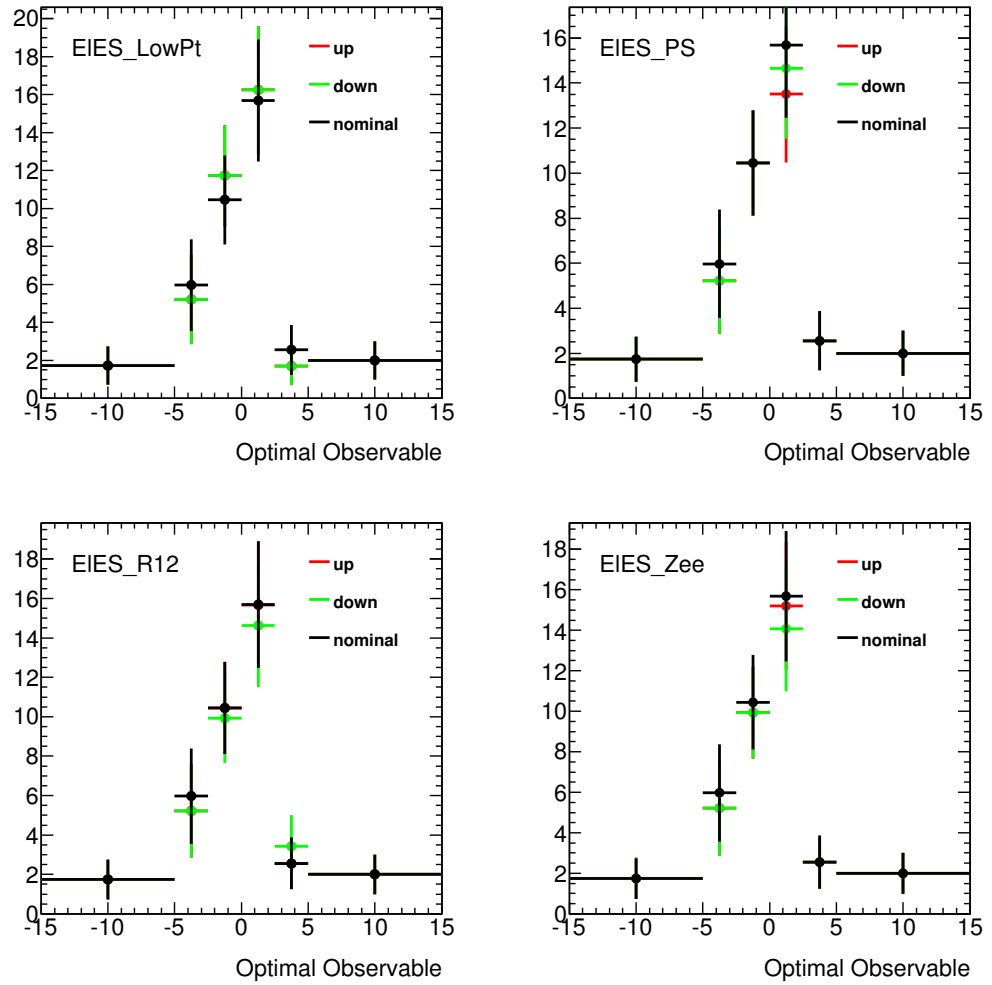


Figure A.1: Variations of the electron energy scale components on the optimal observable distribution of the $Z \rightarrow \tau\tau$ embedding sample plotted in the signal region of the $\tau_\ell\tau_{had}$ decay channel.

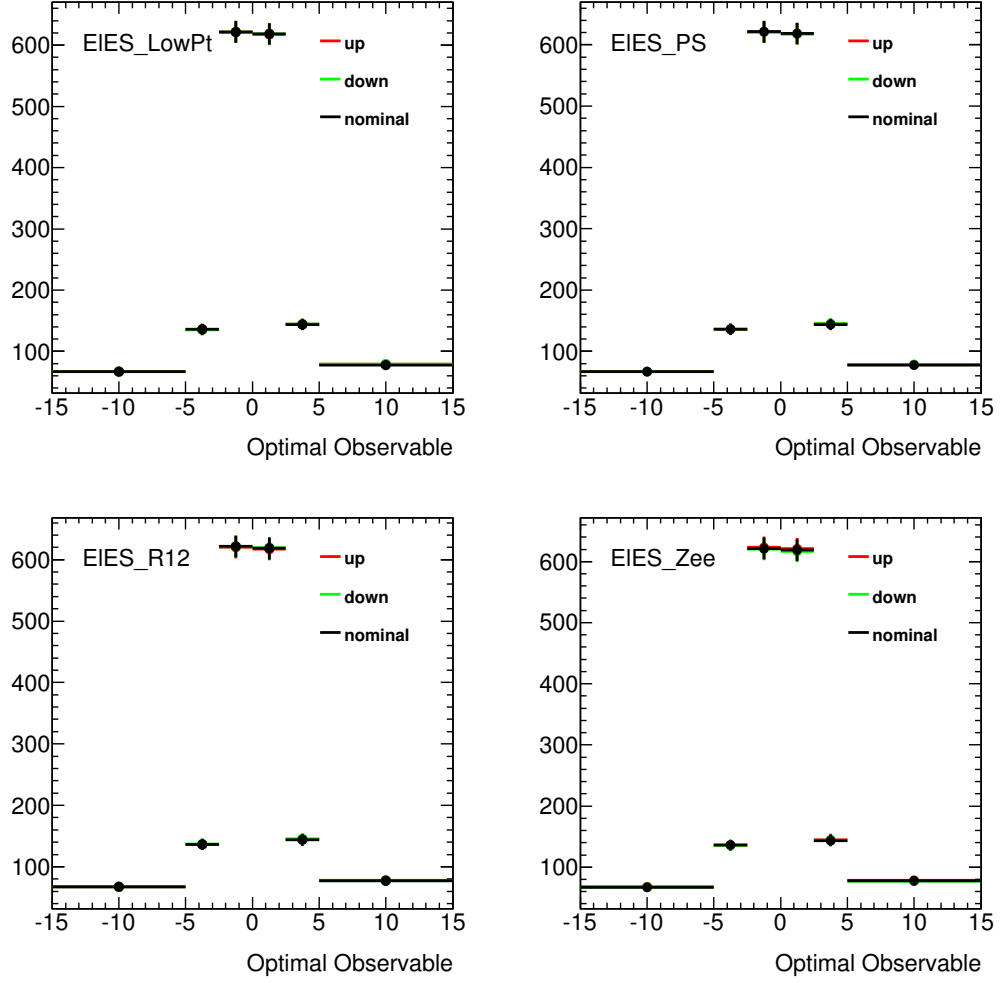


Figure A.2: Variations of the electron energy scale components on the optimal observable distribution of the $Z \rightarrow \tau\tau$ embedding sample plotted in the VBF region of the $\tau_\ell\tau_{\text{had}}$ decay channel.

