Evidence for Standard Model Higgs Boson Decays to Tau Lepton Pairs in the ATLAS Detector Supported by a Search in the Fully Hadronic Final State

by

Edmund Noel Dawe
B.Sc. (Hons.), Mount Allison University, 2008

Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy

in the
Department of Physics
Faculty of Science

SIMON FRASER UNIVERSITY
Fall 2014

APPROVAL

Name: Edmund Noel Dawe

Degree: Doctor of Philosophy

Title of Thesis: *Evidence for Standard Model Higgs Boson Decays to Tau Lepton Pairs in the ATLAS Detector Supported by a Search in the Fully Hadronic Final State*

Examining Committee: Dr. Eldon Emberly
Chair, Associate Professor

Dr. Dugan O'Neil
Senior Supervisor, Professor

Dr. Michel Vetterli
Supervisor, Professor

Dr. Bernd Stelzer
Supervisor, Associate Professor

Dr. Levon Pogosian
Internal Examiner, Associate Professor

Dr. Douglas Bryman
External Examiner
Professor/Warren Chair
Department of Physics and Astronomy
University of British Columbia

Date Approved: December 4, 2014

Partial Copyright Licence

Abstract

This thesis presents the results of a search for Standard Model Higgs bosons decaying to tau lepton pairs, where both tau leptons decay hadronically. The search is performed using 20.3 fb^{-1} of proton-proton collision data with a center-of-mass energy of 8 TeV collected in 2012 by the ATLAS detector at the Large Hadron Collider at CERN.

The search is complicated by a significant source of background events from non-resonant multijet processes and by $Z \to \tau\tau$ events that share similar features with the Higgs signal. The multijet background is estimated from a control region in data, and the $Z \to \tau\tau$ background is estimated from $Z \to \mu\mu$ events in data where the muons are replaced by simulated tau leptons. Other minor backgrounds are estimated from simulation. Signal-sensitive regions in the data are only observed after presenting a satisfactory model of the background processes.

A multivariate analysis using boosted decision trees is optimized to search for Higgs events produced via gluon-gluon fusion and vector boson fusion at m_H = 125 GeV. A cut-based analysis is also developed and presented in support of the multivariate analysis. The compatibility of the multivariate and cut-based analyses at m_H = 125 GeV is evaluated using the bootstrap method.

An excess of events over the expected background is found, with an observed (expected) significance of 2.9 (1.8) standard deviations. The measured signal strength, normalized to the Standard Model expectation is $\mu = 1.77^{+0.93}_{-0.71}$. The combined excess of events from searches in the fully leptonic and semi-leptonic final states corresponds to an observed (expected) significance of 4.5 (3.5) standard deviations and a measured signal strength of $\mu = 1.42^{+0.44}_{-0.38}$. The combined excess provides direct evidence for the coupling of the recently discovered Higgs boson with m_H = 125 GeV to fermions.

To my loving wife, Amy, and dear son, Elliot.

"Everyone wants to be first, but nobody wants to be wrong, and if you're sloppy, someone else will quickly figure it out."

— Sean Carroll

THE PARTICLE AT THE END OF THE UNIVERSE, *2012*

Acknowledgments

I would first like to thank my senior supervisor, Dr. Dugan O'Neil. Dugan's mentorship, counselling, and support have been invaluable to my progress throughout my graduate studies at SFU and in opening up exciting future career opportunities. He has continually challenged me to think of new and creative ways to solve problems, exposed me to new technologies that became essential to this research, taught me how to work with a higher level of efficiency, and encouraged me to take on more significant responsibilities within the ATLAS collaboration. Dugan has also been present at many of the frequent meetings with colleagues, whether at CERN or at SFU, even at particularly bad hours of the day on the West Coast. The work presented in this thesis would not have reached fruition without Dugan's guidance and patience.

I also want to thank past and present members of the HEP group at SFU and the numerous ATLAS colleagues with whom I have collaborated. Graduate studies can often be a lonely pursuit, but I have been fortunate to be embedded in a large community, diverse in wisdom and expertise. In particular, I would like to thank Quentin Buat for his significant help in converging upon the final results presented in this thesis.

The financial support offered by the Department of Physics, the Dean of Graduate Studies Office, the HEP group, and the Natural Sciences and Engineering Research Council of Canada has been essential in making this research possible and providing the privilege to travel to international conferences and meetings, all while keeping food on the table and a roof over my head.

Finally, I am grateful for Amy and my family for their love and support, especially in times of great stress. Thank you for bearing with me these past six years.

Contents

Chapter 1

Introduction

In 2012, the ATLAS and CMS collaborations announced the discovery [1, 2] of a new boson consistent with the Standard Model Higgs boson with a mass of approximately 125 GeV. This observation was mainly supported by searches in the $\gamma\gamma$ and *ZZ* decay modes, due to their superior mass resolutions. Further studies have provided a more precise measurement of the mass and have supported the spin-0 pure scalar hypothesis. These properties are consistent with the long sought-after Higgs boson. Although the $\gamma\gamma$ mode and gluon-gluon fusion Higgs production assume a Higgs coupling to fermions through a one-loop triangle of predominantly top quarks, this discovery did not provide *direct* evidence that the Higgs boson couples to fermions. The fermionic coupling can be directly established by observing Higgs bosons decaying into two fermions.

This thesis presents a search for Standard Model Higgs boson decays in the di-tau channel where both tau decays are hadronic. The search is performed using 20.3 fb^{-1} of proton-proton collision data with a center-of-mass energy of 8 TeV collected in 2012 by the ATLAS detector at the Large Hadron Collider (LHC) at CERN. The results of this search are then combined with parallel searches in the two remaining di-tau channels, where one or both taus decay leptonically.

This introductory chapter will first briefly cover the theoretical framework of electroweak symmetry breaking and the Higgs mechanism, how the Standard Model predicts Higgs bosons can be produced at the LHC and how they should decay. An outline of the analysis strategy employed by this search is then presented.

1.1 Theoretical Motivation

The Standard Model (SM) of particle physics is a successful, although incomplete, model of the known particles and their interactions. In the current model shown in Figure 1.1, twelve fermions with spin 1/2, composed of six quarks and six leptons, are arranged in a three-generation structure. Each generation contains a pair of quarks and a pair of leptons. Higher generations are copies of lower generations with the same quantum numbers but larger masses and each fermion has an anti-fermion with the same mass but opposite electric charge[1]. There is currently no explanation for the three-generation structure or the large variation in fermion masses. Bosons mediate the interactions between the fermions. The massless photon mediates the electromagnetic force between particles with electric charge. The massive W and Z bosons mediate the weak interaction. Together, the electromagnetic and weak interactions are unified under the electroweak interaction. Gluons mediate the strong force between the quarks. Bound states of multiple quarks are called hadrons.

The SM builds upon the mathematical framework of a quantum field theory where particles are described by fields spreading through space and time and where particle dynamics and kinematics are encoded in a Lagrangian density, L_{SM}. The conservation of physical quantities, such as electric charge, arise from certain symmetries imposed on L_{SM}, as declared by Noether's fundamental theorem that associates conservation laws with symmetries of physical systems. A Lagrangian symmetric under *global* gauge transformations only describes free non-interacting particles. Imposing the stronger requirement of invariance under a *local* gauge transformation requires modifications to the Lagrangian that produces new terms describing the interactions of the SM. The symmetries of the SM are described by a unitary group of local gauge transformations:

$$SU(3)_C \times SU(2)_L \times U(1)_Y$$

$SU(3)_C$ governs the strong interaction between quarks, the only fermions with colour charge C, mediated by eight massless gluons. $SU(2)_L \times U(1)_Y$ describes the unified electroweak interaction mediated by the massless photon and the massive weak bosons $W^{\pm}$ and Z.

[1]The neutrino masses are currently unknown and neutrinos are electrically neutral.

L denotes a weak interaction that only takes place between left-handed particles (or right-handed antiparticles) and Y is the weak hypercharge.

1.1.1 Electroweak Symmetry Breaking

Although the weak bosons are known to have mass, there are no mass terms present in the Lagrangian of the electroweak sector; adding Dirac mass terms of the form $-m\bar{\phi}\phi$ does not maintain invariance under the electroweak $SU(2)_L \times U(1)_Y$ local gauge transformations. Particles are instead hypothesized to *acquire* mass after the electroweak symmetry has been spontaneously broken. The symmetry is broken by an energy potential of a new quantized field ϕ where the vacuum state is non-zero and non-unique. The asymmetry is then due to the arbitrary choice of one of the degenerate vacuum states. Requiring the vacuum states to be invariant under Lorentz transformations and under translations implies that this new field must be a scalar field with a constant non-zero vacuum expectation value. This new scalar field is the Higgs field and is an $SU(2)$ doublet with four degrees of freedom.

The simplest such potential that breaks the electroweak symmetry is one of the form:

$$V(\phi) = -\mu^2\phi^\dagger\phi + \frac{\lambda}{2}(\phi^\dagger\phi)^2$$

where ϕ is the complex scalar Higgs field. For $\mu^2 > 0$ and $\lambda > 0$ the potential energy $V(\phi)$ is bounded from below, has a local maximum at $\phi = 0$, and a circle of absolute minima at:

$$\nu = \left(\frac{\mu^2}{\lambda}\right)^{1/2} e^{i\theta}$$

The ground state can be selected where $\theta = 0$:

$$\phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu \end{pmatrix}$$

and any other ground state is related only by a global phase transformation. Expanding the Higgs field around this ground state gives the form:

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} \eta_1 + i\eta_2 \\ \nu + \sigma + i\eta_3 \end{pmatrix}$$

The electroweak Lagrangian with the Higgs potential then yields new mass terms. The three degrees of freedom in η_1, η_2, and η_3 become the longitudinal polarizations of the now massive $W^\pm$ and Z bosons while the photon remains massless. The remaining degree

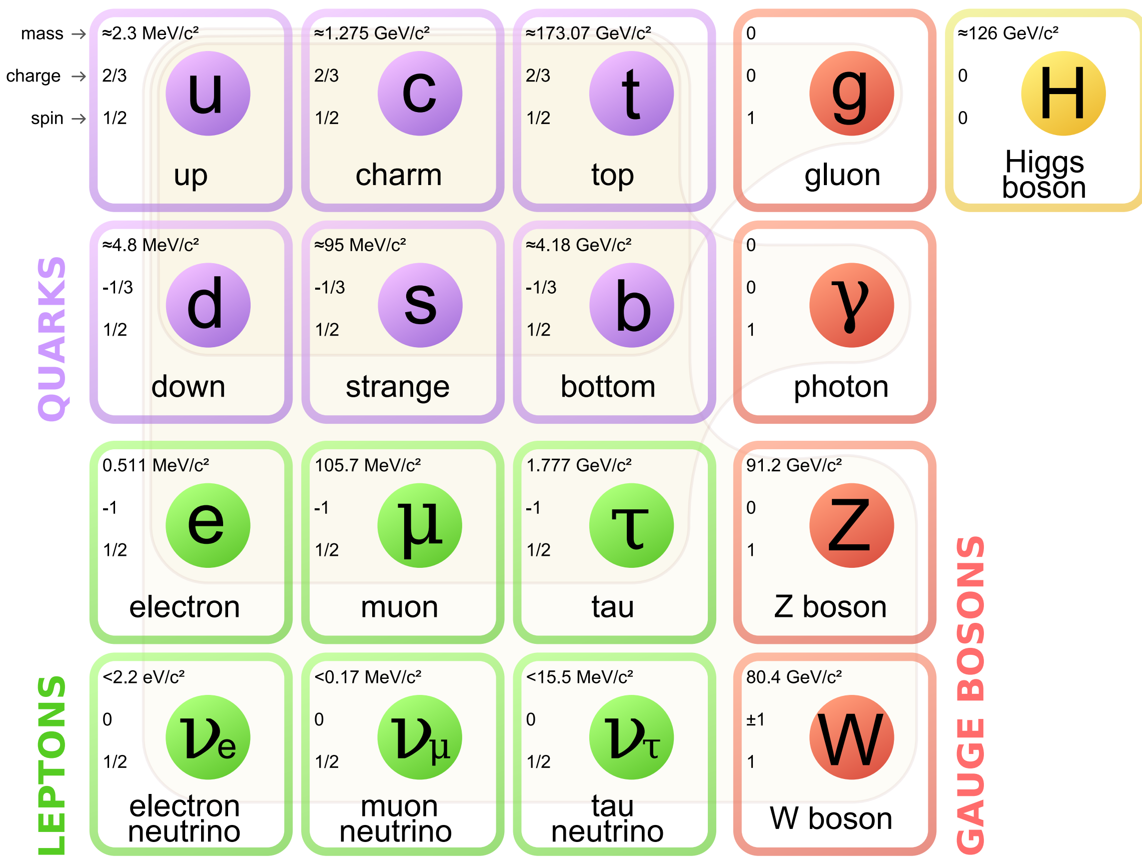


Figure 1.1: The Standard Model of elementary particles and their interactions: the 12 fundamental fermions (including the leptons and quarks) and 4 fundamental bosons [3]. Brown loops indicate which bosons (red) couple to which fermions (purple and green). The mass, charge, and spin of each particle is shown. The Higgs boson couples to all particles with mass.

of freedom in σ predicts the existence of a new electrically neutral scalar Higgs boson. Fermion masses can be generated by the manual inclusion of Yukawa interaction terms of the form $-g_{Hf\bar{f}}\bar{\psi}_f\phi\psi_f$.

1.1.2 Higgs Production at the LHC

The strength of the coupling between the Higgs boson and gauge bosons, fermions and itself is summarized in Table 1.1. A key feature of the Higgs coupling is that it scales with the fermion mass or the square of the gauge boson mass. Higgs production or decays will therefore be dominated by interactions with the heavy W and Z bosons or top and bottom quarks and tau leptons.

At the LHC, Higgs production is dominated by processes that involve interactions of gluons or light quarks in the proton-proton collisions. As the collision energy is increased, the probability of generating loops of heavy quarks or virtual interactions of the heavy weak gauge bosons will increase. An excitation of the Higgs field, in the form of the Higgs boson, will then couple more strongly with these intermediate heavy particles. Four main mechanisms contribute to the Higgs production at the LHC and include gluon fusion (ggF), vector boson fusion (VBF), Higgs-strahlung, and top-Higgs associated production. The leading order diagrams and predicted cross sections for a Higgs mass of 125 GeV computed up to higher orders of perturbation are presented in Table 1.2. The cross sections of the four production mechanisms are also shown in Figure 1.2 as a function of the Higgs mass. Theoretical uncertainties on the predicted cross sections include a component to account for missing higher orders in the QCD calculations and a component representing the underlying uncertainty from the choice of parton distribution functions (PDFs) [4]. PDFs describe the probability of finding a parton (quark or gluon) with a certain longitudinal momentum fraction of the proton and they depend on the energy of the proton-proton interaction. The choice of PDFs then predicts the probability of each type of parton-parton interaction, gluon-gluon interactions being the most frequent at the LHC.

The dominant Higgs production mechanisms leading to the sensitivity of this search are the gluon fusion and vector boson fusion processes. These production mechanisms complement each other by providing handles on both the Higgs Yukawa coupling, in the case of gluon fusion, and the gauge coupling, in the case of vector boson fusion. This analysis will have the ability to test each of these couplings separately and assess compatibility with the Standard Model. The Higgs-strahlung process is also expected to contribute in a visible way and is also accounted for in this analysis. These three production mechanisms are discussed further in the sections below. The top-Higgs associated production is expected to produce an interesting and complex signature in the detector, but the cross section is too low to have a visible impact on the results of this search using the data collected thus far by the LHC. For this reason, the top-Higgs associated production is not accounted for by this analysis.

Gluon Fusion

Gluon fusion is the main Higgs production mechanism at a hadron collider and alone has provided the discovery potential of Higgs searches in the $\gamma\gamma$, ZZ, and WW decay channels over a wide range of Higgs masses. The leading order diagram, shown in Table 1.2, involves a top quark running in the loop created by the interaction of two energetic gluons. Bottom quarks and lighter quarks also contribute to the total cross section in decreasing amounts. This process is sensitive to potential new heavy states in the loop that could enhance the measured gluon fusion cross section, making it an interesting probe into new physics.

Next-to-leading order (NLO) QCD contributions to the leading order diagram in Table 1.2 increase the cross section by 80-100%, which is then further enhanced by about 25% when including next-to-next-to-leading order (NNLO) corrections [5]. The large contribution from higher order processes translates into a frequent radiation of additional gluons in the final state that manifest as additional jets in the ATLAS detector. This energetic radiation will also impart a significant boost to the momentum of the Higgs boson. Unfortunately, this production mechanism is very difficult to distinguish from QCD-initiated $Z \to \tau\tau$ which is also accompanied by a similar spectrum of additional radiation.

Vector Boson Fusion

Although vector boson fusion (VBF) has a much smaller cross section, it has certain distinguishing features that provide the potential for good background suppression. In the leading order t-channel diagram in Table 1.2 two quarks radiate *W* or *Z* bosons that annihilate, creating a Higgs boson. VBF also has contributions from u- and s-channel diagrams [5]. The two quarks in the final state hadronize into jets in the forward regions of the ATLAS detector, and additional radiation between these quarks is suppressed because the quarks are not connected by colour fields. The electroweak production of $Z \to \tau\tau$ is a very similar process but the cross section is negligible. A better theoretical understanding of the VBF process and the reduced contribution from higher-order processes also leads to smaller theoretical uncertainties on its cross section.

Higgs-strahlung

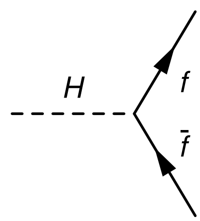
The Higgs-strahlung process involves the radiation of a Higgs boson from *W* or *Z* boson produced by a quark-quark initial state. Leptonic decays of the associated weak boson provide clean detector signatures that can be used to trigger on this process, however, the cross section is much lower than the more significant gluon fusion and vector boson fusion processes. This process is only accounted for to avoid overestimating the gluon fusion or vector boson fusion cross sections since Higgs-strahlung is expected to contribute a non-negligible signal yield.

1.1.3 Higgs Decays

The Higgs boson can decay into a variety of final states, where the branching ratio of each is determined by the Higgs mass. The branching ratios of significant decay modes are shown in Figure 1.3 as a function of the Higgs mass. Figure 1.3 also shows the product of the Higgs production cross section and branching ratios for significant final states that are distinguishable from the backgrounds since this determines the overall rate and relative importance of each channel.

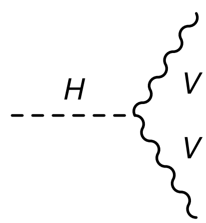
The *WW* and *ZZ* modes dominate at large Higgs masses but become suppressed at low masses as only one of the bosons can be produced on-shell. At low Higgs masses, the

Yukawa Coupling

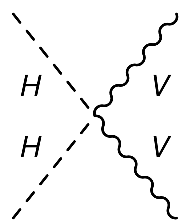


$g_{Hf\bar{f}} = \frac{m_f}{\nu}$

Gauge Coupling

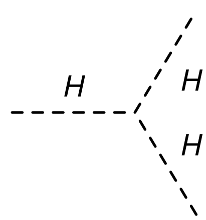


$g_{HVV} = \frac{2m_V^2}{\nu}$

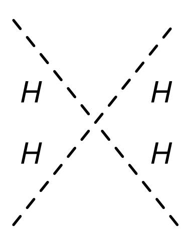


$g_{HHVV} = \frac{2m_V^2}{\nu^2}$

Self-coupling



$g_{HHH} = \frac{3m_H^2}{\nu}$



$g_{HHHH} = \frac{3m_H^2}{\nu^2}$

Table 1.1: SM Higgs boson couplings to gauge bosons ($V = W^{\pm}$ or Z), Higgs bosons and fermions [6]. ν is the vacuum expectation value of the Higgs field, m_f is the fermion mass, m_v is the mass of the vector boson (W or Z), and m_H is the mass of the Higgs boson.