Appendix B

List of Nuisance Parameters

This is a list of the dominant nuisance parameters that enter the final log-likelihood fit described in section 5.1.

- **JES_Eta_Modelling** – Uncertainty on the jet energy scale due to the η intercalibration.
- **EL_SCALE** – Electron energy scale systematic.
- **TES_InSitu_2012** – Component of the tau energy scale uncertainty from the in-situ tau energy correction.
- **norm_LL12_Zll_vbf** – Freely floating normalisation factor on the $Z \rightarrow \ell\ell$ background in $\tau_\ell\tau_\ell$.
- **UE_qq** – Uncertainty on the underlying event for qq -initiated signal processes (VBF, VH).
- **JES_FlavResp** – Uncertainty on the jet energy scale accounting for jet flavour composition.
- **BR_tautau** – Theoretical uncertainty on the $H \rightarrow \tau\tau$ branching ratio.
- **JES_1112_Detector1/2** – Jet energy scale uncertainty from imperfect detector simulation.
- **ANA_EMB_MFS_2012** – $Z \rightarrow \mu\mu$ selection systematic of embedding background, estimated by varying the isolation requirement of muons.
- **TES_Model_2012** – Tau energy scale uncertainty component from modelling uncertainties.
- **QCDscale_ggH_m12** – QCD scale uncertainty for ggH events with ≥ 2 jets.

- **TAU_EFAKE_2012** – Uncertainty on the $e \rightarrow \tau$ fake rate.
- **TAU_ID_2012** – Uncertainty on the tau identification scale factor.
- **JES_2012_PileRho_TAU_QG** – Jet energy scale uncertainty from variation with pile-up for qg -initiated processes.
- **UE_qq_BDT** – Shape uncertainty on underlying event for qq -initiated signal processes (VBF, VH) due to the introduction of the BDT selection cut.
- **JES_2012_Statistical2/3** – Jet energy scale uncertainty from statistical uncertainty on JES determination.
- **JES_2012_Modelling3** – Jet energy scale uncertainty from imperfect modelling when measuring JES.
- **MU_SCALE** – Muon energy scale systematic.
- **ANA_LL12_Zll_vbf_DETAAJ** – Systematic uncertainty on $|\Delta\eta_{jj}|$ reweighting applied to $Z \rightarrow \ell\ell$ ($l \rightarrow \tau$) events for $\tau_\ell\tau_\ell$ in the VBF category.
- **norm_LH12_Top_bv** – Freely floating normalisation factor on the top background in $\tau_\ell\tau_{\text{had}}$, constrained by single bin top CR.

Appendix C

Effects of anomalous couplings in $H \rightarrow W^+W^-$ decays

$(VBF)H \rightarrow W^+W^-$ decays constitute a non-negligible fraction of the total signal event yield in the $\tau_\ell\tau_\ell$ signal region. The number of total signal events increases by approximately 20% when adding these events to $(VBF)H \rightarrow \tau^+\tau^-$. The analysis looks for anomalous CP-odd Higgs couplings to vector bosons, which can therefore not be ignored in the $H \rightarrow WW$ decay vertex, in contrast to processes where the Higgs boson decays into two τ leptons.

In order to disentangle the effects of anomalous couplings in production and decay, a study of $H \rightarrow WW$ has been performed in gluon fusion production instead of VBF. The study makes use of the azimuthal angle $\Delta\phi(e, \mu)$, which is sensitive to CP-violating couplings in the decay vertex. Figure C.1 shows distributions of $\Delta\phi(e, \mu)$ in the $\tau_\ell\tau_\ell$ decay channel for different \tilde{d} values directly generated with MG5_AMC@NLO [68]. Additionally, the Standard Model sample has been re-weighted using $\tilde{d} \approx 1.0$, based on MG5 matrix elements to leading order. The re-weighting procedure is seen to be valid, since the re-weighted sample is compatible with the generated one. The impact of small \tilde{d} admixtures on $\Delta\phi(e, \mu)$ is minor within the relevant range of \tilde{d} values.

The validated re-weighting procedure is then used to estimate the effect of anomalous couplings in the $H \rightarrow WW$ decay on the Optimal Observable. Figure C.2 compares a pure Standard Model signal sample with a re-weighted Standard Model sample, introducing anomalous couplings only in the decay, using a \tilde{d} value of 0.5. The distributions are compatible within the statistical uncertainties, and the effect of anomalous couplings in the Higgs boson decay into two vector bosons is therefore considered to be negligible. In the interest of simplicity, no additional re-weighting procedure with respect to the decay is applied in the analysis.

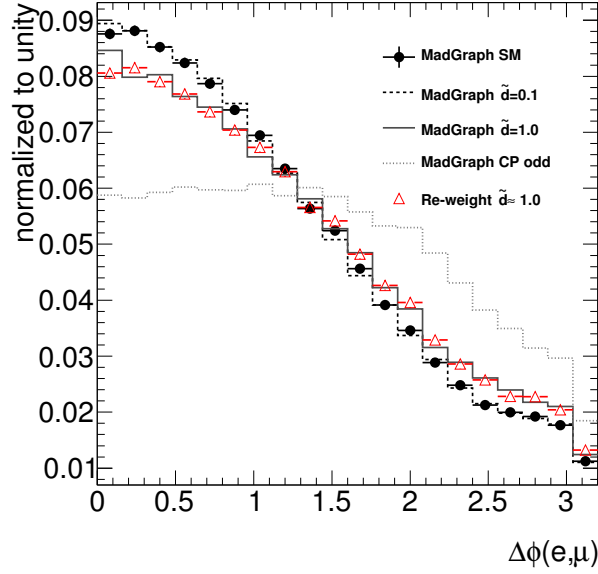


Figure C.1: Comparison of $\Delta\phi(e, \mu)$ in the $\tau_\ell\tau_\ell$ decay channel for different \tilde{d} values using $H \rightarrow WW$ events from gluon fusion production. The samples are generated with MG5. The re-weighting is also based on MG5 and is compatible with the directly generated sample for the corresponding value of \tilde{d} [4].

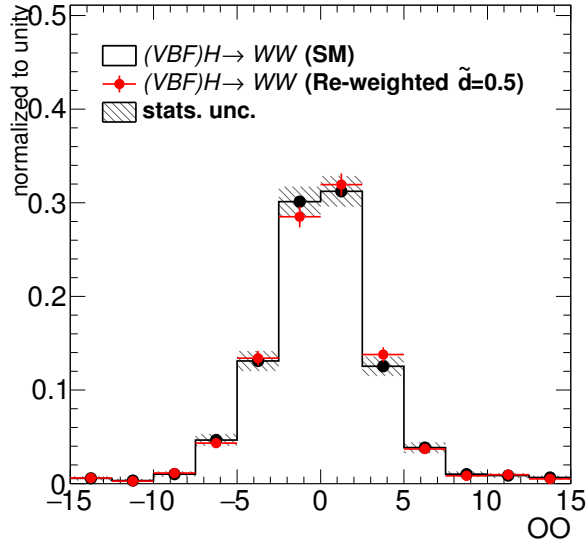


Figure C.2: Comparison of the Optimal Observable in the $\tau_\ell\tau_\ell$ decay channel for a pure Standard Model sample and a re-weighted Standard Model $(VBF)H \rightarrow W^+W^-$ sample. Anomalous couplings are only introduced in the decay with a \tilde{d} value of 0.5. The production vertex is left Standard Model-like [4].

Bibliography

- [1] 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. feb 2016. [dx.doi.org/10.1140/epjc/s10052-016-4499-5](https://doi.org/10.1140/epjc/s10052-016-4499-5).
- [2] ATLAS Collaboration. Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector. *Journal of High Energy Physics*, 2015(4):117, apr 2015. <http://arxiv.org/abs/1501.04943>.
- [3] I. Antcheva, M. Ballintijn, B. Bellenot, et al. ROOT - A C++ framework for petabyte data storage, statistical analysis and visualization. *Computer Physics Communications*, 180(12):2499–2512, dec 2009. <http://arxiv.org/abs/1508.07749>.
- [4] Private communications with Christian Schillo, christian.schillo@cern.ch.
- [5] K.A. Olive et al. (Particle Data Group). *The Review of Particle Physics*. Chin. Phys. C, 2015.
- [6] en.wikipedia.org/wiki/Standard_Model. Date: 2016-04-18.
- [7] W. Pauli. Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. *Zeitschrift für Physik*, 31(1):765–783, feb 1925. <http://link.springer.com/10.1007/BF02980631>.
- [8] Alberto Zannoni. On the Quantization of the Monoatomic Ideal Gas. dec 1999. <http://arxiv.org/abs/cond-mat/9912229>.
- [9] P. A. M. Dirac. On the Theory of Quantum Mechanics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 112(762), 1926.

- [10] Y. Fukuda, T. Hayakawa, E. Ichihara, et al. Evidence for Oscillation of Atmospheric Neutrinos. *Physical Review Letters*, 81(8):1562–1567, aug 1998. <http://arxiv.org/abs/hep-ex/9807003>.
- [11] LHCb Collaboration. Observation of J/ψ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays. *Physical review letters*, 115(7):072001, aug 2015. <http://arxiv.org/abs/1507.03414>.
- [12] The ALEPH Collaboration, the DELPHI Collaboration, the L3 Collaboration, et al. Precision electroweak measurements on the Z resonance. *Physics Reports*, 427(5-6):257–454, may 2006. <http://arxiv.org/abs/hep-ex/0509008>.
- [13] ATLAS Collaboration. Search for down-type fourth generation quarks with the ATLAS detector in events with one lepton and hadronically decaying W bosons. *Physical review letters*, 109(3):032001, jul 2012. <http://arxiv.org/abs/1202.6540>.
- [14] ATLAS Collaboration. Search for same-sign top-quark production and fourth-generation down-type quarks in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2012(4):69, apr 2012. <http://arxiv.org/abs/1202.5520>.
- [15] Alexander Lenz. Constraints on a Fourth Generation of Fermions from Higgs Boson Searches. *Advances in High Energy Physics*, 2013:1–13, 2013. <http://inspirehep.net/record/1223750>.
- [16] David J. Gross and Frank Wilczek. Ultraviolet Behavior of Non-Abelian Gauge Theories. *Physical Review Letters*, 30(26):1343–1346, jun 1973. <http://link.aps.org/doi/10.1103/PhysRevLett.30.1343>.
- [17] H. David Politzer. Reliable Perturbative Results for Strong Interactions? *Physical Review Letters*, 30(26):1346–1349, jun 1973. <http://link.aps.org/doi/10.1103/PhysRevLett.30.1346>.
- [18] G. 't Hooft and M. Veltman. Regularization and renormalization of gauge fields. *Nuclear Physics B*, 44(1):189–213, 1972.
- [19] John F. Donoghue. General relativity as an effective field theory: The leading quantum corrections. may 1994. <http://arxiv.org/abs/gr-qc/9405057>.