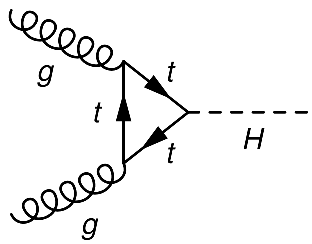
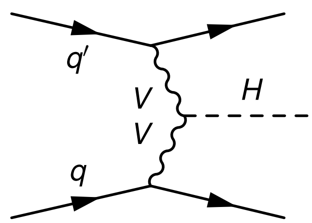
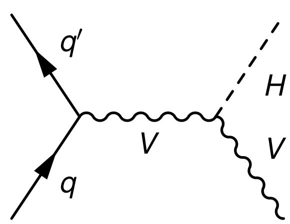
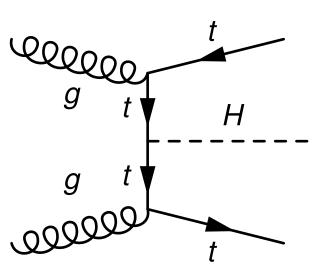
Production Mechanism	Leading Order Diagram	Cross Section [pb]
Gluon Fusion **ggF**		$19.27^{+14\%}_{-14\%}$
Vector Boson Fusion **VBF**		$1.578^{+2.8\%}_{-3.0\%}$
Higgs-strahlung **VH**		WH: $0.7046^{+3.3\%}_{-3.3\%}$ ZH: $0.4153^{+5.6\%}_{5.6\%}$
Top-Higgs **Associated Production** **ttH**		$0.129^{+11.9\%}_{-17.4\%}$

Table 1.2: Standard Model Higgs production mechanisms at the LHC and the corresponding predicted cross sections at a Higgs boson mass of 125 GeV for a center-of-mass proton-proton collision energy of 8 TeV [4].

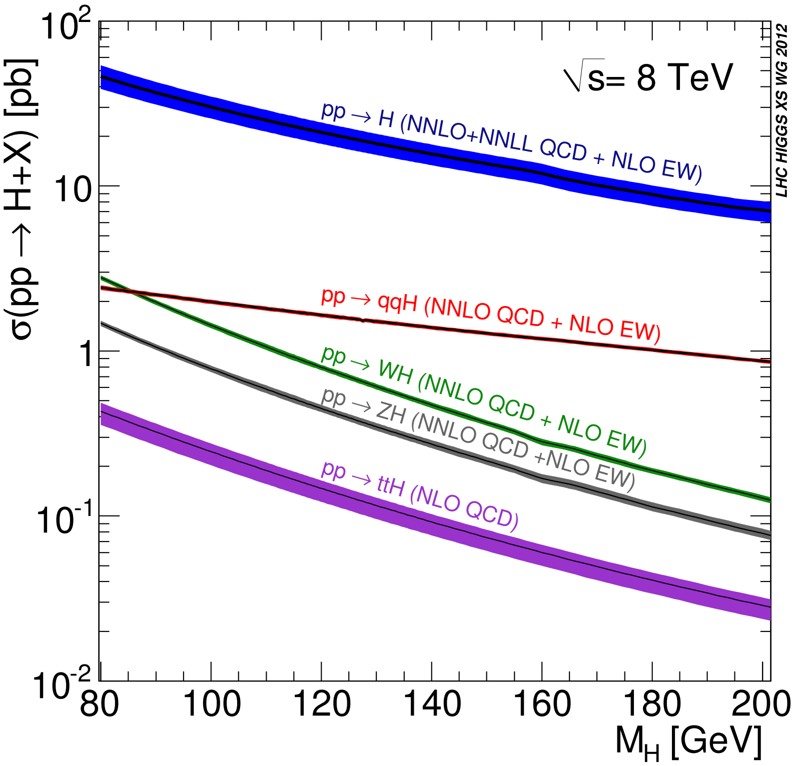


Figure 1.2: Standard Model Higgs production cross sections for a center-of-mass proton-proton collision energy of 8 TeV [4].

$b\bar{b}$ mode dominates as bottom quarks are the heaviest quarks that can be pair-produced on-shell. Extracting the $b\bar{b}$ mode is made difficult by the need to identify jets as originating from bottom quarks ("b-tagging") and the overwhelming direct production of $b\bar{b}$ in QCD background processes. The $c\bar{c}$ mode is similarly difficult to distinguish from the QCD background. The gg mode is essentially the gluon fusion production mechanism in reverse and operates through an intermediate heavy quark loop. gg is of course impossible to separate from QCD background processes. The branching ratio to $\gamma\gamma$ is small due to the required intermediate loop of a W boson or heavy quark since the Higgs boson does not directly couple to photons. Despite the small branching ratio, $\gamma\gamma$ provides a clean signature with a very good invariant mass resolution. The branching ratio to $\mu\mu$ is also small due to the small muon mass, but this mode could also provide a good invariant mass resolution.

The $\tau\tau$ decay mode is the golden channel for probing the fermionic coupling of the Higgs boson at low Higgs masses. This fact motivates the search in the $\tau\tau$ channel documented in this thesis.

1.1.4 Tau Decays

Tau decays and their reconstruction in the ATLAS detector are described in greater detail in Section 3.5. The tau is the only lepton massive enough to decay into hadrons (mostly charged and neutral pions) and it does so in 65% of all decays. Aside from the large hadronic branching fraction, each hadronic tau decay only produces a single neutrino (pictured in Figure 1.3), providing a better resolution on the di-tau mass since neutrinos carry away energy that cannot be directly measured by the ATLAS detector.

As the tau lepton can decay both leptonically and hadronically, three analyses are designed in parallel to separately search for $H \to \tau\tau$ decays in the fully leptonic, semi-leptonic, and fully hadronic final states. The fully leptonic and semi-leptonic channels benefit from the use of electron and muon triggers that operate at a higher efficiency than the hadronic tau triggers, but suffer from additional neutrinos in the final state, worsening the di-tau mass resolution. These final states also contain a larger contribution from a variety of leptonic backgrounds while these backgrounds are minimal in the fully hadronic final state. This thesis presents the analysis of the fully hadronic channel. The three analyses are then combined and the results are shown in Chapter 7.

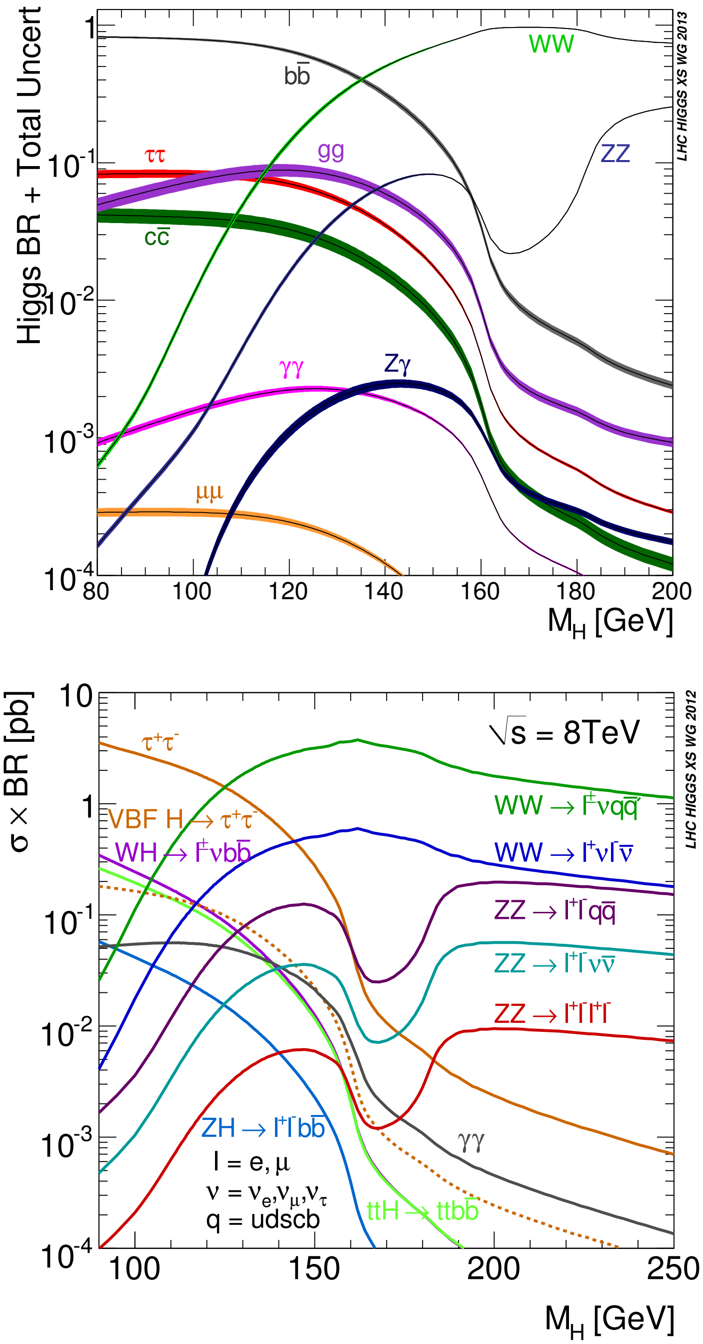


Figure 1.3: Top: Higgs branching ratios and their uncertainties for the low mass range [4]. Bottom: Higgs production cross sections times branching ratios for important search channels at the LHC for a center-of-mass proton-proton collision energy of 8 TeV [5].

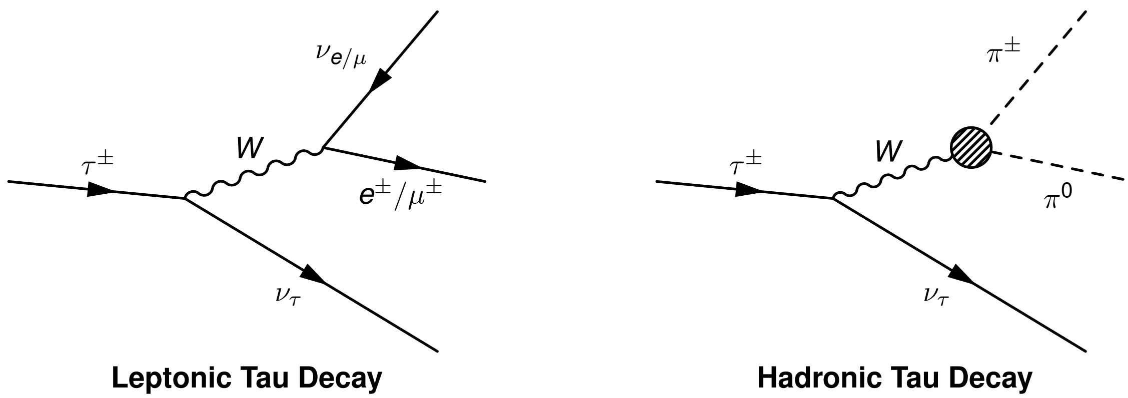


Figure 1.3: Left: A leptonic tau decay. Right: An example of a hadronic tau decay where the hatched blob represents possible intermediate resonances of bound quark pairs. The final state of the hadronic decay includes an odd number of charged mesons (pions, kaons, etc.) and multiple neutral mesons. The majority of hadronic decays produce charged and neutral pions.

1.2 Analysis Strategy

The strategy developed in this analysis of the fully hadronic di-tau channel depends on the form of the signal and background processes considered and the pursued goals. The LHC has not yet delivered enough data to begin making precision Higgs measurements in the di-tau channel, nor does the di-tau channel provide a resolution on a Higgs mass measurement that is competitive with the $\gamma\gamma$ and ZZ channels. The main goal of this search is to determine if the already observed Higgs boson with a mass of about 125 GeV also couples to taus.

At the core of the analysis strategy is the categorization of events by cuts on event topologies and kinematics that separately exploit the signatures of the ggF and VBF processes. A VBF-enhanced category is designed to optimize the sensitivity to the VBF process while reducing the contamination from ggF signal events that have larger theoretical uncertainties. A second category is constructed using the remaining events where the di-tau system is highly boosted. This second category is expected to select gluon fusion events where the Higgs boson has recoiled off of additional radiation. The two major sources of background events in the fully hadronic channel are QCD multijet processes and $Z \to \tau\tau$. QCD multijet processes contribute to both categories but the relative impact on the boosted category is lessened by the requirement on the boost of the di-tau system. $Z \to \tau\tau$ is a prominent background in both categories and is difficult to distinguish from the Higgs signal.

While the event categorization offers some background rejection and sets the stage for probing the two main Higgs production mechanisms, the main challenge confronting this search is the need to further suppress the background contributions within each category. Even before the events are selected by this analysis they must be triggered on by the ATLAS detector, and the fully hadronic di-tau channel presents significant challenges here. QCD jets can also easily fake hadronic tau decays, making tau identification another essential component of this search. Multivariate analysis techniques finally bring the background suppression to a level where the search is sensitive to the presence of $H \to \tau\tau$ decays. Boosted decision trees are trained to separate the 125 GeV Higgs signal from the background using a variety of discriminating features.

1.3 Thesis Outline

The content of this thesis is presented in a series of chapters that progressively build up to the results of this search. Chapter 2 details the operation of the LHC and the capabilities of the ATLAS detector. The reconstruction of important physics objects is then presented in Chapter 3. The design of the analysis is documented in a series of three chapters. Chapter 4 outlines the event selection, Chapter 5 describes the background and signal models, and Chapter 6 presents the analysis techniques used to extract a Higgs signal and quantify the significance of an excess. The results of the search are presented in Chapter 7 followed by final remarks in Chapter 8.

Chapter 2

Experimental Apparatus

Along the French-Swiss border near the city of Geneva lies the world's leading particle physics laboratory, home of the largest and most complex scientific instruments and largest international collaborations of scientists and engineers. The European Laboratory for Particle Physics operated by CERN[1], the European Organization for Nuclear Research, consists of an extensive network of accelerators, particle experiments, and world-class computing facilities. The Large Hadron Collider (LHC[2]) at CERN, is the world's largest and most powerful particle accelerator, colliding protons or lead ions at the sites of four particle detectors situated along its circumference. Opposite each other along the LHC synchrotron ring shown in Figure 2.1, ATLAS (A Toroidal LHC Apparatus) and CMS (the Compact Muon Solenoid), are general purpose high luminosity particle detectors designed to search for the Higgs boson and probe a wide range of new physics beyond the Standard Model.

This chapter outlines the design and operation of the LHC in Section 2.1, the ATLAS detector and its subsystems in Section 2.2, and finally the ATLAS trigger and data acquisition system in Section 4.2.

[1]CERN is derived from the acronym for the original French name *Conseil Européen pour la Recherche Nucléaire*, a provisional body founded in 1952 with the mandate of establishing a world-class fundamental physics research organization in Europe.

[2]Not to be confused with *Les Horribles Cernettes*.

2.1 The Large Hadron Collider

The LHC [7] synchrotron sits in the same 26.7 km circular tunnel from 45 m to 170 m underground that previously contained the Large Electron-Positron (LEP) collider. Construction of the LHC began in 1998 while LEP was still in operation. In 2001 LEP was finally dismantled to make way for the LHC, amid pressure to continue running as the LEP collision data began to show hints of a Higgs-like particle with a mass of about 115 GeV [8]. On September 10, 2008, the first beam of protons was successfully steered around the full LHC ring.

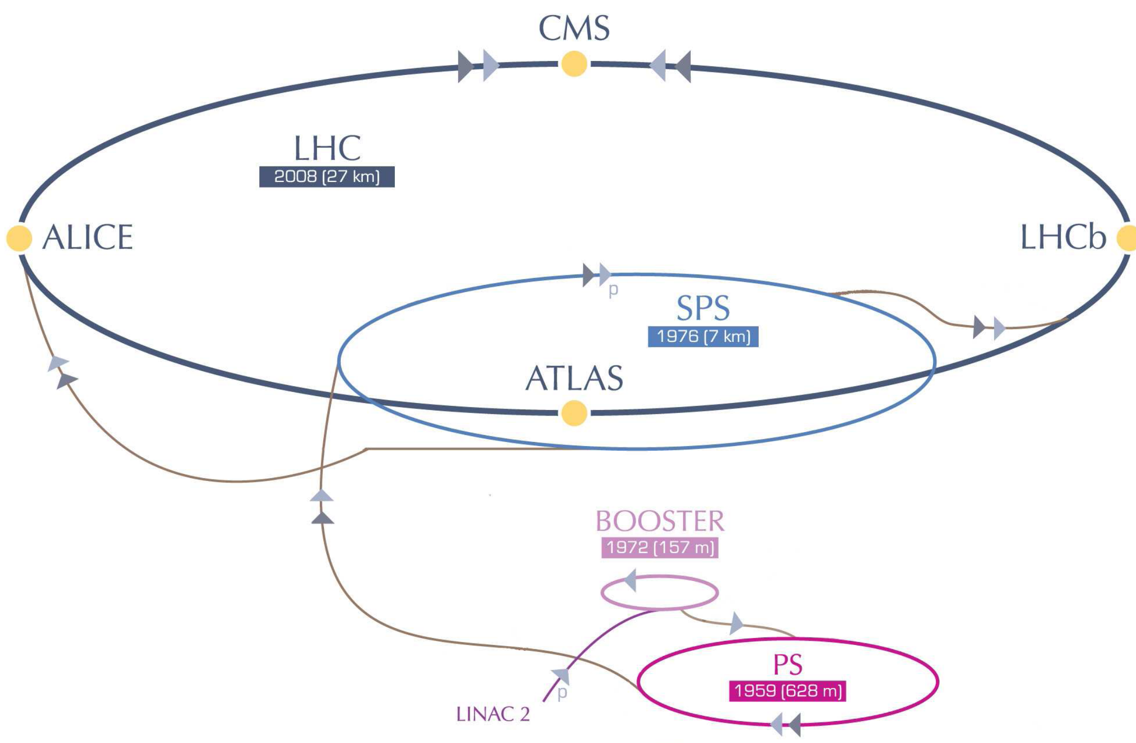


Figure 2.1: The LHC accelerator complex [9]. Before injection into the LHC, the proton beams are accelerated by the Proton Synchrotron Booster, the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS).

The LHC consists of eight straight sections and eight arcs made up of 1232 dipole magnets for bending the beam, 392 quadrupole magnets for focussing the beam, and additional complex magnet systems for beam corrections and squeezing the beams at the collision points. Unlike a particle-antiparticle collider, the counter-rotating beams must be circulated in two separate rings. Due to limited space in the LHC tunnel, a twin-bore magnet design was developed, consisting of two sets of coils and beam channels within the same magnetic and mechanical structure and cryostat. At the design energy of 7 TeV per proton beam, 11850 A of electrical current is required to create an 8.33 T magnetic field in the superconducting dipole magnets. The superconducting magnets are cooled to 1.9 K with super-fluid helium and the beam pipes are maintained under vacuum conditions with pressures below 10^{-13} atmospheres.

Proton beams are accelerated and injected into the main LHC ring in stages as shown in Figure 2.1. Hydrogen gas is first ionized by an electric field stripping away the electrons. The protons are then accelerated to an energy of 50 MeV by the LINAC2 linear accelerator. The Proton Synchrotron Booster forms proton bunches and accelerates them to 1.4 GeV. The bunched proton beams are fed into the Proton Synchrotron accelerating them to 25 GeV, followed by an acceleration to 450 GeV in the Super Proton Synchrotron. Finally, two counter-rotating beams of proton bunches are injected into the LHC where they are accelerated to an energy of up to 7 TeV. Acceleration and the creation of proton bunches is accomplished by the oscillating electromagnetic field inside radiofrequency (RF) cavities. The LHC has 16 RF cavities (8 per beam) operating at 400 MHz in a superconducting state. Protons that are perfectly synchronized with the oscillating RF cavities will see no acceleration while protons with slightly different energies will be accelerated or decelerated until they are synchronized. The process also forms and maintains bunches of protons. The LHC is designed to handle 2808 proton bunches with a 25 ns bunch-spacing.

The number of collision events per second as bunches are crossed at the center of the ATLAS detector is given by:

$$N_{\text{event}} = L\sigma_{\text{process}} \tag{2.1}$$

where σ_{process} is the cross section of the particular process being studied and L is the luminosity. The luminosity depends only on the LHC beam parameters and for a Gaussian beam profile can be written as:

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi\epsilon_n\beta^*} F \tag{2.2}$$

where N_b is the number of protons per bunch, n_b the number of bunches per beam, f_{rev} the revolution frequency, γ_r the relativistic factor, ϵ_n the normalized transverse beam emittance,

β^* a measure of the beam focus at the collision point, and F a geometric reduction factor due to the crossing angle of the beams at the interaction point [7]. Three primary ways of increasing the luminosity include squeezing the beams down to a smaller transverse size, increasing the number of circulating bunches, or increasing the number of protons in each bunch. Squeezing the beams or increasing the number of protons per bunch also leads to an increase in the number of proton-proton interactions within the same bunch, called *in-time pile-up*, while increasing the number of bunches (and therefore decreasing the space between them) leads to an increase in the *out-of-time pile-up*. Out-of-time pile-up occurs when interactions from subsequent bunch-crossings occur during the time required by the detector to process a single event. The LHC and high-luminosity ATLAS and CMS detectors are designed to operate at a peak luminosity of $L = 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ with proton beams, corresponding to a bunch collision rate of 40 MHz and an average of 22 simultaneous proton-proton interactions.