- [20] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, sep 2012. <http://arxiv.org/abs/1207.7214>.
- [21] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, sep 2012. <http://arxiv.org/abs/1207.7235>.
- [22] ATLAS Collaboration. Combined Measurement of the Higgs Boson Mass in p p Collisions at $s = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Physical Review Letters*, 114(19):191803, may 2015. <http://arxiv.org/abs/1503.07589>.
- [23] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Addison-Wesley Publishing Company, 1995.
- [24] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2 pi Decay of the K(2)0 Meson. *Physical Review Letters*, 13(4):138–140, jul 1964. <https://inspirehep.net/record/11878>.
- [25] C. N. Yang and R. L. Mills. Conservation of Isotopic Spin and Isotopic Gauge Invariance. *Physical Review*, 96(1):191–195, oct 1954. <http://inspirehep.net/record/12285>.
- [26] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. Advantages of the color octet gluon picture. *Physics Letters B*, 47(4):365–368, nov 1973. <http://linkinghub.elsevier.com/retrieve/pii/0370269373906254>.
- [27] CMS CMS Collaboration. Measurement of the inclusive 3-jet production differential cross section in proton-proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range. dec 2014. <http://arxiv.org/abs/1412.1633>.
- [28] T. D. Lee and C. N. Yang. Question of Parity Conservation in Weak Interactions. *Physical Review*, 104(1):254–258, oct 1956. <http://journals.aps.org/pr/abstract/10.1103/PhysRev.104.254>.
- [29] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental Test of Parity Conservation in Beta Decay. *Physical Review*, 105(4):1413–1415, feb 1957. <https://inspirehep.net/record/28182>.
- [30] David J. Griffiths. *Introduction to Elementary Particles*. John Wiley & Sons, Inc., 1987.

- [31] Francis Halzen and Alan D. Martin. *Quarks and Leptons, An Introductory Course in Modern Particle Physics*. 1984.
- [32] Makoto Kobayashi and Toshihide Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, feb 1973. <http://inspirehep.net/record/81350>.
- [33] S. W. Herb, D. C. Hom, L. M. Lederman, et al. Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions. *Physical Review Letters*, 39(5):252–255, aug 1977. <http://link.aps.org/doi/10.1103/PhysRevLett.39.252>.
- [34] UA1 Collaboration. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV. *Physics Letters B*, 122(1):103–116, feb 1983. <http://inspirehep.net/record/188736>.
- [35] UA2 Collaboration. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN p collider. *Physics Letters B*, 122(5-6):476–485, mar 1983. <http://inspirehep.net/record/189048>.
- [36] UA1 Collaboration. Experimental observation of lepton pairs of invariant mass around 95 GeV/c² at the CERN SPS collider. *Physics Letters B*, 126(5):398–410, 1983. <http://inspirehep.net/record/190338>.
- [37] UA2 Collaboration. Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN pp collider. *Physics Letters B*, 129(1):130–140, 1983. <https://inspirehep.net/record/191740>.
- [38] D0 Collaboration. Search for High Mass Top Quark Production in p anti-p Collisions at $\sqrt{s} = 1.8$ TeV. nov 1994. <http://arxiv.org/abs/hep-ex/9411001>.
- [39] CDF Collaboration. Observation of Top Quark Production in $\bar{p}p$ Collisions. mar 1995. <http://arxiv.org/abs/hep-ex/9503002>.
- [40] DONUT Collaboration. Observation of Tau Neutrino Interactions. dec 2000. <http://arxiv.org/abs/hep-ex/0012035>.
- [41] project-gfitter.web.cern.ch/project-gfitter/Standard_Model. Date: 2016-11-09.

- [42] Henning Flaecher, Martin Goebel, Johannes Haller, et al. Revisiting the Global Electroweak Fit of the Standard Model and Beyond with Gfitter. nov 2008. <http://arxiv.org/abs/0811.0009>.
- [43] A. D. Sakharov. Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe. *Soviet Physics Uspekhi*, 34(5):392–393, may 1991. <http://stacks.iop.org/0038-5670/34/i=5/a=A08?key=crossref.eaf487d3e170a7b4114a0b99ed598c9a>.
- [44] Bruce T. Cleveland, Timothy Daily, Raymond Davis, Jr., et al. Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector. *The Astrophysical Journal*, 496(1):505–526, mar 1998. <https://inspirehep.net/record/471829>.
- [45] GNO GNO COLLABORATION, M. Altmann, M. Balata, et al. Complete results for five years of GNO solar neutrino observations. apr 2005. <http://arxiv.org/abs/hep-ex/0504037>.
- [46] KamLAND KamLAND Collaboration. First Results from KamLAND: Evidence for Reactor Anti-Neutrino Disappearance. dec 2002. <http://arxiv.org/abs/hep-ex/0212021>.
- [47] The Super-Kamiokande Collaboration and Y. Fukuda et Al. Evidence for oscillation of atmospheric neutrinos. jul 1998. <http://arxiv.org/abs/hep-ex/9807003>.
- [48] T2K Collaboration. Observation of Electron Neutrino Appearance in a Muon Neutrino Beam. nov 2013. <http://arxiv.org/abs/1311.4750>.
- [49] OPERA Collaboration. Discovery of tau neutrino appearance in the CNGS neutrino beam with the OPERA experiment. 2015. <https://arxiv.org/abs/1507.01417>.
- [50] P. Fayet. Supersymmetry and weak, electromagnetic and strong interactions. *Physics Letters B*, 64(2):159–162, sep 1976. <http://linkinghub.elsevier.com/retrieve/pii/0370269376903191>.
- [51] P. Fayet. Spontaneously broken supersymmetric theories of weak, electromagnetic and strong interactions. *Physics Letters B*, 69(4):489–494, jan 1977. <http://linkinghub.elsevier.com/retrieve/pii/0370269377908528>.
- [52] D. I. Kazakov. Beyond the Standard Model (In Search of Supersymmetry). dec 2000. <http://arxiv.org/abs/hep-ph/0012288>.