Nine days after the first circulation of proton beams, on September 19, 2008, a major incident occurred in which a poor electrical connection between two dipole magnets caused an electric arc with enough energy to locally vaporize the beam pipes and to destroy the cryogenic envelope. The liquid helium heated and escaped with explosive force, displacing 26 magnets [10]. Over the following year, 53 magnets were replaced and more than 1 km of contaminated beam vacuum tubes required cleaning. LHC commissioning restarted in November, 2009, soon surpassing the Tevatron to become the world's highest-energy and highest-luminosity hadron-hadron collider. Following the 2008 incident, however, it was decided to operate the LHC at below the design targets. The LHC delivered 5.46 fb^{-1} of collisions with a center-of-mass energy of 7 TeV in 2011 and 22.8 fb^{-1} at 8 TeV in 2012. Figure 2.2 shows the cumulative luminosity and the distributions of the mean number of interactions per bunch crossing for the 7 and 8 TeV runs.

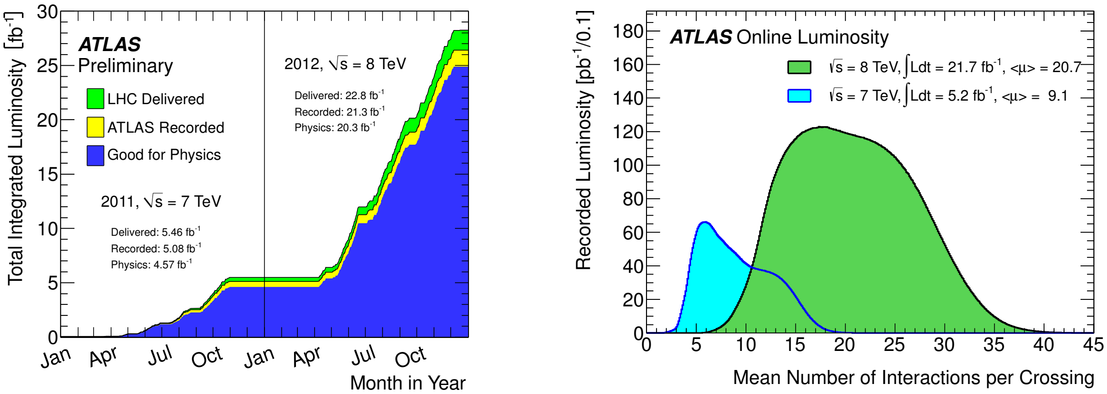


Figure 2.2: Left: Cumulative luminosity versus time delivered to (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams and for pp collisions at 7 and 8 TeV centre-of-mass energy in 2011 and 2012 [11]. Right: The luminosity-weighted distribution of the mean number of interactions per bunch-crossing for the 2011 and 2012 data [11].

2.2 The ATLAS Detector

ATLAS is a multi-purpose detector, 25 m in height and 44 m in length, installed in a cavern along the LHC at CERN's main site in Meyrin, Switzerland. ATLAS is composed of several layers arranged in a cylindrical geometry that is nominally symmetric about the beam collision point. Each layer is designed to measure signals left by certain particles and collectively they build a detailed description of each collision event. A cut-away view of the ATLAS detector is shown in Figure 2.3. Beginning with the components closest to the beam, the Inner Detector is made up of silicon and gas-based particle tracking devices, the calorimeters surround the tracking system and measure energy deposits from neutral and charged particles, and finally the Muon Spectrometer forms the outermost layers recording muon tracks. The central barrel region is enclosed by two end-caps of additional muon and calorimetry systems. A solenoid magnet surrounds the inner detector for the momentum measurement of tracks left by charged particles. A large toroid magnet spanning the barrel region and two smaller toroids in the end-caps provide the magnetic field for muon momentum measurement.

The high-luminosity high-pileup LHC proton-proton collision environment presents certain challenges in the design of the ATLAS detector. In general, the physics goals of the ATLAS detector require fast, radiation-hard electronics and sensors with fine granularity and a large acceptance in solid angle around the collision point. Good energy resolution and fine granularity in the calorimeters are required for the reconstruction of electrons, photons, taus, and jets. The missing energy reconstruction requires high coverage in the calorimeter and muon systems. Tracking components close to the interaction region allow the reconstruction of secondary vertices for tau identification and the tagging of *b*-jets. Efficient triggering on low-momentum objects with a high background rejection is also important to maintain an acceptable trigger rate. This is especially important for the search presented in this thesis where the ability to trigger on taus with low momenta increases the Higgs signal acceptance substantially. The design of the major ATLAS subsystems will be described in Sections 2.2.2-2.2.4 after introducing the ATLAS coordinate system in the next section.

2.2.1 The ATLAS Coordinate System

The ATLAS detector is described by a right-handed coordinate system with an origin at the geometric center of the detector, coinciding with the nominal interaction point. The z-axis lies tangent to the beam with the positive z-axis pointing in the counter-clockwise direction around the LHC (when viewed from above). The x-y plane is transverse to the beam with the positive x-axis pointing at the center of the LHC ring and the positive y-axis pointing upwards. It is often more convenient to describe detector signals and reconstructed physics objects with polar coordinates. In this case, the polar angle, θ, takes on values from 0 at the positive z-axis to π at the negative z-axis. The azimuthal angle, ϕ, takes on values from $-\pi$ at the negative y-axis to π at the positive y-axis, with 0 coinciding with the positive x-axis. The polar direction of physics objects is frequently measured in terms of *rapidity*, y, defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \tag{2.3}$$

where intervals are Lorentz-invariant under boosts along the z-axis. In the ultra-relativistic limit, rapidity is approximated by the *pseudorapidity*, η, defined as:

$$\eta = -\ln \tan(\theta/2) \tag{2.4}$$

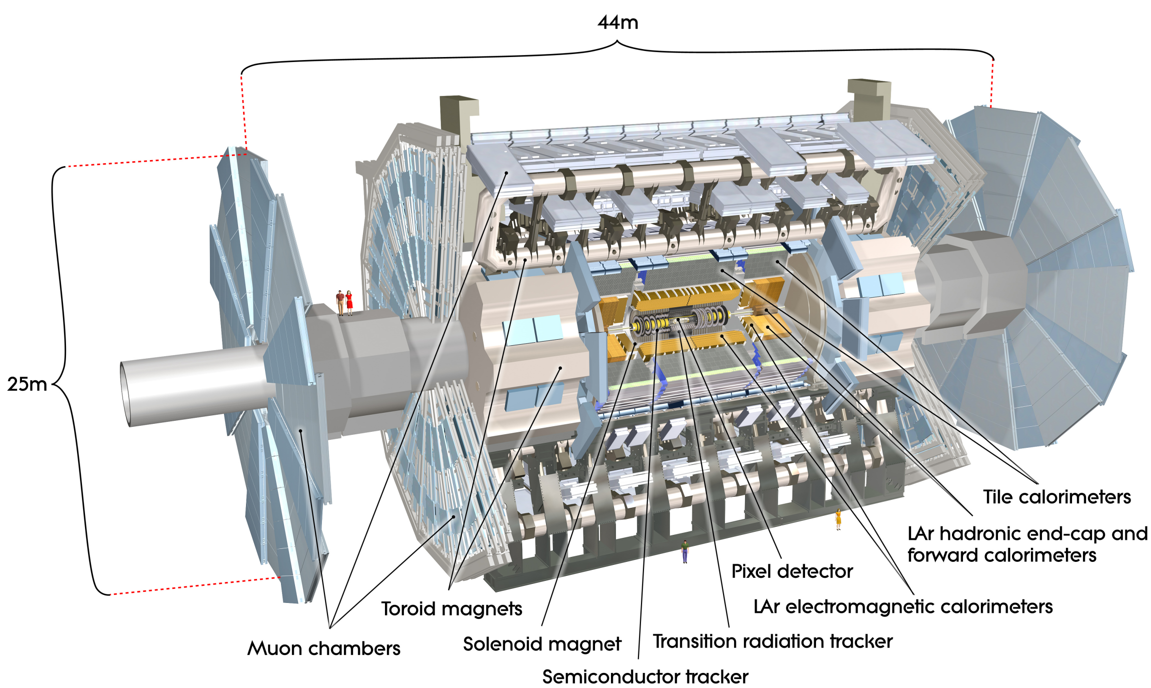


Figure 2.3: A cut-away view of the ATLAS detector, displaying the major subsystems. The detector measures 25 m in height and 44 m in length, weighs approximately 7000 tonnes, and contains about 3000 km of cables [12].

which is equivalent to the rapidity for massless particles. Angular separations in the $\eta - \phi$ plane are measured by the ΔR quantity, defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{2.5}$$

2.2.2 Inner Detector

The Inner Detector (ID) records hits from charged particles within a pseudorapidity range of $|\eta| < 2.5$. The cylindrical system is situated inside a superconducting solenoid with a 2 T magnetic field, allowing a momentum measurement of the curved charged particle tracks as reconstructed by pattern recognition algorithms. The solenoid shares a common cryostat with the liquid argon electromagnetic calorimeter to minimize the amount of dead material in front of the calorimeters. The inner detector is composed of three separate but complementary subsystems: the Pixel Detector, the SemiConductor Tracker (SCT), and the Transition Radiation Tracker (TRT). The barrel modules of the Pixel Detector and SCT are arranged in concentric cylinders around the beam axis and the end-cap modules form disks that are perpendicular to the beam. The granularity of the ID increases inward toward the beam. A cut-away view of the ID displaying the barrel and end-cap components is shown in Figure 2.4 and Figure 2.5 shows the radial scale of each ID layer in a view transverse to the beam axis.

Pixel Detector

The Pixel Detector has the finest granularity, sits closest to the beam, and has full coverage in azimuthal angle and provides coverage within a pseudorapidity range of $|\eta| < 2.5$. 1774 identical pixel sensor modules are arranged into three concentric cylindrical layers in the barrel region and three disks in each end-cap. Each sensor contains 47,232 pixels, for a total of more than 80 million readout channels. Sensors have a minimum pixel size in $R - \phi \times z$ of 50×400 μm^2 and an intrinsic accuracy of 10 μm in $R - \phi$ 115 μm along z in the barrel and along R in the end-cap disks. To prevent radiation damage to the pixels and to reduce noise and leakage current, the Pixel Detector and SCT are operated between -5 to $-10°$C.

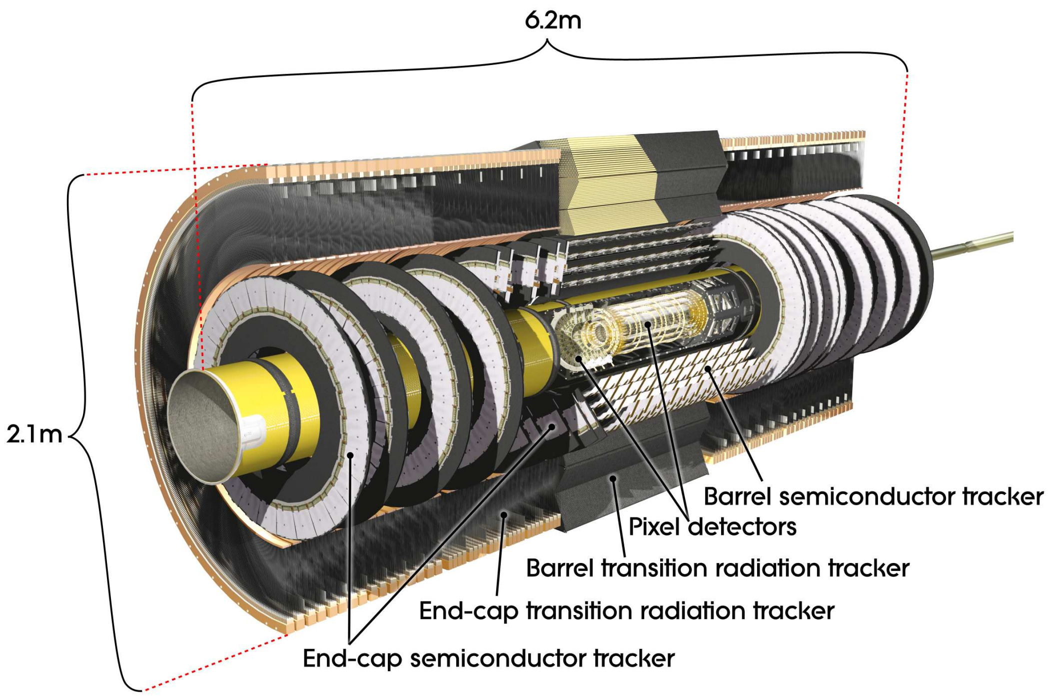


Figure 2.4: A cut-away view of the ATLAS inner detector showing the pixel detector, semiconductor tracker, and the transition radiation tracker [12].

SemiConductor Tracker

The SCT surrounds the Pixel Detector and provides additional precision tracking but with a silicon microstrip sensor technology that is more cost-effective than the pixel sensors. The SCT consists of 4088 modules tiling four cylindrical layers in the barrel region and nine disk layers in each end-cap with a total of approximately 6.3 million readout channels. Like the Pixel Detector, the SCT provides full coverage in azimuthal angle and coverage within a pseudorapidity range of $|\eta| < 2.5$. Each module has silicon strips on both sides that are parallel to the beam axis but the strips have a small stereo angle of 40 mrad between the top and bottom layer such that the coordinate along the module axis can be measured. The intrinsic accuracy of SCT modules is 17 μm in $R - \phi$ and 580 μm along z in the barrel and R in the end-caps.

Transition Radiation Tracker

The TRT surrounds the SCT and is the outermost component of the Inner Detector. The TRT is made up of many straw tubes 4 mm in diameter containing mainly xenon gas. An electric potential of 1530 V is maintained between the conductive tube walls and a gold-plated tungsten wire suspended within the center of each tube. As a charged particle passes through the gas mixture, freed electrons drift toward the anode wire and generate an electrical signal. The TRT also provides discrimination between electrons and pions. Minimum-ionizing charged particles like pions generate a much smaller signal than electrons that emit transition radiation in the form of x-ray photons as they pass through the inhomogeneous TRT material. These photons are absorbed by the xenon gas and result in a larger avalanche of freed electrons toward the anode wire. Straws 144 cm long are packed parallel to the beam axis in the barrel region and straws 37 cm long are arranged radially in the end-caps. The TRT covers the pseudorapidity range of $|\eta| < 2.0$ and only provides a measurement of the $R - \phi$ coordinates in each straw with an intrinsic accuracy of 130 μm, but typically registers 36 hits per track. This much higher number of hits spread over a larger distance scale compensates for the lower precision when compared with the SCT and Pixel Detector. The TRT has a total of approximately 351,000 readout channels and operates at room temperature.

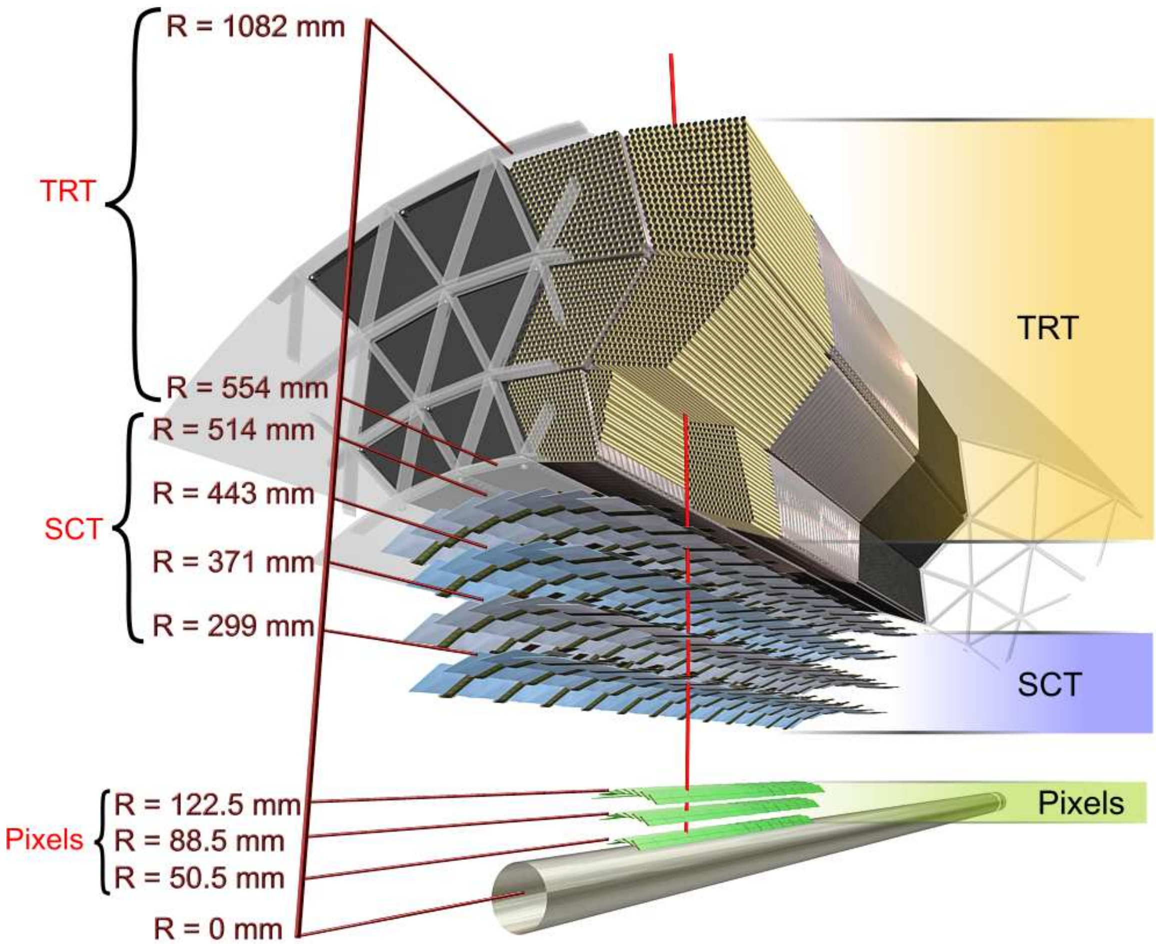


Figure 2.5: Drawing showing the sensors and structural elements traversed by a charged track in the barrel inner detector. The track traverses successively the beryllium beam-pipe, the three cylindrical silicon-pixel layers with individual sensor elements of 50x400 μm^2 , the four cylindrical double layers (one axial and one with a stereo angle of 40 mrad) of barrel silicon-microstrip sensors (SCT) of pitch 80 μm, and approximately 36 axial straws of 4 mm diameter contained in the barrel transition-radiation tracker modules within their support structure [12].

2.2.3 Calorimeters

ATLAS measures particle energies with sampling calorimeters that contain layers of dense material to absorb incident particles and initiate showers of additional particles and layers of active material to measure an output signal that is proportional to the input energy. An electromagnetic calorimeter is designed to absorb and measure the energies of electrons and photons, and a hadronic calorimeter measures energies deposited by hadrons. A cut-away view of all calorimeter systems is shown in Figure 2.6. The calorimetry systems have a combined pseudorapidity coverage of $|\eta| < 4.9$. Containment of the electromagnetic and hadronic showers in the calorimeters is important for accurate energy measurement and the reconstruction of the missing energy, as well as to limit the punch-through into the muon system. The total thickness of the calorimetry (including dead material) is 11 interaction lengths[3] at $\eta = 0$.

Electromagnetic Calorimeters

The electromagnetic calorimeter (ECAL) is a sampling calorimeter using absorber plates made of lead and active regions filled with liquid argon. The lead and liquid argon layers are shaped in an accordion-like geometry, shown in Figure 2.7, such that the calorimeter has full coverage in ϕ without cracks. As high energy electrons or positrons pass through and interact with the absorber material, bremsstrahlung photons are emitted. High energy photons entering the calorimeter or produced via the bremsstrahlung effect generate electron-positron pairs as they interact with the absorber material. In this way, a tree-like shower of photons, electrons, and positrons develops through the layers of the calorimeter until the energies of the shower particles fall below the critical energy required to produce additional particles (approximately 10 MeV), at which point the shower stops. Charged particles passing through the active layers ionize the liquid argon. The liberated charges are collected by electrodes and an electrical impulse signal proportional to the energy deposited in each liquid argon layer is measured.