- [53] Planck Collaboration, P. A. R. Ade, N. Aghanim, et al. Planck 2013 results. XVI. Cosmological parameters. mar 2013. <http://arxiv.org/abs/1303.5076>.
- [54] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. jan 2011. <http://arxiv.org/abs/1101.0593>.
- [55] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 2. Differential Distributions. jan 2012. <http://arxiv.org/abs/1201.3084>.
- [56] LHC Higgs Cross Section Working Group. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. jul 2013. <http://arxiv.org/abs/1307.1347>.
- [57] A. Djouadi, J. Kalinowski, and M. Spira. HDECAY: a Program for Higgs Boson Decays in the Standard Model and its Supersymmetric Extension. apr 1997. <http://arxiv.org/abs/hep-ph/9704448>.
- [58] Michael Spira. QCD Effects in Higgs Physics. may 1997. <http://arxiv.org/abs/hep-ph/9705337>.
- [59] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber. Precise predictions for the Higgs-boson decay $H \rightarrow WW/ZZ \rightarrow 4$ leptons. apr 2006. <http://arxiv.org/abs/hep-ph/0604011>.
- [60] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber. Radiative corrections to the semileptonic and hadronic Higgs-boson decays $H \rightarrow WW/ZZ \rightarrow 4$ fermions. nov 2006. <http://arxiv.org/abs/hep-ph/0611234>.
- [61] W. Buchmüller and D. Wyler. Effective lagrangian analysis of new interactions and flavour conservation. *Nuclear Physics B*, 268(3-4):621–653, may 1986. <http://linkinghub.elsevier.com/retrieve/pii/0550321386902622>.
- [62] S. Schael, R. Barate, R. Brunelière, et al. Improved measurement of the triple gauge-boson couplings gamma WW and ZWW in e+e- collisions. *Physics Letters B*, 614(1-2):7–26, may 2005. <http://inspirehep.net/record/669198>.
- [63] G. Abbiendi et al., F. Odorici, T.E. Marchant, et al. Measurement of W boson polarisations and CP-violating triple gauge couplings from $W^+ W^-$ production at LEP. *The European Physical Journal C*, 19(2):229–240, feb 2001. <https://inspirehep.net/record/533112>.

- [64] L3 Collaboration. Search for anomalous couplings in the Higgs sector at LEP. aug 2000. <http://arxiv.org/abs/hep-ex/0008023>.
- [65] ATLAS Collaboration. Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector. jun 2015. <http://arxiv.org/abs/1506.05669>.
- [66] T. Figy, V. Hankele, G. Klamke, and D. Zeppenfeld. Anomalous Higgs boson couplings in vector boson fusion at the CERN LHC. sep 2006. <http://arxiv.org/abs/hep-ph/0609075>.
- [67] Markus Schumacher. Determination of the CP Quantum Numbers of the Higgs Boson and test of CP invariance in the Higgs-strahlung process at a future e^+e^- collider. *LC-PHSM-2001-003*, 2001. <http://inspirehep.net/record/559389>.
- [68] J. Alwall, R. Frederix, S. Frixione, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. may 2014. <http://arxiv.org/abs/1405.0301>.
- [69] D. Atwood and A. Soni. Analysis for magnetic moment and electric dipole moment form factors of the top quark via $e^+e^- \rightarrow t\bar{t}$. *Physical Review D*, 45(7):2405–2413, apr 1992. <http://inspirehep.net/record/322206>.
- [70] M. Davier, L. Duflot, F. Le Diberder, and A. Rougé. The optimal method for the measurement of tau polarization. *Physics Letters B*, 306(3-4):411–417, jun 1993. <http://linkinghub.elsevier.com/retrieve/pii/037026939390101M>.
- [71] M. Diehl and O. Nachtmann. Optimal observables for the measurement of three gauge boson couplings in $e^+e^- \rightarrow W^+W^-$. *Zeitschrift fur Physik C Particles and Fields*, 62(3):397–411, sep 1994. <http://link.springer.com/10.1007/BF01555899>.
- [72] M. Diehl and O. Nachtmann. Anomalous triple gauge couplings in $e^+e^- \rightarrow W^+W^-$ and 'optimal' strategies for their measurement. jan 1997. <http://arxiv.org/abs/hep-ph/9702208>.
- [73] M. Diehl, O. Nachtmann, and F. Nagel. Triple gauge couplings in polarised $e^-e^+ \rightarrow W^-W^+$ and their measurement using optimal observables. *The European Physical Journal C*, 27(3):375–397, apr 2003. <http://link.springer.com/10.1140/epjc/s2002-01096-y>.

- [74] R. Akers, G. Alexander, J. Allison, et al. A test of CP-invariance in $Z^0 \rightarrow \tau^+\tau^-$ using optimal observables. *Zeitschrift fur Physik C Particles and Fields*, 66(1-2):31–44, mar 1995. <http://link.springer.com/10.1007/BF01496578>.
- [75] K. Ackerstaff et al., F. Odorici, F.G. Oakham, et al. Search for CP violation in $Z^0 \rightarrow \tau^+\tau^-$ and an upper limit on the weak dipole moment of the τ lepton. *Zeitschrift fur Physik C Particles and Fields*, 74(3):403–412, may 1997. <http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s002880050403>.
- [76] The OPAL Collaboration. W+W- production and triple gauge boson couplings at LEP energies up to 183 GeV. nov 1998. <http://arxiv.org/abs/hep-ex/9811028>.
- [77] The OPAL Collaboration. Measurement of triple gauge boson couplings from WW production at LEP energies up to 189 GeV. sep 2000. <http://arxiv.org/abs/hep-ex/0009022>.
- [78] The OPAL Collaboration. Measurement of charged current triple gauge boson couplings using W pairs at LEP. aug 2003. <http://arxiv.org/abs/hep-ex/0308067>.
- [79] M Ciccolini, A Denner, and S Dittmaier. Strong and electroweak corrections to the production of a Higgs boson+2 jets via weak interactions at the Large Hadron Collider. *Physical review letters*, 99(16):161803, oct 2007. <http://arxiv.org/abs/0707.0381>.
- [80] M. Ciccolini, Denner A., and Dittmaier S. Electroweak and QCD corrections to Higgs production via vector-boson fusion at the CERN LHC. *Physical Review D*, 77(1):013002, jan 2008. <http://arxiv.org/abs/0710.4749>.
- [81] Ansgar Denner, Stefan Dittmaier, Stefan Kallweit, and Alexander Mück. HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders. *Computer Physics Communications*, 195:161–171, oct 2015. <http://arxiv.org/abs/1412.5390>.
- [82] A. Elagin, P. Murat, A. Pranko, and A. Safonov. A New Mass Reconstruction Technique for Resonances Decaying to di-tau. *Nucl. Instrum. Meth. A654*, dec 2010. <http://arxiv.org/abs/1012.4686>.