The ECAL has three layers in the barrel covering $|\eta| < 1.475$ and two layers in the

[3]The nuclear interaction length is the mean distance travelled by a hadronic particle before undergoing an inelastic nuclear interaction.

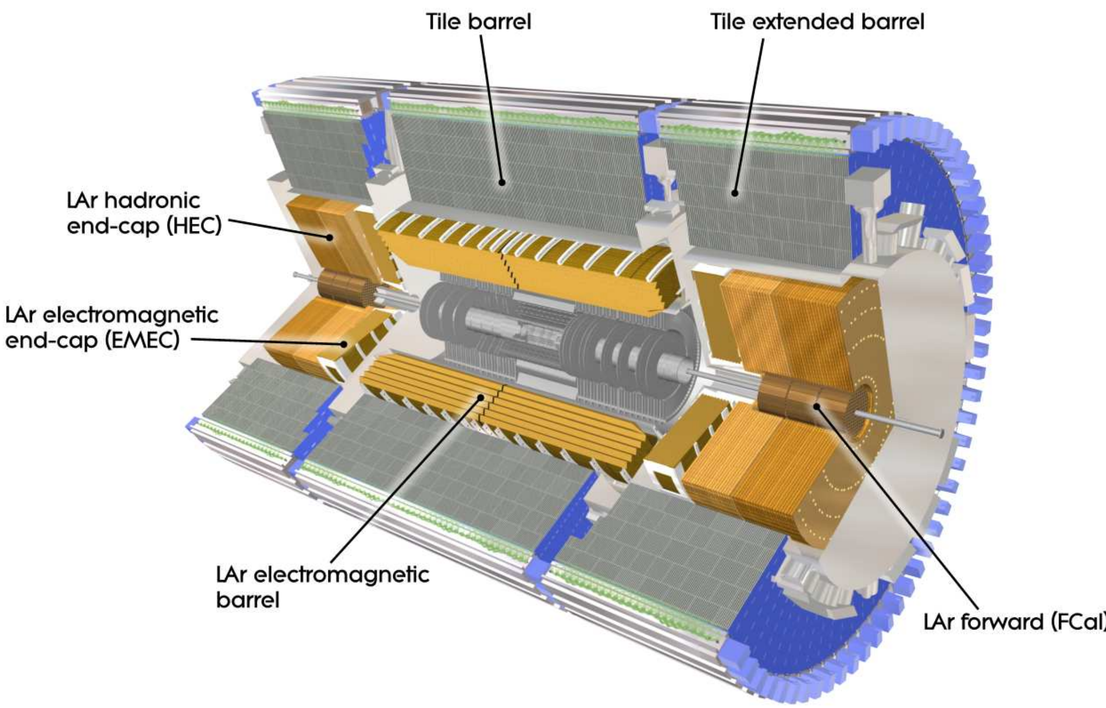


Figure 2.6: A cut-away view of the ATLAS calorimeter system showing the electromagnetic and hadronic calorimeters [12].

end-caps covering $1.375 < |\eta| < 3.2$. The first layer of the ECAL is called the strip layer and has a very fine granularity in η ($0.025/8 \times 0.1$ in $\Delta\eta \times \Delta\phi$) up to $|\eta| < 1.4$ and plays an important role in identifying collimated photons from neutral pion decays. The majority of the calorimeter cells have a granularity ranging from 0.025×0.025 in the inner layers of the barrel region to 0.050×0.025 in the outer layer and 0.1×0.1 in the end-caps. In the region of $|\eta| < 1.8$ in front of the ECAL a presampler detector is used to correct for energy lost by electrons and photons in dead material before entering the ECAL. The total thickness of the ECAL ranges more than 22 radiation lengths[4] in the barrel to 26 radiation lengths in the end-caps.

Hadronic Calorimeters

The hadronic calorimeters (HCAL) are designed to absorb and measure the energies of hadrons, and despite having a coarser granularity than the ECAL, they are sufficient for jet and missing transverse momentum reconstruction. Hadrons such as charged pions initiate showers in the detector absorber material via electromagnetic or inelastic nuclear interactions. More interaction lengths are required to contain the hadronic showers than for the electromagnetic showers in the ECAL. More than 8 additional interaction lengths are offered by the HCAL beyond the 2 interaction lengths in the ECAL.

The HCAL consists of the Tile Calorimeter surrounding the ECAL in the barrel region covering $|\eta| < 1.7$, followed by the Hadronic End-cap Calorimeter (HEC) placed on the outside of the ECAL end-cap layers covering $1.5 < |\eta| < 3.2$, and finally the Forward Calorimeter (FCal) covering $3.1 < |\eta| < 4.9$ dedicated to the measurement of forward jets. The Tile Calorimeter is made of alternate layers of steel absorbers and plastic scintillator tiles. Hadronic shower particles excite the scintillator material which then emits light that is transformed into electrical signals by photo-multiplier tubes. Due to the higher radiation density in the forward regions, the end-cap HCAL detectors use liquid argon as the active material. The HEC consists of two independent wheels per end-cap and uses copper absorbers. The FCal consists of three modules in each end-cap. The first layer of the FCAL uses copper absorbers optimized for electromagnetic interactions and the remaining

[4] A radiation length is the mean distance travelled by a high-energy electron through a material leaving it with $1/e$ of its original energy.

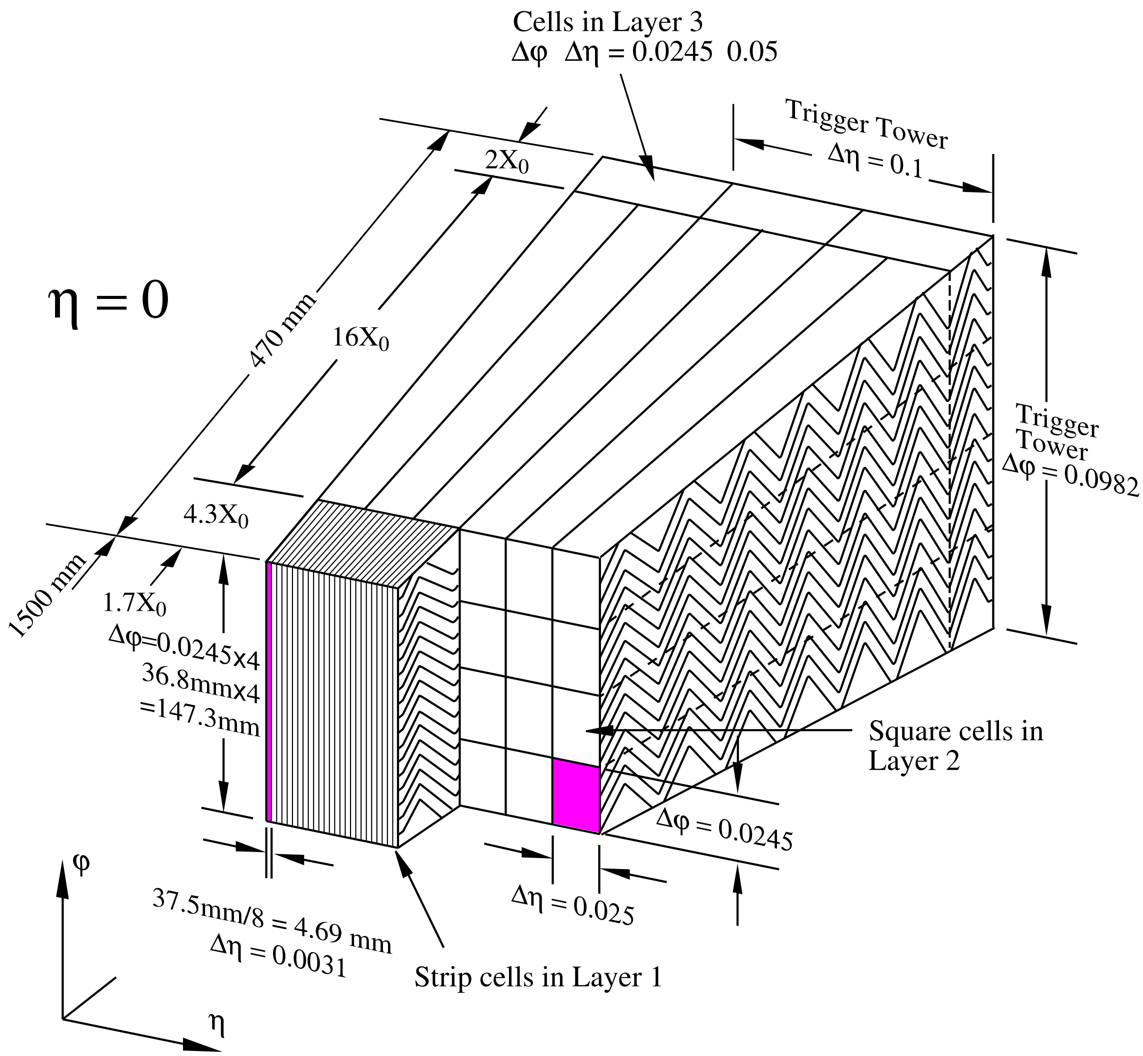


Figure 2.7: Diagram of a barrel module of the liquid argon electromagnetic calorimeter showing the accordion geometry and granularity in η and ϕ of the cells of each of the three layers [12].

two layers use tungsten optimized for hadronic interactions.

2.2.4 Muon Spectrometer

Most muons in the ATLAS detector are minimum ionizing particles, so while they can leave tracks in the inner detector, they typically deposit only a few GeV of energy in the calorimeters. The outermost layers of the ATLAS detector are dedicated to the detection of muons. The Muon Spectrometer (MS), shown in Figure 2.8, covers $|\eta| < 2.7$ and is composed of four types of detectors that provide either precision tracking or fast triggering.

Monitoring drift tubes (MDTs) cover $|\eta| < 2.7$ with modules in both the barrel and end-cap regions. MDTs provide a muon momentum measurement with precision coordinate measurements in the bending direction of the toroidal magnetic field. MDTs are composed of rows of drift tubes in different configurations to allow position measurements in both η and ϕ. Each drift tube contains an anode wire kept at 3080 V and a gas that is ionized when traversed by a muon. Cathode Strip Chambers (CSCs) are multi-wire proportional chambers and handle a higher muon flux in the forward directions. The long charge drift times in the CSCs and MDTs (20 ns in the CSCs but with a large tail and up to 700 ns in the MDTs) make them inadequate for triggering purposes. Resistive Plate Chambers (RPCs) are installed in the barrel region and have an intrinsic time resolution of only 1.5 ns. Two parallel resistive plates are separated by 2 mm. A large electric field between the plates allows avalanches to form around the ionizing muon tacks in the contained gas mixture. Thin Gap Chambers (TGCs) are installed in the forward regions and like the CSCs, are multi-wire proportional chambers, but have a smaller time resolution. TGC signals arrive with 99% probability inside a time window of 25 ns.

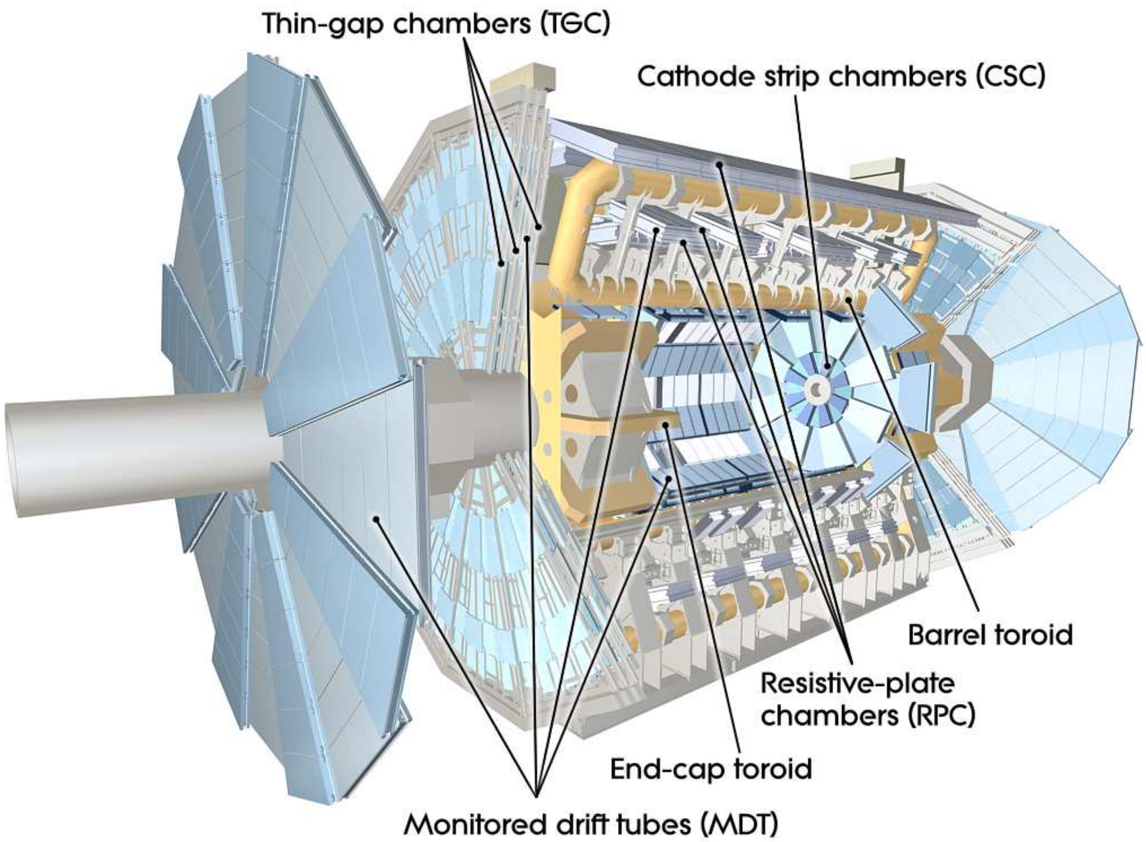


Figure 2.8: A cut-away view of the ATLAS muon system and toroid magnets [12].

2.3 Triggering and Data Acquisition

The LHC delivered collisions at a rate of 20 MHz in 2012, however the ATLAS data acquisition system is designed to record only hundreds of events per second. The vast majority of collisions produce low-p_{T} multijet events and it is the job of the trigger systems to filter events down to a small subset that contain interesting physics at a rate that can be recorded for further analysis. The trigger system has three levels, each progressively reducing the event rate while using an increasing amount of detector information.

Level 1 (L1) is hardware-based and uses the calorimeter information at a reduced granularity, and the fast muon trigger chambers, to make a decision within 2.5 μs. Fast trigger algorithms search for high transverse-momentum muons, electrons, photons, jets, and hadronically decaying tau leptons, defining regions of interest (ROIs) around the identified physics objects. L1 can also trigger on a large imbalance in the transverse momentum or large total transverse energy. L1 reduces the event rate to about 70 kHz. Data from L1 is sent through a readout system where is it passed along to the high-level trigger (HLT).

The HLT is a software-based trigger system consisting of Level 2 (L2) and the Event Filter (EF). L2 further refines the selected events by using the full detector granularity within the ROIs defined by L1 as well as the inner detector. L2 makes a decision within 75 ms and reduces the event rate to 6.5 kHz. Finally, the EF trigger has access to the full detector information and executes identification algorithms very similar to the offline reconstruction. EF takes 4 s on average to make a decision and reduces the event rate to approximately 1 kHz.

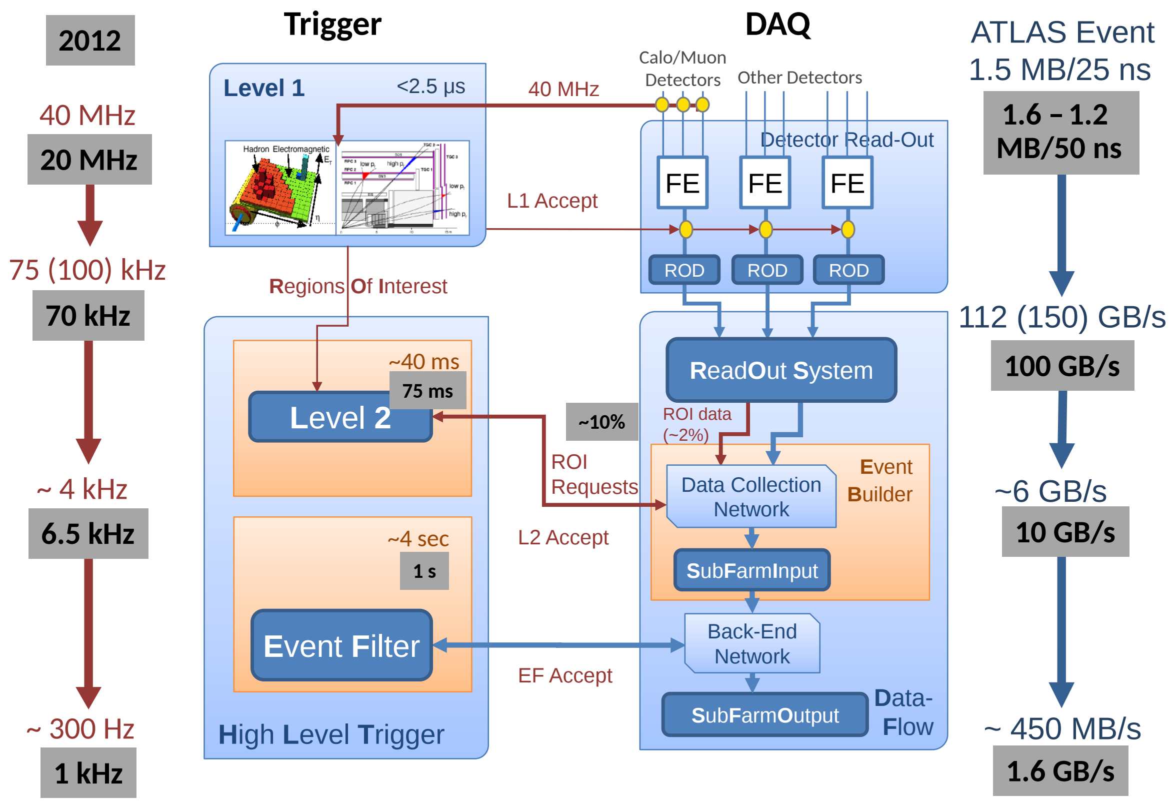


Figure 2.9: The ATLAS trigger and data acquisition system design in 2012 [11].

Chapter 3

Event Reconstruction

This analysis relies on an understanding of the ATLAS event reconstruction from the basic trigger level up to the high level of reconstructing the kinematics of the di-tau system – a process combining the missing transverse momentum, hadronic tau candidates and their associated tracks, and any additional jets in each event. The reconstruction of tracks, calorimeter clusters, jets, electrons, muons, taus, and missing transverse momentum are performed using standard ATLAS algorithms described in this chapter. A more detailed description of the tau reconstruction and identification is presented as the work devoted to these studies formed the foundation for this analysis, as well as many other ATLAS measurements involving taus.

3.1 Tracks and Vertices

As charged particles pass through the layers of the inner detector, they register signals in the pixels, silicon microstrips of the SCT, and straw tubes of the TRT along their path. These signals build a 3-dimensional picture of where all charged particles have interacted with the inner detector. Tracks are an algorithmic construction that aim to “connect the dots” and determine the most likely path of individual charged particles. Track reconstruction in the ATLAS detector presents a challenging problem as a result of ambiguities and high-multiplicity pile-up events.

After the raw data from the pixel and SCT detectors are converted into 3-dimensional space-points and calibrated drift circles are built from the raw TRT timing information, track reconstruction is performed in two stages [13, 14]. In the first stage, an *inside-out* algorithm begins with 3-point seeds in the three pixel layers and first SCT layer and iteratively adds hits throughout the SCT using a Kalman filter [15] outward from the interaction point to construct track candidates. Track candidates are fitted, outlier clusters are removed, ambiguities are resolved, and fake tracks are rejected by cutting on track quality criteria. Tracks are then extended into the TRT. The full tracks are then refitted combining the information from all three detectors. The tracks reconstructed by the inside-out algorithm are required to have transverse momentum $p_{\mathrm{T}} > 400$ MeV. This first stage is designed to reconstruct the primary charged particles directly produced in a *pp* interaction or from the subsequent decays or interactions of particles with a lifetime shorter than 3×10^{-11}s. A second stage, known as back-tracking, begins with unused segments in the TRT and extends them inward into the SCT and pixel detectors. This stage is designed to reconstruct charged secondary particles from photon conversions or decays of long-lived primary particles.