- [83] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *J. Instrum.*, 3:S08003. 437 p, 2008. Also published by CERN Geneva in 2010.
- [84] Lyndon Evans and Philip Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001–S08001, aug 2008. <http://inspirehep.net/record/796247>.
- [85] Forthommel. *Map of the CERN accelerator complex*. Date: 2012-09-07. <http://en.wikipedia.org/wiki/CERN>,.
- [86] John C. Collins, Davison E. Soper, and George Sterman. Factorization of Hard Processes in QCD. sep 2004. <http://arxiv.org/abs/hep-ph/0409313>.
- [87] J. M. Campbell, J. W. Huston, and W. J. Stirling. Hard Interactions of Quarks and Gluons: a Primer for LHC Physics. nov 2006. <http://arxiv.org/abs/hep-ph/0611148>.
- [88] Hung-Liang Lai, Marco Guzzi, Joey Huston, et al. New parton distributions for collider physics. jul 2010. <http://arxiv.org/abs/1007.2241>.
- [89] J. Pumplin, D. R. Stump, J. Huston, et al. New Generation of Parton Distributions with Uncertainties from Global QCD Analysis. jan 2002. <http://arxiv.org/abs/hep-ph/0201195>.
- [90] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. jan 2009. <http://arxiv.org/abs/0901.0002>.
- [91] K.A. Olive et al. (Particle Data Group). *Monte Carlo Techniques*. Chin. Phys. C38, 2014.
- [92] ATLAS Collaboration. The ATLAS Inner Detector commissioning and calibration. apr 2010. <http://arxiv.org/abs/1004.5293>.
- [93] R Hauser. The ATLAS Data Acquisition and High Level Trigger Systems: Experience and Upgrade Plans. Technical report, 2012. <http://cds.cern.ch/record/1497132>.
- [94] The ATLAS Collaboration. ATLAS Twiki Public Results: twiki.cern.ch/twiki/bin/view/AtlasPublic. Accessed September 2016.
- [95] ATLAS Collaboration. ATLAS computing: Technical design report. Technical report, 2005. <http://cds.cern.ch/record/837738>.

- [96] T Cornelissen, M Elsing, W Liebig, S Fleischmann, and E Moyse. Concepts, Design and Implementation of the ATLAS New Tracking (NEWT). Technical report, 2007. <http://cds.cern.ch/record/1020106>.
- [97] ATLAS Collaboration. Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment, 2012. <https://inspirehep.net/record/1204277>.
- [98] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. mar 2016. <http://arxiv.org/abs/1603.02934>.
- [99] W Lampl, P Loch, S Menke, et al. Calorimeter Clustering Algorithms, 2008. <http://cdsweb.cern.ch/record/1099735>.
- [100] T Barillari, A Jantsch, P Stavina, et al. Local Hadronic Calibration, 2008. <https://cds.cern.ch/record/1112035>.
- [101] ATLAS Collaboration. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. dec 2011. <http://arxiv.org/abs/1112.6426>.
- [102] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A Brief Introduction to PYTHIA 8.1. oct 2007. <http://arxiv.org/abs/0710.3820>.
- [103] G. Corcella, I. G. Knowles, G. Marchesini, et al. HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes). nov 2000. <http://arxiv.org/abs/hep-ph/0011363>.
- [104] Ryan Atkin, Hodgkinson M., Schieferdecker P., et al. Review of jet reconstruction algorithms. *Journal of Physics: Conference Series*, 645(1):012008, oct 2015. <https://inspirehep.net/record/1398448>.
- [105] Victor Coco, Gregory Soyez, Juan Rojo-Chacon, Christian Sander, and Pierre-Antoine Delsart. Jets and jet algorithms. pages 182–204, 2009. <https://inspirehep.net/record/866539>.
- [106] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti- k_t jet clustering algorithm. feb 2008. <http://arxiv.org/abs/0802.1189>.
- [107] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. nov 2011. <http://arxiv.org/abs/1111.6097>.

- [108] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Catchment Area of Jets. feb 2008. <http://arxiv.org/abs/0802.1188>.
- [109] ATLAS Collaboration. Pile-up subtraction and suppression for jets in ATLAS. *ATLAS-CONF-2013-083*, 2013. <https://cds.cern.ch/record/1570994>.
- [110] ATLAS Collaboration. Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data. *ATLAS-CONF-2012-020*, 2012. <http://cds.cern.ch/record/1430034>.
- [111] ATLAS Collaboration. Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data. *ATLAS-CONF-2011-102*, 2011. <http://cds.cern.ch/record/1369219>.
- [112] Marc Lehmann. b-Tagging Algorithms and their Performance at ATLAS. sep 2008. <http://arxiv.org/abs/0809.4896>.
- [113] ATLAS Collaboration. Calibration of *b*-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment, 2014. <https://cds.cern.ch/record/1664335>.
- [114] ATLAS Collaboration. Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb^{-1} of Data from the ATLAS Detector, 2012. <http://cds.cern.ch/record/1435197>.
- [115] ATLAS Collaboration. Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data. In *ATLAS-CONF-2014-032*, jun 2014. <http://cds.cern.ch/record/1706245>.
- [116] ATLAS Collaboration. Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung, 2012. <https://inspirehep.net/record/1204271>.
- [117] ATLAS Collaboration. Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data. apr 2014. <http://arxiv.org/abs/1404.2240>.
- [118] ATLAS Collaboration. Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data. oct 2011. <http://arxiv.org/abs/1110.3174>.
- [119] ATLAS Collaboration. Expected electron performance in the ATLAS experiment, 2011. <https://cds.cern.ch/record/1345327>.

- [120] ATLAS Collaboration. Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s} = 8$ TeV. *The European physical journal. C, Particles and fields*, 75(3):120, jan 2014. <http://arxiv.org/abs/1408.3179>.
- [121] ATLAS Collaboration. Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data, 2013. <http://cds.cern.ch/record/1580207>.
- [122] ATLAS Collaboration. Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s} = 8$ TeV. *The European physical journal. C, Particles and fields*, 75(7):303, jan 2014. <http://arxiv.org/abs/1412.7086>.
- [123] ATLAS Collaboration. Identification of the Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector, 2013. <http://cds.cern.ch/record/1562839>.
- [124] ATLAS Collaboration. Performance of the Reconstruction and Identification of Hadronic Tau Decays in ATLAS with 2011 Data, 2012. <http://cds.cern.ch/record/1485531>.
- [125] ATLAS Collaboration. Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC in 2011, 2012. <https://inspirehep.net/record/1204263>.
- [126] ATLAS Collaboration. Determination of the tau energy scale and the associated systematic uncertainty in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC in 2012, 2013. <https://cds.cern.ch/record/1544036>.
- [127] ATLAS Collaboration. Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV. 2013. <http://inspirehep.net/record/1260962>.
- [128] ATLAS Collaboration. Pile-up Suppression in Missing Transverse Momentum Reconstruction in the ATLAS Experiment in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV, 2014. <http://cds.cern.ch/record/1702055>.
- [129] M. Ciccolini, A. Denner, and S. Dittmaier. Strong and electroweak corrections to the production of Higgs+2jets via weak interactions at the LHC. jul 2007. <http://arxiv.org/abs/0707.0381>.