Primary vertices are reconstructed using an iterative vertex finding algorithm [16] on reconstructed tracks that are compatible with originating from the primary interaction region. Tracks are extrapolated into the luminous region and the z-coordinate at the point of closest approach to the beam spot center is computed for each track. The global maximum in the density of the z-coordinates yields the initial seed for a χ^2-based adaptive vertex fitting algorithm [17]. Tracks that are incompatible with the vertex by more than 7σ are used to seed the fitting of additional vertices. The *primary vertex* (PV) is defined as the vertex with the maximum sum of transverse momenta of the associated tracks. Photon conversions and secondary vertices are reconstructed by dedicated algorithms.

The helical path of a track is parametrized at the point of closest approach to the z-axis of the ATLAS coordinate system or the z-axis of a coordinate system centered at a vertex. The *perigee parameters* $(d_0, z_0, \phi_0, \theta, \frac{q}{p})$ are illustrated in Figure 3.1 and are defined as follows:

- d_0: The transverse impact parameter is the distance of closest approach to the z-axis in the $x - y$ plane. The sign of d_0 is positive when $\phi - \phi_0 = \frac{\pi}{2} \mod 2\pi$.
- z_0: The longitudinal impact parameter is the z-coordinate at the perigee.
- ϕ_0: The azimuth angle of the track at the perigee, measured in the range $[-\pi, \pi]$.
- θ: The polar angle of the track, measured in the range $[0, \pi]$.
- $\frac{q}{p}$: The ratio of the track charge to the magnitude of the track momentum.

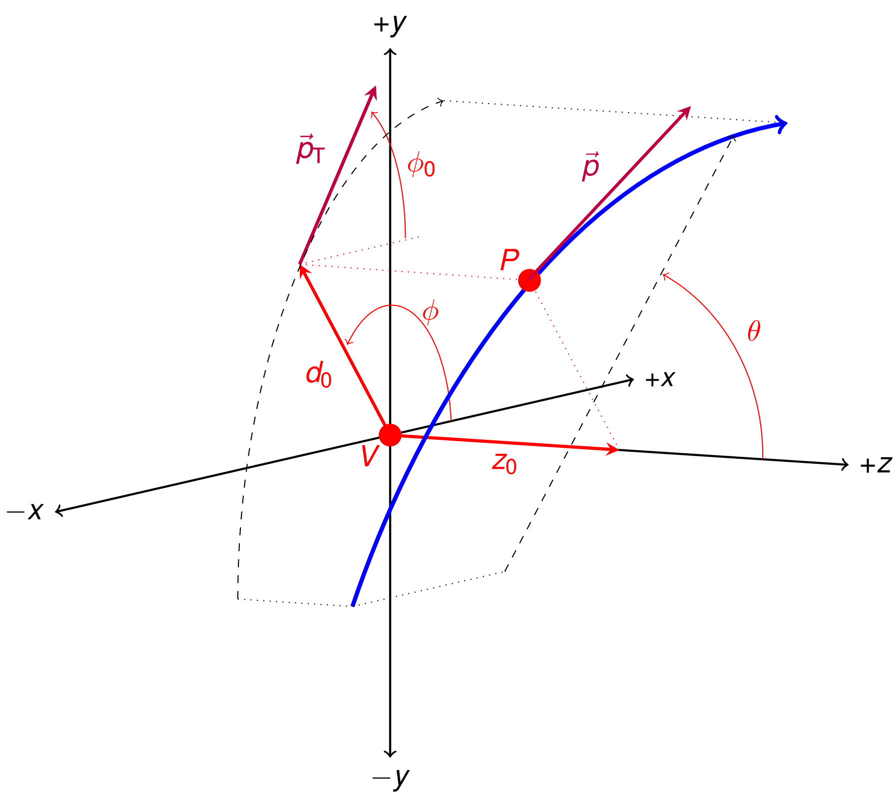


Figure 3.1: Illustration of the perigee parameters of the point of closest approach P of a track (blue curve) to the z-axis of a coordinate system centered at vertex V. The projections of the track onto the $x - y$ and $y - z$ planes are shown with the dashed black paths.

3.2 Jets

Quarks and gluons from the initial hard interaction generate collimated showers of additional quarks and gluons as energy is dissipated through radiated gluons and as gluons with sufficient energy generate quark-antiquark pairs. This *parton shower* continues, as depicted in Figure 3.2, until the particles coalesce into colour-neutral hadronic states – a process called *hadronization*. These collimated sprays of hadrons are the physically measurable objects in the ATLAS detector, and they present an indirect means of inferring properties of the initial hard interaction. The ATLAS reconstruction software groups the resulting clusters of energy deposits in the EM and hadronic calorimeters into physical abstractions[1] called "jets"[2].

Before jet reconstruction begins, topological clusters, or *topoclusters*, of energy deposits in the calorimeters are identified [19]. Calorimeter cells with a signal-to-noise ratio of at least 4 seed the construction of each topocluster. Cells with a signal-to-noise ratio of at least 2 surrounding the seed cells are included iteratively, and finally one layer of all neighbouring cells is added to the topocluster. Topoclusters are merged if a cell is to be included by more than one topocluster. Topoclusters are then split if they contain multiple local maxima. Topoclusters have an energy equal to the sum of the energies of the contained cells, zero mass, and a direction equal to a unit vector from the origin of the ATLAS coordinate system pointing at the energy-weighted topocluster barycentre.

Topocluster energies are then calibrated according to the Local Hadron Calibration scheme [20]. This calibration classifies topoclusters as mainly electromagnetic or hadronic depending on cluster shape variables. The energies of hadronic clusters are corrected for invisible energies released by non-ionization processes or escaped energy in the form of non-interacting neutrinos or escaped muons. Out-of-cluster corrections recover low-energy deposits in the tails of hadronic showers that did not satisfy the noise thresholds in the topocluster construction. All topocluster energies are also corrected for the presence of dead material. The calibrated topocluster four-momenta are then the inputs to the jet reconstruction software.

[1] The concept of a jet is algorithm-dependent.

[2] Jet reconstruction can also use tracks as inputs, however, this analysis only uses calorimeter-based jets.

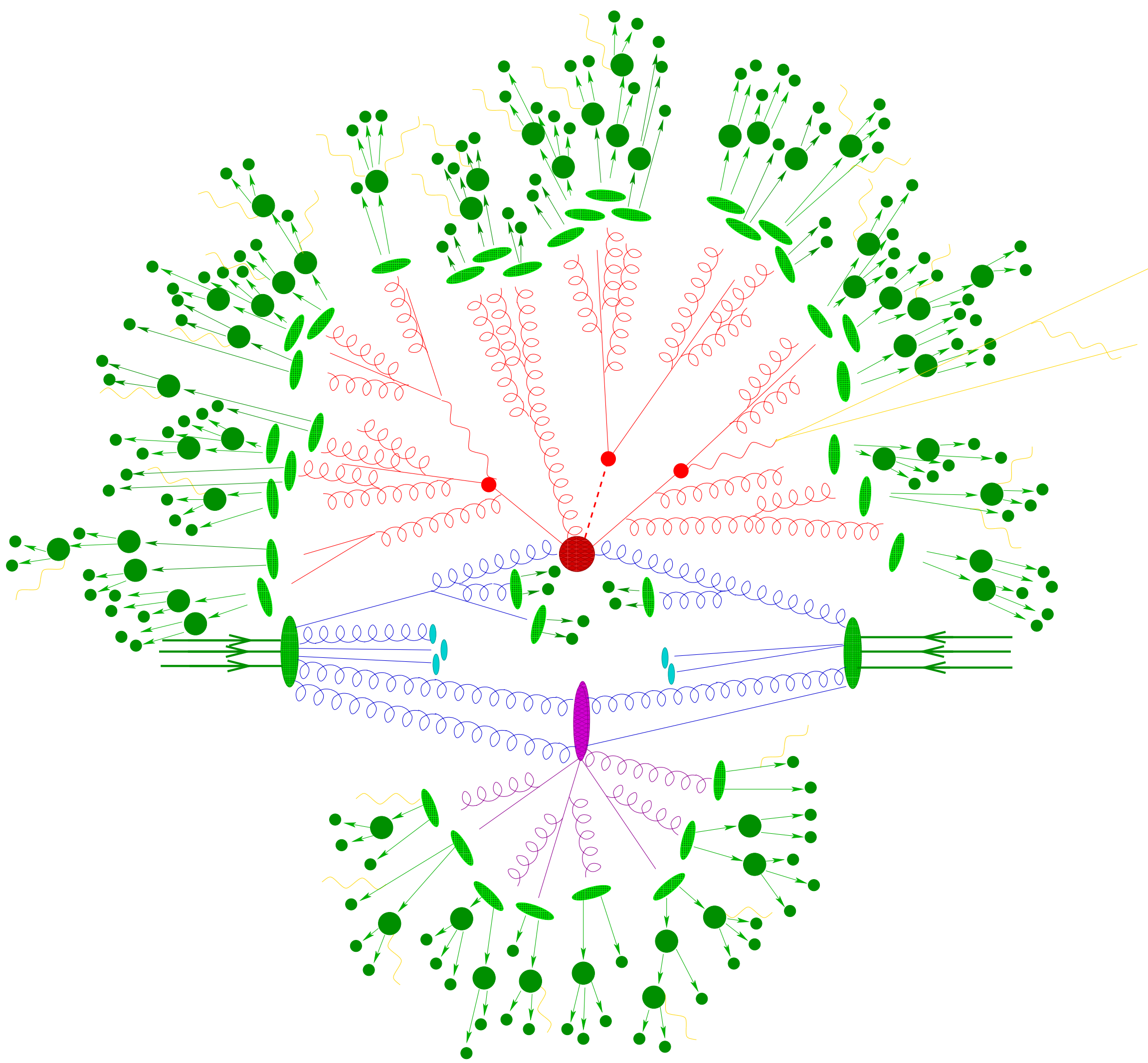


Figure 3.2: A pictorial representation of a proton-proton collision event [18]. The hard interaction (large red blob) generates a shower of quarks and gluons (red) that then hadronize (light green blobs). Unstable hadrons decay into stable hadrons (dark green blobs). Photons can radiate at any stage (yellow). Secondary interactions (purple blob) and remnants from the collision form the “underlying event”. Events from additional proton-proton interactions, referred to as “pile-up”, may be superimposed and can arise from interactions within the same bunch crossing (“in-time pileup”) or from subsequent closely spaced bunch crossings (“out-of-time pile-up”).

Jet reconstruction algorithms generally fall into one of two categories. *Cone* algorithms start with seeds (e.g. topoclusters above some energy threshold) and then determine stable cones of radius *R* containing the constituents in the vicinity of the seeds. Some cone algorithms perform this cone-finding on all seeds and employ a split-merge step to either merge overlapping jets or assign the shared constituents to the closest jet, depending on the fraction of the constituents in the overlapping region. Other cone algorithms determine a stable cone around the hardest seed in the event, then remove the constituents in this cone before repeating the process until there are no remaining seeds. The second category is composed of successive recombination jet algorithms where pairs of constituents are iteratively merged according to a distance metric. Cone algorithms are generally infrared and collinear unsafe, meaning that the output depends strongly on the presence of additional soft constituents or a collinear splitting of constituents, as shown in Figure 3.3.

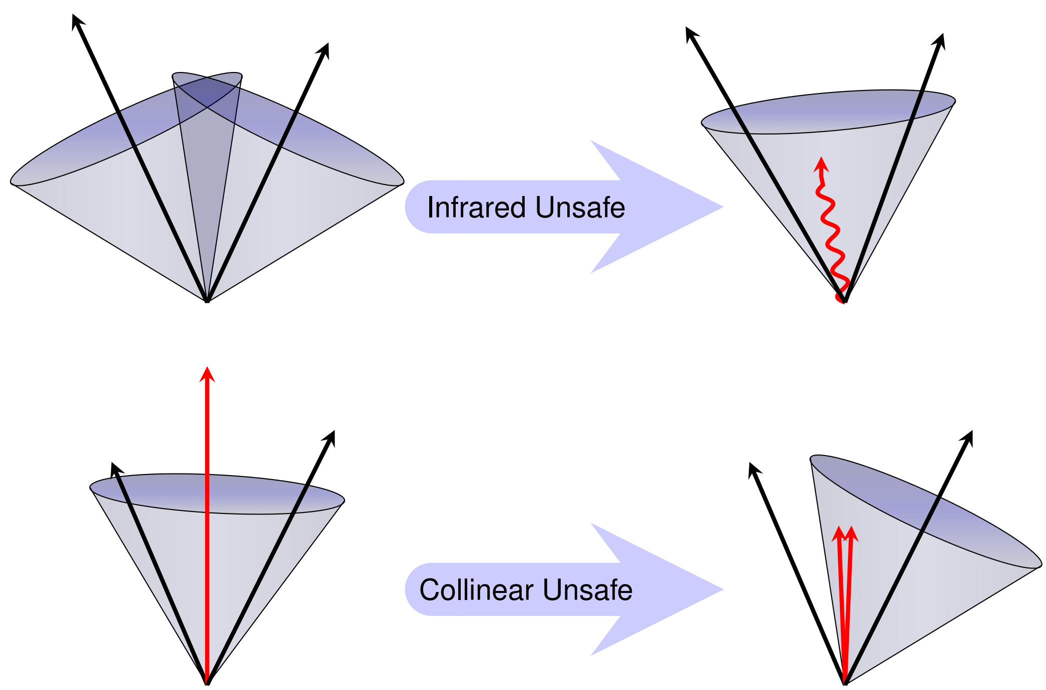


Figure 3.3: Top: An infrared unsafe jet algorithm is sensitive to the presence of soft constituents that may, for example, influence the jet splitting. Bottom: The output of a collinear unsafe jet algorithm is not invariant under the collinear splitting of constituents.

ATLAS uses the anti-k_t successive recombination algorithm [21] since it is both infrared and collinear safe, and the resulting jet boundaries are more resilient to additional soft radiation. Soft-resilient jet boundaries ease theoretical calculations, facilitate jet calibration, and are more robust against the effects of the underlying event and pile-up. The anti-k_t algorithm considers the following distance metric between pairs of constituents i and j:

$$d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2})\frac{\Delta_{ij}^2}{R^2} \tag{3.1}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, y_i and ϕ_i are the transverse momentum, rapidity, and azimuth angle, respectively. d_{ij} is larger between soft constituents than soft and hard constituents that are similarly separated. In this way, priority is given to the clustering of soft with hard constituents over clustering among soft constituents. Naively the anti-k_t algorithm appears to require $\mathcal{O}(n^3)$ operations: $\mathcal{O}(n^2)$ to generate the d_{ij} table and $\mathcal{O}(n)$ for the clustering. This would translate into excessive computing times required to cluster $\mathcal{O}(2000)$ constituents in a typical ATLAS collision event. ATLAS instead uses the `FastJet` [22] implementation with a reduced complexity of $\mathcal{O}(n \log n)$ by querying nearest neighbouring constituents with a Voronoi diagram.

Tracks can be associated to jets that lie within the coverage of the ATLAS tracking system ($|\eta| < 2.5$). To select jets that are more likely produced by the primary hard-scattering (identified by the primary vertex), a cut can be applied on the fraction of the associated track momenta from tracks that originate from the primary vertex. This so-called Jet Vertex Fraction (JVF) is defined as:

$$\mathrm{JVF} = \sum_{\text{jet tracks from PV}} p_{\mathrm{T}}^{\text{track}} \Big/ \sum_{\text{all jet tracks}} p_{\mathrm{T}}^{\text{track}} \tag{3.2}$$

This analysis uses anti-k_t jets with a distance parameter $R = 0.4$ built from topoclusters within the full coverage of the calorimeters ($|\eta| < 4.9$). Cuts on jet kinematics, quality, and the JVF are applied as outlined in Chapter 4. Jets are calibrated according to the jet energy scale (JES) [23].

3.3 Electrons and Muons

Although electrons and muons are not directly selected[3] by this analysis of the fully hadronic ditau final state, a brief description of their reconstruction is justified by the vetoing of events containing these objects, as discussed in Chapter 4.

Electron candidates considered by this analysis are reconstructed in the central region of the detector within the coverage of the inner detector. Clusters in the EM calorimeter are reconstructed and then matched with reconstructed tracks in the inner detector in the following three stages [24]. A sliding-window search is first performed over the entire EM calorimeter, using a window size of 3×5 in units of 0.025×0.025 in (η, ϕ) space, looking for windows that contain at least 2.5 GeV. Windows that satisfy this energy requirement seed the reconstruction of topoclusters. Tracks with $p_{\mathrm{T}} > 0.5$ GeV are extrapolated from their last measured point in the inner detector and are considered matched to a cluster if the point at which they intersect the middle layer of the EM calorimeter is within $|\eta| < 0.05$ of the cluster barycentre and $|\Delta\phi| < 0.1$ in the direction that the track bends in the magnetic field. The closest matching track is taken if multiple tracks satisfy the above requirements for a given cluster, giving priority to tracks with hits in the pixel detector or SCT. If no tracks match a cluster, then the cluster is considered as an unconverted photon candidate. Finally, the cluster window sizes are optimized for the different regions of the calorimeter. An electron identification is then applied using sequential cuts on discriminating features composed of calorimeter and/or tracking information. This identification rejects hadrons faking electrons, non-isolated electrons, and electrons from photon conversions. Three identification thresholds with increasing purity are defined for the purpose of physics analysis: *loose*, *medium*, and *tight*. Typical electron reconstruction and identification efficiencies, after these selections cuts, range between 80% and 90% depending on η and p_{T}.

Muon candidates are reconstructed from the association of a track in the inner detector and a track in the muon system using the STACO [25] algorithm. The momentum is evaluated from the combination of these tracks. Three levels of tightness (*loose*, *medium*, and *tight*) on the reconstructed muon quality are defined to reject fake, non-prompt, and cosmic muons. Typical muon efficiencies are greater than 98% [26].

[3] Muons are selected in data to construct the embedded $Z \rightarrow \tau\tau$ background model presented in Section 5.1.

3.4 Missing Transverse Momentum

The vector sum of the transverse components of all reconstructed momenta in a hadron collider event is in general non-zero due to the presence of neutrinos, which escape without interacting with the ATLAS detector, or from other sources that generate an imbalance in the distribution of energy deposits in the detector. The *missing* transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) is obtained from the negative vector sum of the momenta of all detected particles. Fake $E_{\mathrm{T}}^{\mathrm{miss}}$ can be generated by activity in detector cracks and dead or transition regions used for services, or from cosmic-ray or beam-halo muons crossing the detector, detector noise and other temporal defects. The presence of pile-up also degrades the resolution of true $E_{\mathrm{T}}^{\mathrm{miss}}$ from neutrinos.

The signal events in this analysis are characterized by true $E_{\mathrm{T}}^{\mathrm{miss}}$ due to the presence of the neutrinos from tau decays. Cutting on the $E_{\mathrm{T}}^{\mathrm{miss}}$ and requiring that the di-tau kinematics be consistent with the $E_{\mathrm{T}}^{\mathrm{miss}}$ direction suppresses events from multijet processes and selects more signal-like events, as discussed in Chapter 4. The $E_{\mathrm{T}}^{\mathrm{miss}}$ also plays an important role in the reconstruction of the di-tau resonance, discussed in Section 3.6.