- [130] Mariano Ciccolini, Ansgar Denner, and Stefan Dittmaier. Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC. oct 2007. <http://arxiv.org/abs/0710.4749>.
- [131] K. Arnold, M. Bahr, G. Bozzi, et al. VBFNLO: A parton level Monte Carlo for processes with electroweak bosons. nov 2008. <http://arxiv.org/abs/0811.4559>.
- [132] A. Djouadi, M. Spira, and P.M. Zerwas. Production of Higgs bosons in proton colliders. QCD corrections. *Physics Letters B*, 264(3-4):440–446, aug 1991. <http://linkinghub.elsevier.com/retrieve/pii/037026939190375Z>.
- [133] S. Dawson. Radiative corrections to Higgs boson production. *Nuclear Physics B*, 359(2-3):283–300, aug 1991. <http://linkinghub.elsevier.com/retrieve/pii/0550321391900612>.
- [134] M. Spira, A. Djouadi, D. Graudenz, and R.M. Zerwas. Higgs boson production at the LHC. *Nuclear Physics B*, 453(1-2):17–82, oct 1995. <https://arxiv.org/abs/hep-ph/9504378>.
- [135] Robert V. Harlander and William B. Kilgore. Next-to-Next-to-Leading Order Higgs Production at Hadron Colliders. *Physical Review Letters*, 88(20):201801, may 2002. <https://arxiv.org/abs/hep-ph/0201206>.
- [136] Charalampos Anastasiou and Kirill Melnikov. Higgs boson production at hadron colliders in NNLO QCD. *Nuclear Physics B*, 646(1):220–256, 2002. <https://arxiv.org/abs/hep-ph/0207004>.
- [137] V. Ravindran, J. Smith, and W.L. van Neerven. NNLO corrections to the total cross section for Higgs boson production in hadron–hadron collisions. *Nuclear Physics B*, 665:325–366, 2003. <https://arxiv.org/abs/hep-ph/0302135>.
- [138] Oliver Brein, Abdelhak Djouadi, and Robert Harlander. NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders. 2004. <https://arxiv.org/abs/hep-ph/0307206>.
- [139] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini. Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO. mar 2009. <http://arxiv.org/abs/0903.2120>.
- [140] Stefano Catani and Massimiliano Grazzini. Next-to-Next-to-Leading-Order Subtraction Formalism in Hadron Collisions and its Application to

- Higgs-Boson Production at the Large Hadron Collider. *Physical Review Letters*, 98(22):222002, may 2007. <http://link.aps.org/doi/10.1103/PhysRevLett.98.222002>.
- [141] T. Gleisberg, S. Hoeche, F. Krauss, et al. Event generation with SHERPA 1.1. nov 2008. <http://arxiv.org/abs/0811.4622>.
 - [142] Matteo Cacciari, Michal Czakon, Michelangelo L. Mangano, Alexander Mitov, and Paolo Nason. Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation. nov 2011. <http://arxiv.org/abs/1111.5869>.
 - [143] Peter Baernreuther, Michal Czakon, and Alexander Mitov. Percent level precision physics at the Tevatron: first genuine NNLO QCD corrections to $q\bar{q} \rightarrow t\bar{t} + X$. apr 2012. <http://arxiv.org/abs/1204.5201>.
 - [144] Michal Czakon and Alexander Mitov. NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction. oct 2012. <http://arxiv.org/abs/1210.6832>.
 - [145] Michal Czakon, Paul Fiedler, and Alexander Mitov. The total top quark pair production cross-section at hadron colliders through $O(\alpha_S^4)$. mar 2013. <http://arxiv.org/abs/1303.6254>.
 - [146] Michal Czakon and Alexander Mitov. Top++: a program for the calculation of the top-pair cross-section at hadron colliders. dec 2011. <http://arxiv.org/abs/1112.5675>.
 - [147] Nikolaos Kidonakis. Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-. may 2010. <http://arxiv.org/abs/1005.4451>.
 - [148] Nikolaos Kidonakis. NNLL resummation for s-channel single top quark production. jan 2010. <http://arxiv.org/abs/1001.5034>.
 - [149] Nikolaos Kidonakis. Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production. mar 2011. <http://arxiv.org/abs/1103.2792>.
 - [150] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Vector boson pair production at the LHC. apr 2011. <http://arxiv.org/abs/1105.0020>.

- [151] T. Binoth, M. Ciccolini, N. Kauer, and M. Krämer. Gluon-induced W-boson pair production at the LHC. nov 2006. <http://arxiv.org/abs/hep-ph/0611170>.
- [152] Paolo Nason and Carlo Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. nov 2009. <http://arxiv.org/abs/0911.5299>.
- [153] Keith Hamilton, Paolo Nason, and Giulia Zanderighi. MINLO: multi-scale improved NLO. *Journal of High Energy Physics*, 2012(10):155, oct 2012. <http://arxiv.org/abs/1206.3572>.
- [154] S. Jadach, Z. Was, R. Decker, and J.H. Kühn. The τ decay library TAUOLA, version 2.4. *Computer Physics Communications*, 76(3):361–380, aug 1993. <http://linkinghub.elsevier.com/retrieve/pii/001046559390061G>.
- [155] N. Davidson, T. Przedzinski, and Z. Was. PHOTOS Interface in C++; Technical and Physics Documentation. nov 2010. <http://arxiv.org/abs/1011.0937>.
- [156] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. jun 2002. <http://arxiv.org/abs/hep-ph/0206293>.
- [157] Michelangelo L. Mangano, Mauro Moretti, and Roberto Pittau. Multijet matrix elements and shower evolution in hadronic collisions: $W b \bar{b} + n$ jets as a case study. aug 2001. <http://arxiv.org/abs/hep-ph/0108069>.
- [158] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour. Multiparton Interactions in Photoproduction at HERA. jan 1996. <http://arxiv.org/abs/hep-ph/9601371>.
- [159] Paolo Nason. A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms. sep 2004. <http://arxiv.org/abs/hep-ph/0409146>.
- [160] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. sep 2007. <http://arxiv.org/abs/0709.2092>.
- [161] Borut Paul Kersevan and Elzbieta Richter-Was. The Monte Carlo Event Generator AcerMC 2.0 with Interfaces to PYTHIA 6.2 and HERWIG 6.5. may 2004. <http://arxiv.org/abs/hep-ph/0405247>.