In this analysis, the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction [27] uses calorimeter cells calibrated according to the reconstructed physics objects to which they are associated. Calorimeter cells are associated with a reconstructed and identified high-p_{T} parent object in the following order: electrons, photons, hadronically decaying tau leptons, jets, and muons. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is then calculated as the sum of terms associated to each of these physics objects:

$$E_{\mathrm{x,y}}^{\mathrm{miss}} = E_{\mathrm{x,y}}^{\mathrm{miss},e} + E_{\mathrm{x,y}}^{\mathrm{miss},\gamma} + E_{\mathrm{x,y}}^{\mathrm{miss},\tau} + E_{\mathrm{x,y}}^{\mathrm{miss},\mathit{jets}} + E_{\mathrm{x,y}}^{\mathrm{miss},\mu} + \mathrm{STVF} \times E_{\mathrm{x,y}}^{\mathrm{miss},\mathit{soft}} \tag{3.3}$$

The p_{T} of muons identified in the events are also taken into account in the $E_{\mathrm{x,y}}^{\mathrm{miss},\mu}$ term. Calorimeter cells not associated with any objects contribute to the $E_{\mathrm{x,y}}^{\mathrm{miss},\mathit{soft}}$ term and are scaled by the soft term vertex fraction (STVF) to suppress contributions from pile-up that worsen the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution, as shown in Figure 3.4. This fraction is the ratio of the scalar sum of the p_{T} of tracks from the primary vertex unmatched to physics objects to the scalar sum p_{T} of all tracks in the event also unmatched to physics objects:

$$\mathrm{STVF} = \sum_{\mathrm{tracks_{soft}},\mathrm{PV}} p_{\mathrm{T}} \Big/ \sum_{\mathrm{tracks_{soft}}} p_{\mathrm{T}} \tag{3.4}$$

The magnitude and direction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the transverse plane is then:

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{(E_{\mathrm{x}}^{\mathrm{miss}})^2 + (E_{\mathrm{y}}^{\mathrm{miss}})^2} \qquad \phi^{\mathrm{miss}} = \arctan \frac{E_{\mathrm{y}}^{\mathrm{miss}}}{E_{\mathrm{x}}^{\mathrm{miss}}} \tag{3.5}$$

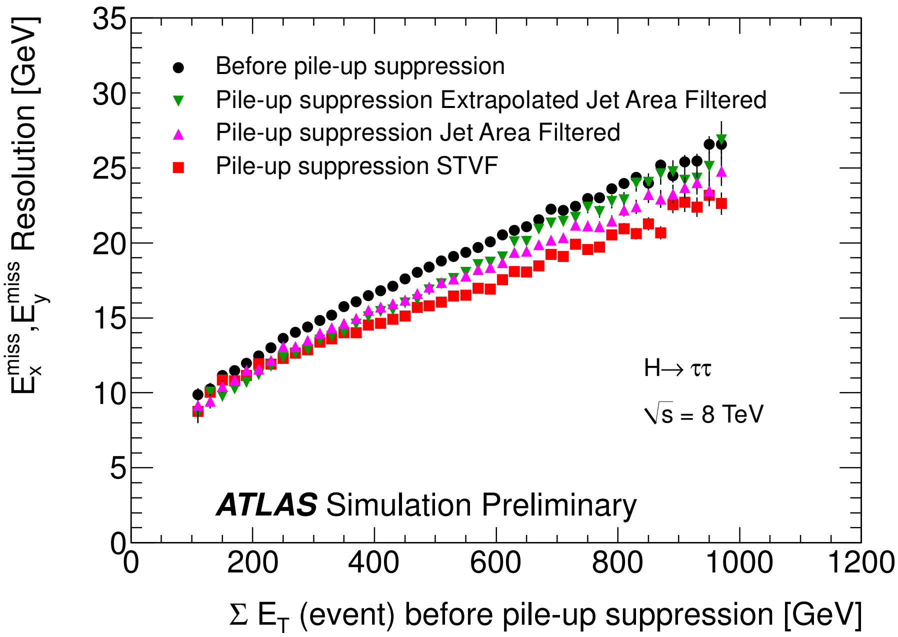


Figure 3.4: $E_{\mathrm{x}}^{\mathrm{miss}}$ and $E_{\mathrm{y}}^{\mathrm{miss}}$ resolution as a function of the total transverse energy in the event calculated by summing the p_{T} of muons and the total transverse energy in the calorimeter in MC $H \to \tau\tau$ events (m_H = 125 GeV) [27]. Results are shown before and after pile-up suppression. This analysis uses the STVF pile-up suppression.

3.5 Hadronic Tau Decays

An efficient and accurate reconstruction of tau decays with a strong rejection of processes that create similar detector signatures is essential to the feasibility of this search in the $H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ channel. Triggers are tuned to maintain a sufficiently high acceptance of true tau decays within the alloted bandwidth constraints. A tau reconstruction is designed to yield accurate tau *candidate* four-momenta and categorization by decay type. Finally, an identification stage combines many features of each reconstructed tau decay candidate and assigns a score which conveys the probability that the candidate is a true tau. Tau triggers also borrow techniques from the offline tau identification to improve the efficiency of true tau decays at the trigger level. As shown in Chapter 4, the efficiencies of the tau trigger and identification have the largest impacts on the signal acceptance.

The tau is the heaviest lepton with a mass of 1776.82 $\pm$ 0.16 MeV [6], making it the only lepton that can decay into hadrons. As presented in Table 3.1, 65% of tau decays are hadronic, resulting in predominantly one or three charged pions, possible neutral pions and photons, and one neutrino. Taus have a mean lifetime of τ = (290.6 $\pm$ 1.0) $\times$ 10^{-15}s, corresponding to a proper decay length of $c\tau$ = 87μm [6]. With this relatively short lifetime, only the tau decay products are observed by the ATLAS detector. The tau reconstruction and identification in the ATLAS detector is designed solely for the hadronic decays due to the difficulty in distinguishing prompt leptons from leptonic tau decays. Hadronic tau decays will be denoted as τ_{had} and $\tau_{\text{had-vis}}$ when excluding the invisible neutrino.

Approximately 50%, 15%, and 0.1% of all tau decays result in one, three, and five charged pions, respectively. These decay modes are labelled as 1-prong, 3-prong, etc. Reconstructed tau decays are correspondingly categorized by the number of associated tracks. Candidates with one, three, and two or more associated track(s) are labelled as $\tau_{\text{1-prong}}$, $\tau_{\text{3-prong}}$, and $\tau_{\text{multi-prong}}$, respectively. A reconstructed tau candidate is said to be a *true* or *fake* candidate in simulated events depending on whether it is matched to a true hadronically decaying tau lepton within $\Delta R < 0.2$.

This chapter will first cover tau reconstruction in Section 3.5.1, followed by the tau identification and measurement of the tau efficiency in Section 3.5.2, the tau energy calibration in Section 3.5.3, and the tau trigger in Section 3.5.4.

Decay Mode	Branching Ratio [%]
Leptonic Modes	35.2
$e^- \bar{\nu_e} \nu_\tau$	17.8
$\mu^- \bar{\nu_\mu} \nu_\tau$	17.4
Hadronic Modes	64.8
$\pi^- \nu_\tau$	10.8
$\pi^- \pi^0 \nu_\tau$	25.5
$\pi^- 2\pi^0 \nu_\tau$	9.3
$2\pi^- \pi^+ \nu_\tau$	9.0
$2\pi^- \pi^+ \pi^0 \nu_\tau$	2.7

Table 3.1: Leptonic and dominant hadronic decay modes of the τ^- lepton and the corresponding branching ratios expressed as percentages of the overall decay width [6].

3.5.1 Tau Reconstruction

The reconstruction [28] of tau candidates is seeded by anti-k_t jets (see Section 3.2) with a distance parameter R = 0.4 using topoclusters calibrated with the Local Hadron Calibration [20]. All jets with a transverse momentum $p_T \geq 10$ GeV and $|\eta| \leq 2.5$ (within the coverage of the ATLAS tracking system) are considered as tau candidates. The four-momenta of reconstructed tau candidates are defined in terms of three degrees of freedom: p_T, η, and ϕ. The mass of $\tau_{\text{had-vis}}$ candidates is defined to be zero, therefore the transverse momentum, p_T, and transverse energy, $E_T = E\sin(\theta)$, are equal. The tau reconstruction algorithm associates tracks to the topocluster jets in the calorimeter and calculates a set of discriminating variables used by the identification algorithms. A custom tau energy scale (TES), separate from the JES, is also derived since hadronic tau decays consist of a particular mix of charged and neutral pions. Due to the typically narrow collimation of tau decays, a *core region* within $\Delta R \leq 0.2$ is used to determine the primary tau attributes, such as the number of prongs, the charge, and the energy scale.

Vertex Association

Before the tau track selection is performed, a dedicated vertex association algorithm is used to determine the most likely vertex of origin for each tau candidate. Within a high

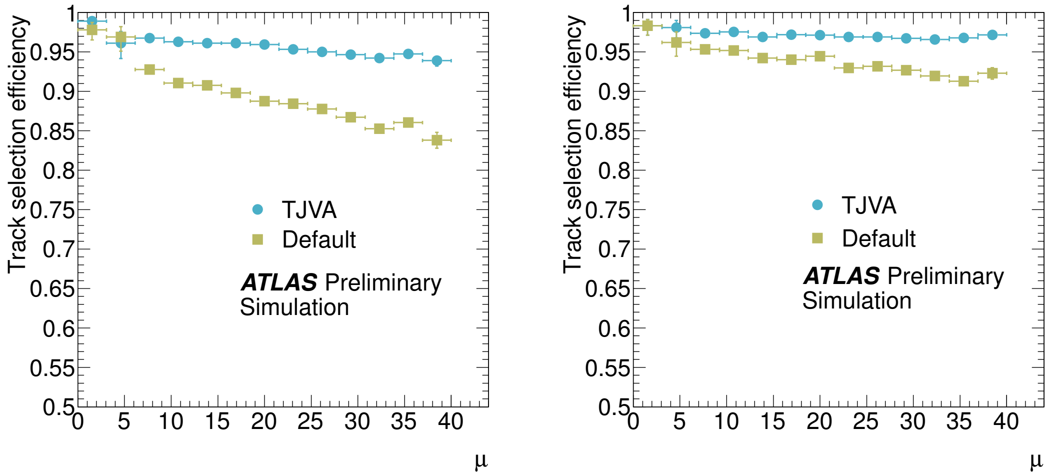


Figure 3.5: The tau track selection efficiency for true (matched) 1-prong (left) and 3-prong candidates in simulated $Z \to \tau\tau$ events versus the average number of pileup interactions per bunch crossing (μ). The track multiplicity is less sensitive to pileup with the more accurate vertex association provided by the TJVA algorithm [29].

pileup environment the probability of assigning a pileup vertex as the primary vertex increases. Incorrectly assigning the primary vertex as the origin of a tau candidate then leads to significant changes in the reconstructed track multiplicity as tracks fail a cut on the z_0 impact parameter. The Tau Jet Vertex Association (TJVA) algorithm [29] builds on the JVF quantity introduced in Section 3.2 by finding for each tau candidate a vertex resulting in the highest JVF. As shown in Figure 3.5 the TJVA algorithm is able to maintain a stable track selection efficiency in simulated $Z \to \tau\tau$ events as the average number of pileup interactions per bunch crossing (μ) is increased.

The Tau Axis

The tau axis is calculated using the topoclusters within $\Delta R \leq 0.2$ of the barycenter of the jet seed (see Section 3.2). The four-momenta of these core topoclusters are recalculated using the coordinate system centered at the tau vertex and the vectors are then summed. The (η, ϕ) components of this vector define the *intermediate* tau axis. The *final* tau axis is calculated after the energy calibration is performed below in Section 3.5.3, where a correction to the pseudorapidity is applied.

Track Association

Tracks are associated to each tau candidate within $\Delta R \leq 0.4$ of each tau axis if they satisfy the following quality criteria:

- $p_{\mathrm{T}} \geq 1$ GeV
- At least two hits in the pixel detector
- A total of at least seven hits in the pixel and SCT detectors
- $|d_0| \leq 1.0$ mm
- $|z_0 \sin\theta| \leq 1.5$ mm

where the transverse and longitudinal impact parameters, d_0 and z_0 are calculated with respect to the vertex determined by the TJVA algorithm. The number of tracks within the core region of $\Delta R \leq 0.2$ determines the number of prongs. The tau charge is determined by the sum of the track charges in the core region. Tracks within the annulus of $0.2 < \Delta R \leq 0.4$ are used to calculate discriminating features used by the identification.

3.5.2 Tau Identification

The tau reconstruction algorithm outlined above provides very little rejection against fake candidates from QCD-initiated jets or electrons and muons. A separate set of algorithms have been designed to specifically reject these sources of fake candidates using the information contained in a set of discriminating variables calculated by the tau reconstruction software. The tau identification used during the 2012 data-taking period is largely a re-optimization of the identification algorithms pioneered in the earlier data [29, 30] with improvements on the robustness in a higher pileup environment. This section will focus on the rejection of QCD-initiated jets since they represent the dominant source of fake tau candidates in this analysis.

QCD Jet Rejection

QCD multijet production occurs at a rate that is roughly six to nine orders of magnitude greater than $Z \rightarrow \tau\tau$ or $H \rightarrow \tau\tau$, respectively, and quark- or gluon-initiated jets produce detector signatures that are very similar to hadronically decaying taus. There are, however, several features that distinguish taus from QCD jets that, when combined by an identification algorithm, yield sufficient QCD jet rejection. QCD jets tend to have a higher number of constituents. Jets from hadronically decaying taus tend to be more collimated in both the distribution of topoclusters in the calorimeters, and associated tracks. Tracks from tau decays also tend to have a high impact parameter significance due to displacement of the tau decay. These characteristics are captured by the following set of discriminating variables:

- **Pile-up-corrected core energy fraction ($f^{\mathrm{corr}}_{\mathrm{core}}$):** Pile-up-corrected fraction of transverse energy in the central region $\Delta R < 0.1$ of the tau cone:

$$f^{\mathrm{corr}}_{\mathrm{core}} = \frac{\sum_{i}^{\Delta R_i < 0.1} E^{\mathrm{EM}}_{\mathrm{T},i}}{\sum_{j}^{\Delta R_j < 0.2} E^{\mathrm{EM}}_{\mathrm{T},j}} + (0.003 \times N_{\mathrm{vtx}} \text{ if } p_{\mathrm{T}} < 80 \text{ GeV}) \tag{3.6}$$

  where $E^{\mathrm{EM}}_{\mathrm{T},i}$ ($E^{\mathrm{EM}}_{\mathrm{T},j}$) is the transverse energy, calibrated at the EM energy scale, deposited in cell i (j), and i iterates over the cells associated with the tau candidate within $\Delta R < 0.1$ of the intermediate tau axis, while j iterates over all cells within $\Delta R < 0.2$. N_{vtx} is the the number of pile-up vertices with at least two associated tracks plus the primary vertex, which is required to have at least four associated tracks. The p_{T} of the tau candidate is calibrated at the tau energy scale. The more collimated jets from hadronic tau decays will tend to deposit a higher fraction of the total jet energy in the core region, giving higher values of $f^{\mathrm{corr}}_{\mathrm{core}}$. Signal and background distributions of $f^{\mathrm{corr}}_{\mathrm{core}}$ are shown in the top left plot in Figure 3.6.

- **Pile-up-corrected leading track momentum fraction ($f^{\mathrm{corr}}_{\mathrm{track}}$):**

$$f^{\mathrm{corr}}_{\mathrm{track}} = \frac{p^{\mathrm{leadtrk}}_{\mathrm{T}}}{\sum_{j}^{\Delta R_j < 0.2} E^{\mathrm{EM}}_{\mathrm{T},j}} + 0.003 \times N_{\mathrm{vtx}} \tag{3.7}$$

  where $p^{\mathrm{leadtrk}}_{\mathrm{T}}$ is the transverse momentum of the leading p_{T} core track of the tau candidate. $E^{\mathrm{EM}}_{\mathrm{T},j}$ and N_{vtx} are defined as stated above for $f^{\mathrm{corr}}_{\mathrm{core}}$. $f^{\mathrm{corr}}_{\mathrm{track}}$ Tau decays tend to impart a larger fraction of the total momentum to the leading track while the momentum tends to be more evenly distruted across all tracks in a QCD jet.

- **Track radius (R_{track}):** p_{T}-weighted track width:

$$R_{\mathrm{track}} = \frac{\sum_{i}^{\Delta R_i \leq 0.4} p_{\mathrm{T},i} \Delta R_i}{\sum_{i}^{\Delta R_i \leq 0.4} p_{\mathrm{T},i}} \tag{3.8}$$

where i iterates over all tracks within the $\Delta R \leq 0.4$ cone around the intermediate tau axis and $p_{\mathrm{T},i}$ is the track transverse momentum. This is a track-based counterpart of $f_{\mathrm{core}}^{\mathrm{corr}}$ defined above and tends to be smaller for true tau candidates. Signal and background distributions of R_{track} are shown in the top right plot in Figure 3.6.

- **Maximum ΔR (ΔR_{max}):** The maximum ΔR between a track in the core region and the intermediate tau axis will tend to be smaller for true tau candidates. Signal and background distributions of ΔR_{max} are shown in the bottom left plot in Figure 3.6. This variable is only used for multi-prong candidates.

- **Transverse flight path significance ($S_{\mathrm{T}}^{\mathrm{flight}}$):** The decay length significance in the transverse plane of the secondary vertex for multi-prong tau candidates:

$$S_{\mathrm{T}}^{\mathrm{flight}} = \frac{L_{\mathrm{T}}^{\mathrm{flight}}}{\delta L_{\mathrm{T}}^{\mathrm{flight}}} \tag{3.9}$$

  where $L_{\mathrm{T}}^{\mathrm{flight}}$ is the reconstructed signed decay length and $\delta L_{\mathrm{T}}^{\mathrm{flight}}$ is its estimated uncertainty. Only tracks in the core region are used to fit a secondary vertex. The displacement of a reconstructed secondary vertex tends to be more significant for true tau candidates. Signal and background distributions of $S_{\mathrm{T}}^{\mathrm{flight}}$ are shown in the bottom right plot in Figure 3.6.

- **Leading track IP significance ($S_{\mathrm{lead\ track}}$):** Transverse impact parameter significance of the leading associated track in the core region:

$$S_{\mathrm{lead\ track}} = \frac{d_0}{\delta d_0} \tag{3.10}$$

  where d_0 is the distance of closest approach of the track to the tau vertex in the transverse plane, and δd_0 is its estimated uncertainty. This variable offers similar information to $S_{\mathrm{T}}^{\mathrm{flight}}$ defined above but is used for 1-prong candidates where a secondary vertex is not reconstructed.

- **Track mass m_{tracks}:** Invariant mass of the track system, including all associated tracks with the cone $\Delta R < 0.4$. m_{tracks} is constrained by the tau mass and will tend to be smaller than for a QCD jet where it will scale more with the jet energy. This variable is only used for multi-prong candidates.

- **Number of tracks in the isolation annulus ($N_{\mathrm{track}}^{\mathrm{iso}}$):** Number of associated tracks within the annulus $0.2 < \Delta R \leq 0.4$ around the intermediate tau axis. $N_{\mathrm{track}}^{\mathrm{iso}}$ tends to be smaller for true tau candidates. This variable is only used for 1-prong candidates.

Calorimeter-based variables are affected more by contributions from pile-up than track-based variables, with pile-up events producing extra energy deposits within the tau cone. To compensate for the increased pile-up conditions in the 2012 data-taking period, the size of the cone in which calorimeter cells are summed in the definitions of $f^{\text{corr}}_{\text{core}}$ and $f^{\text{corr}}_{\text{track}}$ is decreased from the 2011 value of $\Delta R < 0.4$ to $\Delta R < 0.2$, and an additional correction dependent on N_{vtx} is applied. The 2012 tau identification also uses a subset of the 2011 variables that were seen to be more robust against pile-up.

The above discriminating variables are combined into one identification variable using boosted decision trees (BDTs) or a projective likelihood method (LLH) [31]. The multivariate BDT technique has become the primary tau identification method used by the ATLAS collaboration, forming a critical component of many measurements of the Standard Model and searches for new physics. BDTs are also used to separate signal from background at the event level within the context of this search for $H \to \tau_{\text{had}}\tau_{\text{had}}$ and a detailed description of the BDT algorithm is presented in Section 6.2. Using the TMVA package [32], BDTs are trained separately for 1-prong and 3-prong candidates where the signal sample is populated with reconstructed tau candidates in simulated $Z, Z' \to \tau\tau$ and $W \to \tau\nu$ events that are matched with true hadronically decaying taus and the background sample is populated with candidates from jet-enriched data. The 3-prong BDT is then used to classify all multi-prong candidates. The desired signal efficiency or background rejection (defined here as the inverse background efficiency) of the tau identification can be attained by placing a cut on the BDT output, where higher values correspond to higher purity. One metric used to compare the performance of identification methods is the inverse background efficiency versus signal efficiency curve. At a fixed signal efficiency, the method with the lower background efficiency and therefore higher inverse background efficiency is the more powerful method. Figure 3.7 compares the performances of the BDT and LLH identification methods.

For the purpose of standardizing tau selection across physics analyses in the ATLAS collaboration, three signal efficiency *working points* – *loose*, *medium*, and *tight* – are defined, corresponding to 70%, 60% and 40% for 1-prong candidates and 65%, 55%, and

35% for multi-prong candidates, respectively. The working points are determined separately in p_{T} bins to compensate for the p_{T} dependence of the BDT output. Signal efficiencies of the working points are therefore roughly flat as a function of p_{T}, as shown in Figure 3.8. Figure 3.8 also shows that the revised list of variables and pile-up corrections used by the 2012 tau identification yields efficiencies that are approximately independent of the pile-up conditions.

Electron and Muon Rejection

Electrons can create tracks and clusters that resemble 1-prong tau decays, however, properties including the emission of transition radiation by the electron track, and a shower in the calorimeter that is typically shorter and more narrow, can be used to discriminate against them. Similar to the QCD jet rejection, BDTs are trained using several track and shower shape variables to discriminate between true 1-prong tau decays and electrons [28]. Three working points – *loose*, *medium*, and *tight* – are defined on the BDT output, corresponding to 95%, 85% and 75% signal efficiency, respectively.

Most muons passing through the ATLAS detector are minimum ionizing particles and rarely deposit enough energy in the calorimeters to create fake tau candidates. Physics analyses also employ an object overlap removal procedure such that tau candidates matching reconstructed muons are not considered. Muons can fail reconstruction if they have low energy and are stopped in the calorimeter, the track is skewed enough by the calorimeter and reconstruction fails in the Muon Spectrometer, or if the muon passes through an inefficient region of the Muon Spectrometer. A fake reconstructed tau candidate can then remain after overlap removal if such a muon track is wrongly associated with a calorimeter cluster with sufficient energy. A muon veto is designed using cuts on the fraction of energy deposited in the electromagnetic calorimeter and on the fraction of energy carried by the highest momentum track.

Tau Identification Efficiency

The signal efficiencies of the tau identification working points are tuned using simulated events and it is important to measure the same efficiencies in data and to determine scale

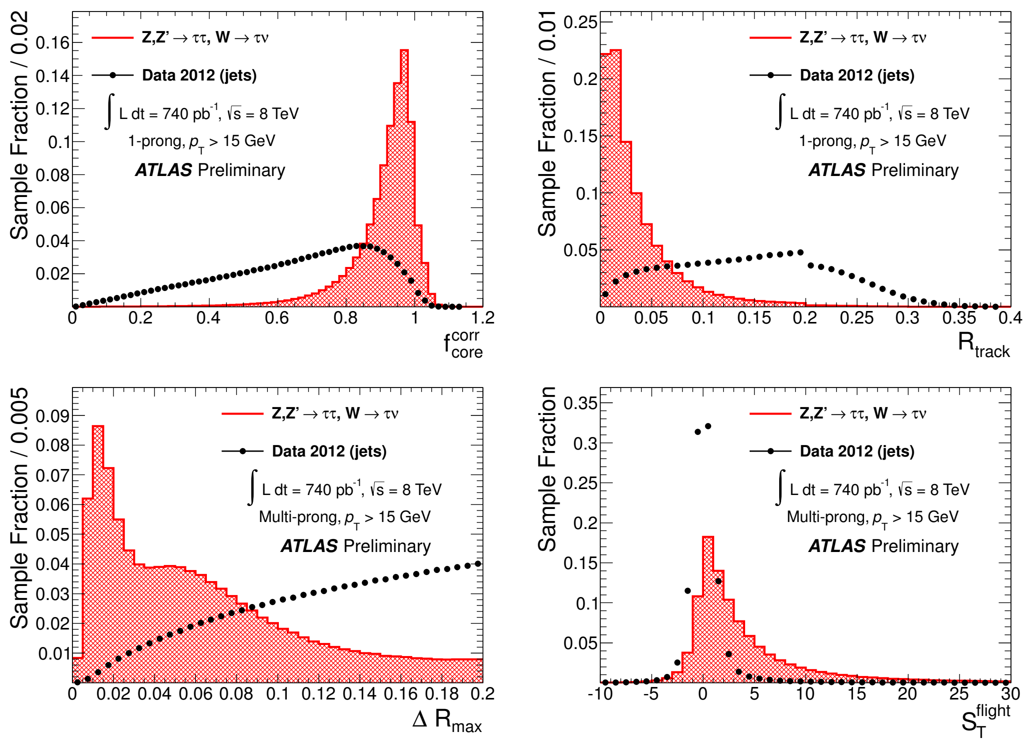


Figure 3.6: Distributions of a selection of jet discriminating variables for simulated $Z, Z' \to \tau\tau$ and $W \to \tau\nu$ signal samples and a jet background sample selected from 2012 data [28]. The distributions are normalized to unit area.

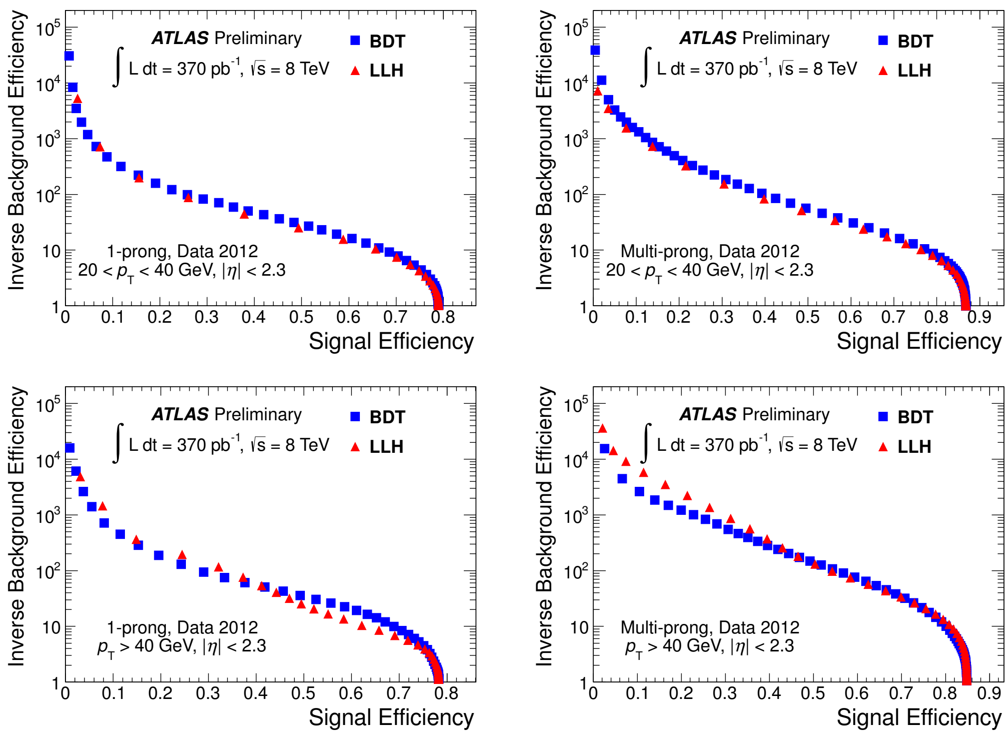


Figure 3.7: Inverse background efficiencies as a function of signal efficiency for 1-prong (left) and multi-prong (right) candidates, in low (top) and high (bottom) p_T ranges, for the two tau ID methods: boosted decision trees (BDT) and a projective likelihood method (LLH). The signal efficiencies were obtained using simulated $Z, Z' \to \tau\tau$ and $W \to \tau\nu$ samples and the inverse background efficiencies from data multijet events [28]. The statistical uncertainty is comparable to the size of the markers.

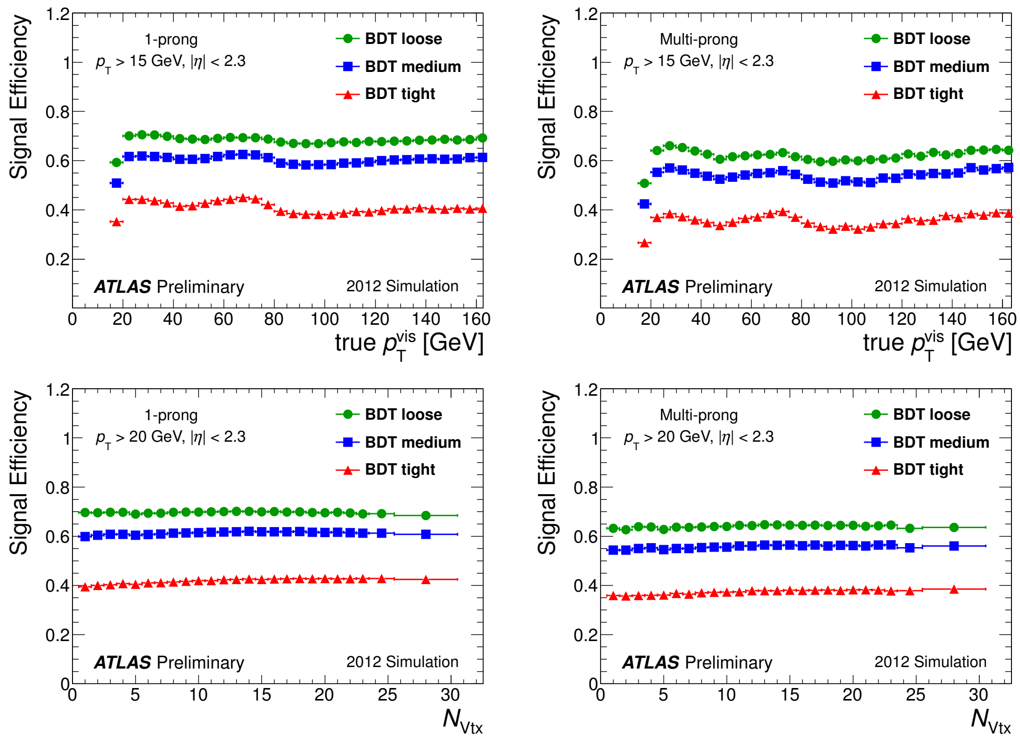


Figure 3.8: Signal efficiencies for 1-prong (left) and multi-prong (right) candidates for the three working points of the BDT tau ID as a function of true visible tau p_T (top) and number of vertices (bottom). The efficiencies were obtained using simulated $Z, Z' \to \tau\tau$ and $W \to \tau\nu$ samples for signal and multijet events from data for background [28].

factors that correct for differences. Since it is impossible to select a pure sample of real taus in data, a simultaneous fit of several background templates along with a tau template is performed using a variable that discriminates between real and fake taus. This fit is performed before and after each tau identification threshold and the corresponding signal tau efficiencies in data are measured.

The track multiplicity distribution of reconstructed true taus has the signature peaks in the 1 and 3-track bins and presents a useful template to discriminate between true and fake taus. To improve the constraint on the fake tau background from jets, which typically generate tracks in a much wider cone, additional tracks within the $0.2 < \Delta R \leq 0.6$ annulus around the core region are included according to a p_{T}-correlated track counting algorithm inspired by the anti-k_t jet algorithm. For all tracks in this annulus with $p_{\mathrm{T}} > 500$ MeV and that satisfy the tau reconstruction track quality criteria the following distance metric is calculated with respect to all tracks in the core region:

$$D = \frac{p_{\mathrm{T}}^{\mathrm{core}}}{p_{\mathrm{T}}^{\mathrm{outer}}} \times \Delta R(\mathrm{core}, \mathrm{outer}) \tag{3.11}$$

If $D < 4$ for any combination of core and outer track, the other track is included. The threshold of 4 was optimized to populate the high track multiplicity tail from QCD jets while mostly preserving the track multiplicity of true tau candidates.

A fit of the recounted track multiplicity in selected $Z \to \tau_\mu \tau_{\mathrm{had}}$ data events is performed before and after tau ID, as shown in Figure 3.10, and efficiency scale factors are derived. The systematic uncertainty on the scale factors is determined by varying the signal templates by changing the generator, hadronic shower model, fragmentation model parameters, and detector geometry [28, 29]. This fitting procedure has been repeated on selected $W \to \tau_{\mathrm{had}} \nu_\tau$ and $t\bar{t} \to \tau_{\mathrm{had}}$ + jets events as a cross-check in a low p_{T} and busy high p_{T} environments. A comparison of efficiency scale factors is shown in Figure 3.9.

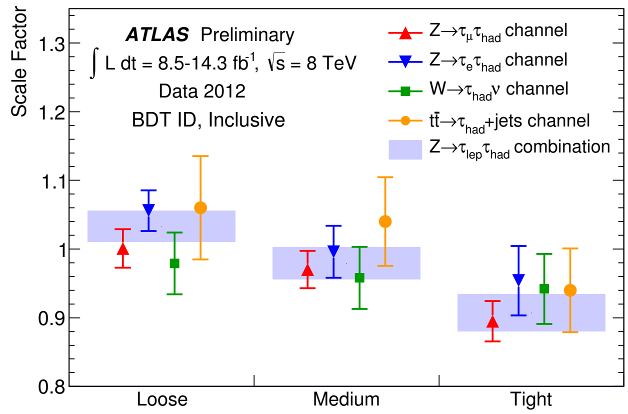


Figure 3.9: Inclusive scale factors for all channels and all working points [28].

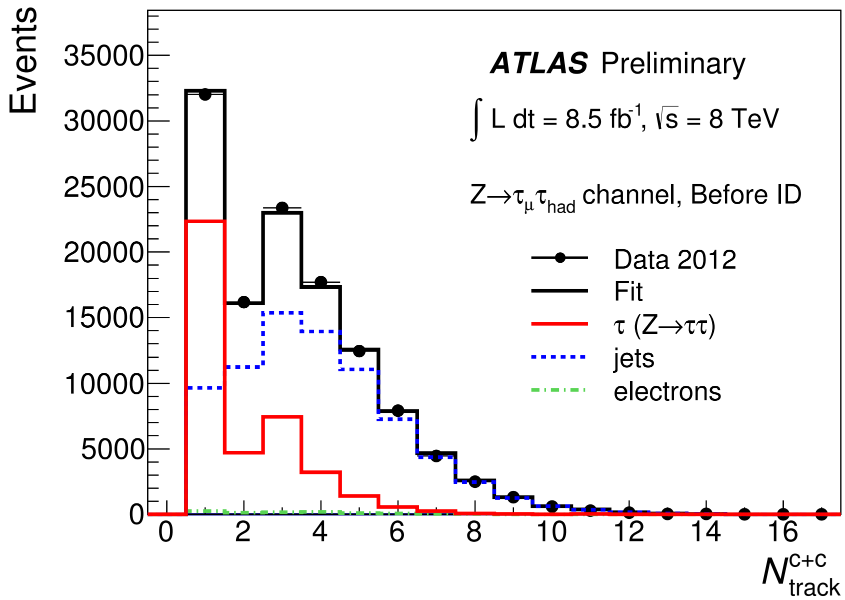
3.5.3 The Tau Energy Scale

Jets seeding tau reconstruction are composed of topoclusters calibrated by the Local Hadron Calibration (LC) that improves the tau energy resolution with respect to topoclusters at the electromagnetic (EM) energy scale. While LC compensates for invisible or escaped energy, out-of-cluster deposits, or dead material, LC does not correct for energy lost before the calorimeters, underlying event and pileup contributions, or out-of-cone effects. A tau energy scale (TES) separate from the jet energy scale (JES) is also justified by the typically higher fraction of energy deposited in the electromagnetic calorimeters than in QCD jets. The TES calibration compensates for the above effects and restores the reconstructed momentum to the true visible tau momentum scale.

The TES is derived from simulated $W \rightarrow \tau\nu$ and $Z, Z' \rightarrow \tau\tau$ events that include contributions from in-time and out-of-time pileup. Only reconstructed tau candidates passing the medium BDT identification (c.f. Section 3.5.2) and matched with a true hadronically decaying tau lepton with transverse momentum of at least 10 GeV are used. Events are excluded where a reconstructed jet with $p_{\mathrm{T}} > 15$ GeV is within $\Delta R < 0.5$ of a selected tau candidate. The calibrated momentum is then:

$$p^{\tau} = \frac{p^{\tau}_{\mathrm{LC}} - p^{\tau}_{\mathrm{pileup}}}{R(p^{\tau}_{\mathrm{LC}}, |\eta^{\tau}{}_{\mathrm{reco}}|, n^{\tau}_{p})} \tag{3.12}$$

where R is the response curve defined as the ratio of the reconstructed momentum at LC



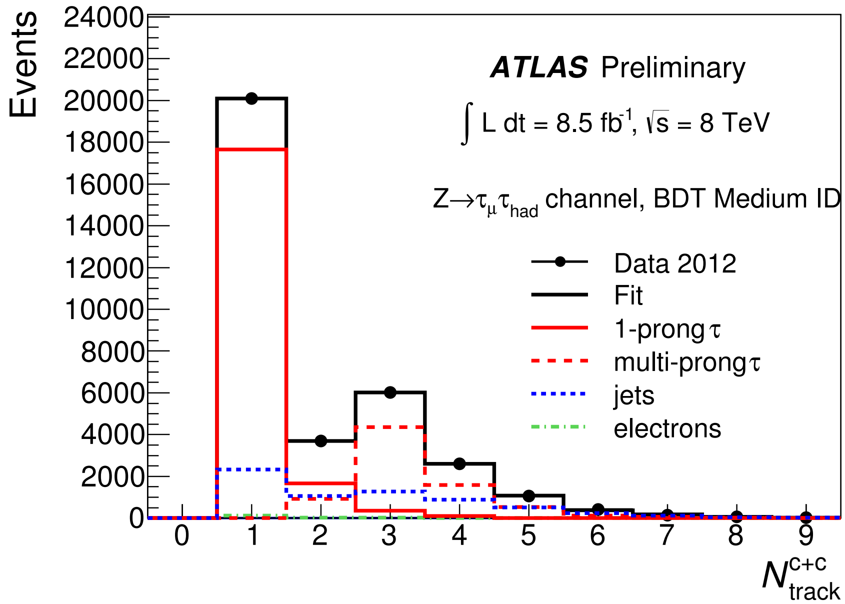
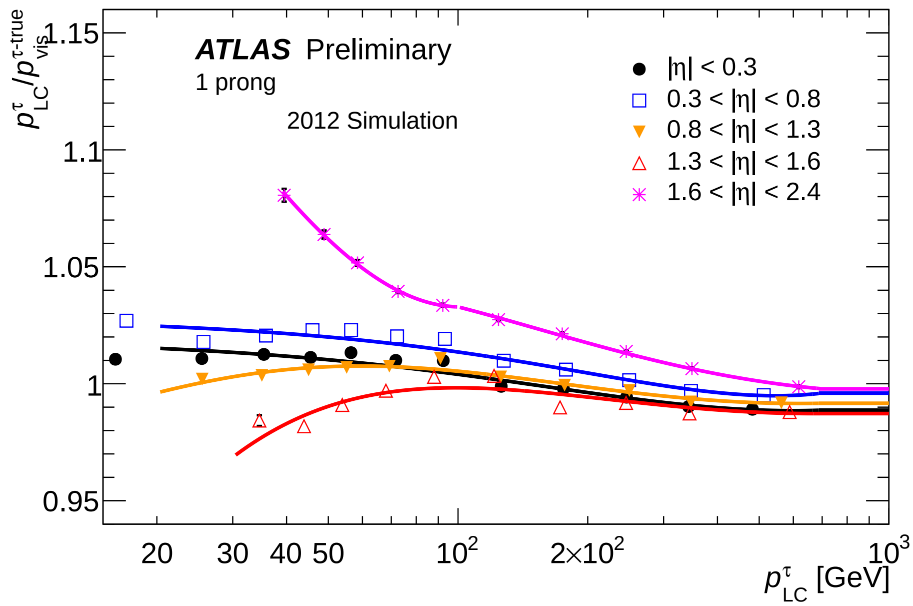


Figure 3.10: Fit before tau ID (top) and after BDT medium tau ID (bottom) for the measurement of the 1-prong and multi-prong identification efficiencies. All ID methods and working points are included in the same fit. The tau signal template and the electron template are taken from simulations. The jet template is obtained from data in a control region. The distribution shown is the sum of core and p_{T}-correlated tracks [28].

scale p^{τ}_{LC} to the visible momentum of the matched true tau decay. R, shown in Figure 3.11, is a function of p^{τ}_{LC} and is determined separately in bins of $|\eta^{\tau}{}_{\mathrm{reco}}|$ and number of prongs n^{τ}_{p}. $p^{\tau}_{\mathrm{pileup}}$ represents a pileup correction dependent on the number of reconstructed primary vertices. The pseudorapidity of tau candidates is also corrected to compensate for topoclusters in poorly instrumented regions of the calorimeters. The final tau pseudorapidity is then $|\eta^{\tau}| = |\eta^{\tau}{}_{\mathrm{reco}}| - \eta_{\mathrm{bias}}$ where $\eta_{\mathrm{bias}} = \langle|\eta^{\tau}{}_{\mathrm{reco}}| - |\eta^{\tau}{}_{\mathrm{true}}|\rangle$. The momentum resolution for 1-prong and multi-prong candidates after the above calibration is shown in Figure 3.12. 1-prong candidates have a smaller momentum resolution because of the higher fraction of energy deposited in the electromagnetic calorimeter from π^0 decays.