- [162] B. P. Kersevan I. Hinchliffe. A Consistent Prescription for the Production Involving Massive Quarks in Hadron Collisions. mar 2006. <http://arxiv.org/abs/hep-ph/0603068>.
- [163] T. Binoth, M. Ciccolini, N. Kauer, and M. Krämer. Gluon-induced W-boson pair production at the LHC. nov 2006. <http://arxiv.org/abs/hep-ph/0611170>.
- [164] J. Alwall, A. Ballestrero, P. Bartalini, et al. A standard format for Les Houches Event Files. sep 2006. <http://arxiv.org/abs/hep-ph/0609017>.
- [165] Fabio Maltoni, Kentarou Mawatari, and Marco Zaro. Higgs characterisation via vector-boson fusion and associated production: NLO and parton-shower effects. nov 2013. <http://arxiv.org/abs/1311.1829>.
- [166] Andy Buckley, Jonathan Butterworth, David Grellscheid, et al. Rivet user manual, 2010. <https://arxiv.org/abs/1003.0694>.
- [167] N Möser, J Kroseberg, M Schumacher, M Schmitz, and N Wermes. Estimation of $Z \rightarrow \tau\tau$ Background in VBF $H \rightarrow \tau\tau$ Searches from $Z \rightarrow \mu\mu$ Data using an Embedding Technique, 2009. <http://cds.cern.ch/record/1201000>.
- [168] Z. Czynzula, T. Przedzinski, and Z. Was. TauSpinner program for studies on spin effect in tau production at the LHC. *The European Physical Journal C*, 72(4):1988, apr 2012. <http://www.springerlink.com/index/10.1140/epjc/s10052-012-1988-z>.
- [169] L. Breinman, J. Friedman, C. J. Stone, and R. A. Olshen. *Classification and Regression Trees*. CRC Press, 1984.
- [170] A. Hoecker, P. Speckmayer, J. Stelzer, et al. TMVA - Toolkit for Multivariate Data Analysis. mar 2007. <http://arxiv.org/abs/physics/0703039>.
- [171] Jerome H. Friedman. Stochastic gradient boosting. *Computational Statistics and Data Analysis*, 38(4):367–378, 2002.
- [172] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2):1–19, feb 2011. <http://arxiv.org/abs/1007.1727>.
- [173] Glenn Cowan. *Statistical Data Analysis*. Oxford University Press, 1998.

- [174] Kyle Cranmer, Akira Shibata, Wouter Verkerke, Lorenzo Moneta, and George Lewis. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. jun 2012. <https://inspirehep.net/record/1236448>.
- [175] Lorenzo Moneta, Kevin Belasco, Kyle Cranmer, et al. The RooStats Project. page 11, sep 2010. <http://arxiv.org/abs/1009.1003>.
- [176] ATLAS Collaboration., Collaboration. CMS, and The LHC Higgs Combination Group. Procedure for the LHC Higgs boson search combination in summer 2011. 2011. <https://cds.cern.ch/record/1379837>.
- [177] ATLAS Collaboration. Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC. *The European physical journal. C, Particles and fields*, 73(8):2518, jan 2013. <http://arxiv.org/abs/1302.4393>.
- [178] ATLAS Collaboration. Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *The European physical journal. C, Particles and fields*, 75:17, jan 2014. <http://arxiv.org/abs/1406.0076>.
- [179] ATLAS Collaboration. Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector. *The European physical journal. C, Particles and fields*, 73(3):2306, jan 2012. <http://arxiv.org/abs/1210.6210>.
- [180] The ATLAS Collaboration. ATLAS Twiki Public Results: twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmisApproved2013JESUncertainty. Accessed May 2016.
- [181] ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data. *The European Physical Journal C*, 74(10):3071, oct 2014. <http://arxiv.org/abs/1407.5063>.
- [182] ATLAS Collaboration. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data. *The European physical journal. C, Particles and fields*, 74(11):3130, jan 2014. <http://arxiv.org/abs/1407.3935>.
- [183] ATLAS Collaboration. Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at $\sqrt{s} = 7$ TeV in 2011. 2012. <http://inspirehep.net/record/1204270>.

- [184] ATLAS Collaboration. Calibration of the performance of b -tagging for c and light-flavour jets in the 2012 ATLAS data, 2014. <https://cds.cern.ch/record/1741020>.
- [185] Bo Andersson, Sandipan Mohanty, and Fredrik Soderberg. Recent Developments in the Lund Model. dec 2002. <http://arxiv.org/abs/hep-ph/0212122>.
- [186] Alexander Kupco. Cluster Hadronization in HERWIG 5.9. jun 1999. <http://arxiv.org/abs/hep-ph/9906412>.
- [187] J. Neyman. Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 236(767):333–380, aug 1937. <http://rsta.royalsocietypublishing.org/content/236/767/333>.
- [188] ATLAS Collaboration. Constraints on non-Standard Model Higgs boson interactions in an effective Lagrangian using differential cross sections measured in the $H \rightarrow \gamma\gamma$ decay channel at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Physics Letters B*, 753:69–85, feb 2016. <http://arxiv.org/abs/1508.02507>.
- [189] Joel Heinrich and Louis Lyons. Systematic Errors. *Annual Review of Nuclear and Particle Science*, 57(1):145–169, nov 2007. <http://www.annualreviews.org/doi/10.1146/annurev.nucl.57.090506.123052>.