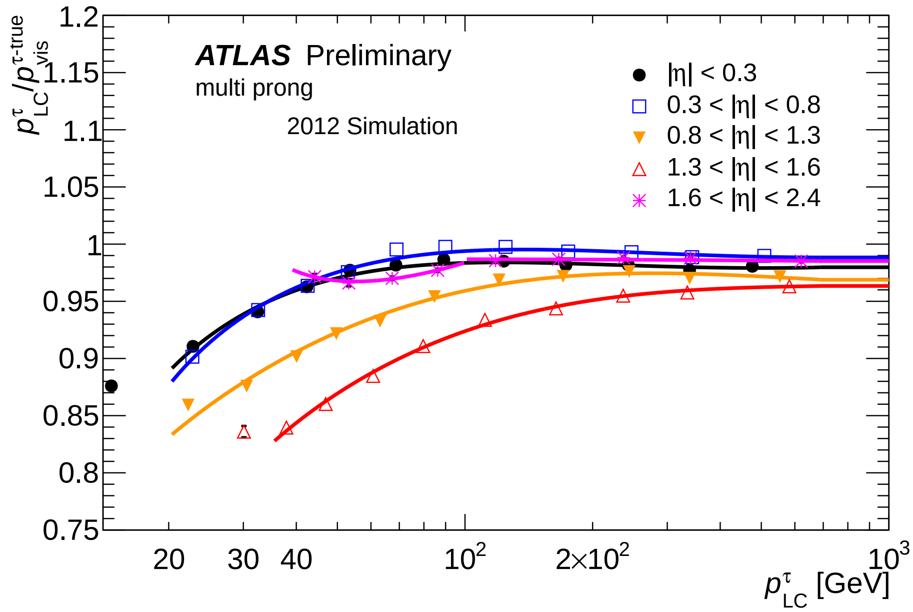


Figure 3.11: Response curves as a function of the reconstructed tau momentum at LC scale for 1-prong (top) and multi-prong (bottom) candidates in bins of $|\eta_{\text{reco}}|$ [33].

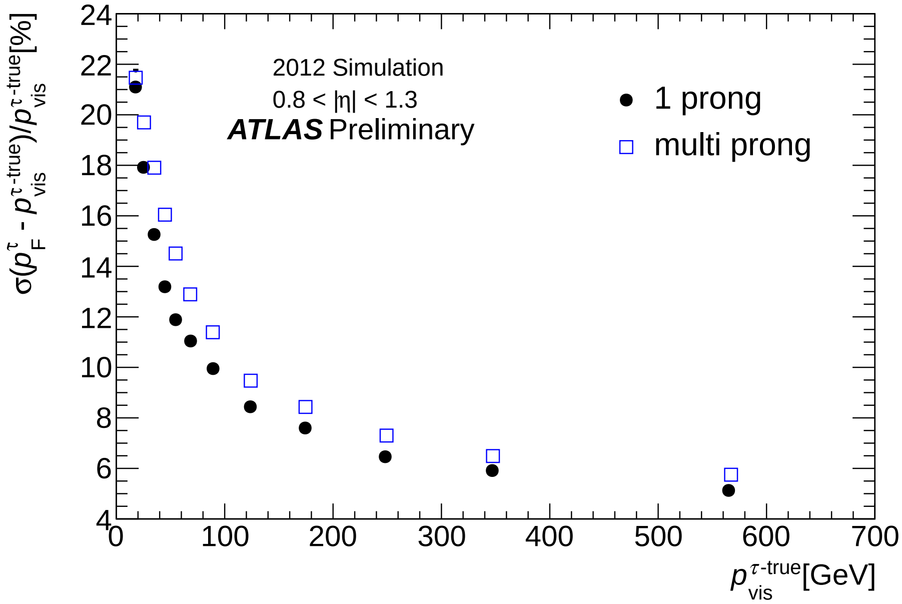


Figure 3.12: Momentum resolution for 1-prong and multi-prong candidates in the region $0.8 < |\eta| < 1.3$ [33].

The simulation-based TES is also applied to tau candidates in data but a mismodelling of the detector response can result in a different energy scale. Therefore an in-situ correction to data must be applied such that the momenta of candidates in data match the predicted momenta in simulation. $Z \to \tau_\mu \tau_{\text{had}}$ events are selected in data and the invariant mass distribution of the muon and hadronic tau is compared with simulation. An in-situ TES shift is determined by a scaling of the p_{T} that produces the best statistical agreement between the mass peak in data and simulation.

The TES is influenced by uncertainties on the calorimeter response, the underlying event and pile-up models, detector model, the hadronic shower model, as well as the non-closure of the calibration method on reconstructed simulated tau jets. The calorimeter response component is estimated by measuring the single particle responses and convolving with the particle composition of hadronic tau decays. The total size of the TES uncertainty ranges from 2-3% and is shown in Figure 3.13.

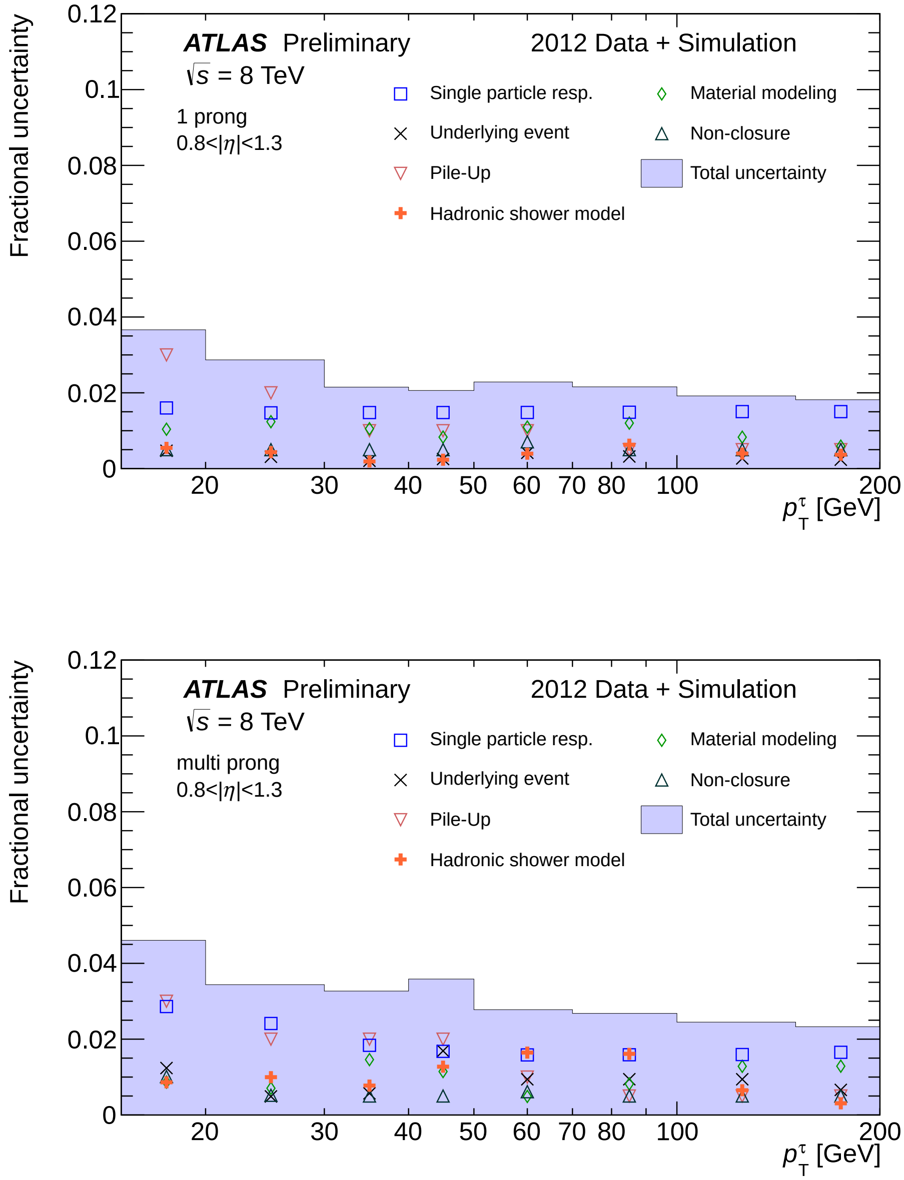


Figure 3.13: The systematic uncertainty on the tau energy scale in bins of the calibrated visible transverse tau momentum for 1-prong (top) and multi-prong (bottom) candidates in the pseudorapidity range $0.8 < |\eta| < 1.3$ [33].

3.5.4 Triggering on Taus

A dedicated set of tau triggers are designed to provide an efficient selection of taus for physics analysis. The optimization at the trigger level is critical to the success of many measurements and searches involving taus due especially to strict bandwidth limits while the instantaneous luminosity delivered by the LHC has increased. Tau trigger components have been implemented at all three levels of the ATLAS trigger system.

At L1, the tau trigger uses electromagnetic and hadronic calorimeter towers of cells to calculate energy in a core region and an isolation region around the core [34]. Trigger calorimeter towers have a granularity of $\Delta\eta \times \Delta\phi$ = 0.1 $\times$ 0.1. The core region is composed of 2$\times$2 trigger towers and the isolation region is composed of the 4$\times$4 box of towers around the core region. Regions of interest (ROIs) are identified with core regions that contain at least 8, 11, 15, 20, or 40 GeV. These ROIs are labeled as `L1_TAUX` where `X` is the energy threshold in the core region. Triggering on events with at least one ROI would require an energy threshold of well over 100 GeV to keep within the maximum trigger rate. This would however result in a very low efficiency selection of $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ or $H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ events. Instead, L1 tau trigger ROIs must be combined with other trigger items (electrons, muons, or $E_{\text{T}}^{\text{miss}}$) or combined with each other to create di-tau triggers. To cope with increased rates at L1 in the later part of 2011 and for the 2012 data-taking periods, tau trigger items with isolation cuts were implemented, requiring no more than 4 GeV in the isolation trigger tower regions of the 8 and 11 GeV ROIs.

The full granularity of all calorimeter layers and the tracking system is available at L2. L2 refines the positions of the L1 ROIs by considering all calorimeter cells in $\Delta\eta \times \Delta\phi$ = 0.8 $\times$ 0.8. After noise suppression is applied to all calorimeter cells, shower shape variables similar to those used by the offline tau identification are calculated. Tracks are reconstructed and track-based variables are also calculated. The L2 tau trigger then applies cuts on these variables to reject fake tau ROIs while optimizing true tau efficiencies.

At the EF level, more sophisticated identification algorithms are possible due to the higher latency. A multivariate trigger using boosted decision trees very similar to the offline tau identification presented in Section 3.5.2 is used, however, since full vertex reconstruction is not possible at the trigger level, discriminating variables are not pile-up corrected.

3.6 Kinematic Reconstruction of the Ditau Resonance

The kinematics of a di-tau system is described by the following underconstrained system of four equations:

$$
\begin{aligned}
E_{\mathrm{x}}^{\mathrm{miss}} &= p_{\mathrm{mis}_1} \sin\theta_{\mathrm{mis}_1} \cos\phi_{\mathrm{mis}_1} + p_{\mathrm{mis}_2} \sin\theta_{\mathrm{mis}_2} \cos\phi_{\mathrm{mis}_2} \\
E_{\mathrm{y}}^{\mathrm{miss}} &= p_{\mathrm{mis}_1} \sin\theta_{\mathrm{mis}_1} \sin\phi_{\mathrm{mis}_1} + p_{\mathrm{mis}_2} \sin\theta_{\mathrm{mis}_2} \sin\phi_{\mathrm{mis}_2} \\
M_\tau^2 &= m_{\mathrm{mis}_1}^2 + m_{\mathrm{vis}_1}^2 + 2\sqrt{p_{\mathrm{vis}_1}^2 + m_{\mathrm{vis}_1}^2}\sqrt{p_{\mathrm{mis}_1}^2 + m_{\mathrm{mis}_1}^2} - 2p_{\mathrm{vis}_1} p_{\mathrm{mis}_1} \cos\alpha_{vm_1} \\
M_\tau^2 &= m_{\mathrm{mis}_2}^2 + m_{\mathrm{vis}_2}^2 + 2\sqrt{p_{\mathrm{vis}_2}^2 + m_{\mathrm{vis}_2}^2}\sqrt{p_{\mathrm{mis}_2}^2 + m_{\mathrm{mis}_2}^2} - 2p_{\mathrm{vis}_2} p_{\mathrm{mis}_2} \cos\alpha_{vm_2}
\end{aligned} \tag{3.13}
$$

where $\alpha_{vm_{1,2}}$ is the angle separating the “visible” and “missing” tau decay products and $M_\tau = 1.777\ \mathrm{GeV}/c^2$ is the tau mass. The unknowns include the components of the momenta carried by the neutrinos for each tau decay as well as the invariant mass of the two neutrinos, $m_{\mathrm{mis}_{1,2}}$, from any leptonic decay. This system therefore has a maximum of eight unknowns if both tau decays are leptonic, where each decay produces two neutrinos, and a minimum of six unknowns if both taus are hadronic decays since each decay only produces one neutrino and $m_{\mathrm{mis}_{1,2}} = 0$.

The Missing Mass Calculator (MMC) [35] uses constraints from the measured *x*- and *y*-components of $E_{\mathrm{T}}^{\mathrm{miss}}$ and the visible masses of both tau candidates to perform a grid scan over the two components of the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector and the $(\phi_{\mathrm{mis}_1}, \phi_{\mathrm{mis}_2})$ plane. The corresponding $\alpha_{vm_{1,2}}$ and di-tau mass can be determined at each grid point and is assigned a probability according to the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution and the tau decay topologies. The MMC is then able to produce an estimate of the most probable di-tau mass $m_{\tau\tau}^{\mathrm{MMC}}$ for $\sim$99% of $H \to \tau\tau$ and $Z \to \tau\tau$ events. The small loss rate of about 1% is due to large fluctuations of the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement or other scan variables. In figure 3.14, reconstructed $m_{\tau\tau}^{\mathrm{MMC}}$ mass distributions are shown for $Z \to \tau\tau$ and $H \to \tau\tau$ with a Higgs mass of 125 GeV in the *VBF* and *Boosted* categories.

Another important quantity is the transverse momentum $p_{\mathrm{T}}^{\mathrm{H}}$ of the resonance. This is reconstructed using the vector sum of the event $E_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse momentum of the visible tau decay products.

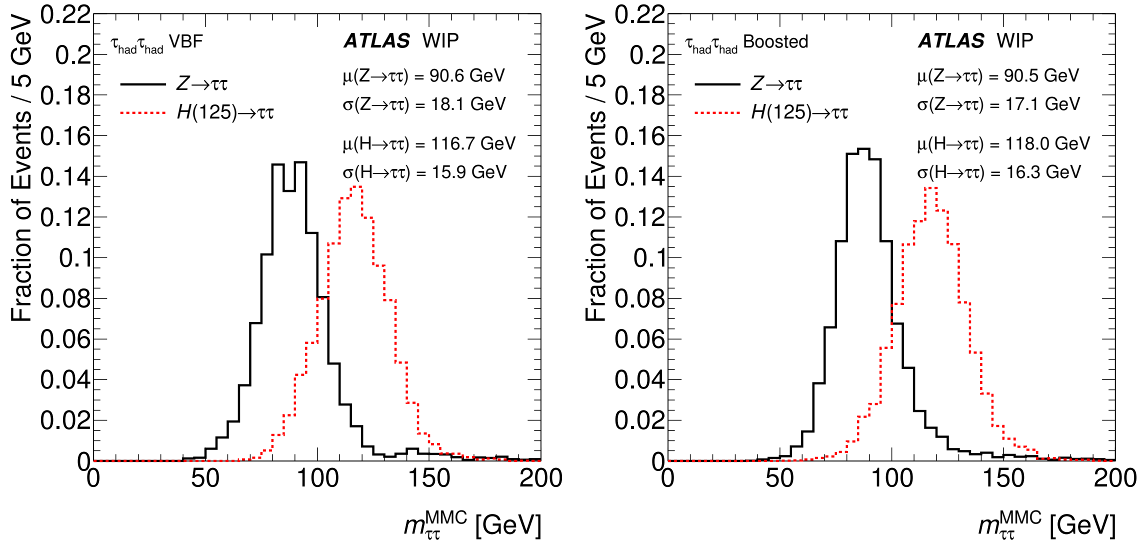


Figure 3.14: The reconstructed $m_{\tau\tau}^{\text{MMC}}$ mass distributions for $H \to \tau\tau$ (m_H = 125 GeV) and $Z \to \tau\tau$ events in MC simulation and embedding, respectively, for events in the *VBF* (left) and *Boosted* (right) categories. The mean (μ) and root mean square (σ) of the distributions are given as well.

Chapter 4

Event Selection and Categorization

This chapter outlines the event selection and the categorization used by both the cut-based and multivariate analyses. The purpose of the first stage of selection, called the "preselection", is to accept events that pass the data quality requirements and that fire the double hadronic tau trigger, to reject events that contain reconstructed electron or muon candidates, and finally to accept events where two likely tau candidates are flagged. The events that pass this preselection are then partitioned into categories that separately exploit the different kinematics and event topologies exhibited by the VBF and gluon-fusion Higgs production mechanisms. Before presenting the details of the event preselection and categorization, the data samples collected by the ATLAS Experiment during the 2012 run and the tau triggers used by this analysis are first discussed.

4.1 Data Samples

The data used for this search have been collected in proton-proton collisions by the ATLAS Experiment at the LHC at the center of mass energy of 8 TeV in 2012. Only the events recorded when all the subsystems of the ATLAS detector were operating and that pass quality checks are used. This amounts to an integrated luminosity of 20.3 fb^{-1}. The total integrated luminosity versus time that was delivered by the LHC, recorded by the ATLAS detector, and finally passing the data quality assessments is shown in Figure 4.1.

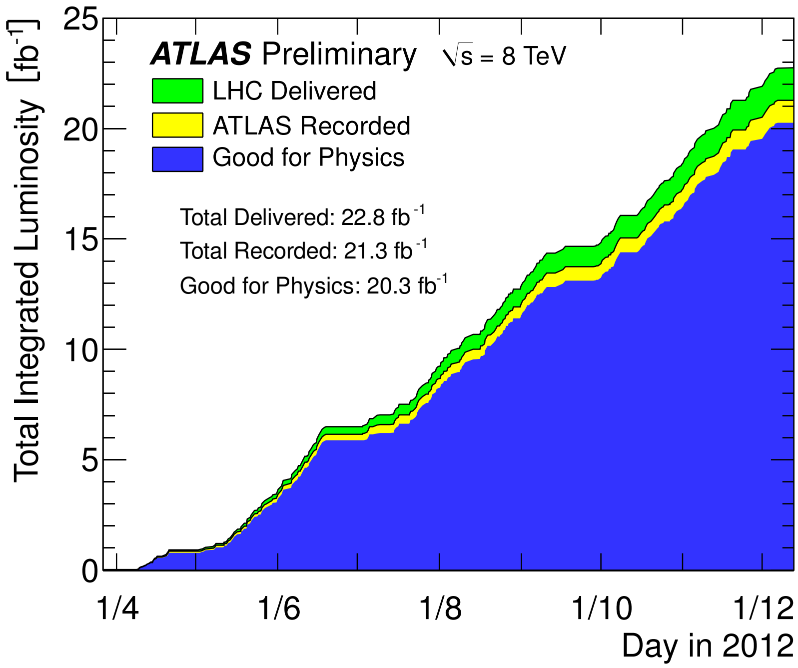


Figure 4.1: Cumulative luminosity versus time delivered to (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams and for pp collisions at 8 TeV centre-of-mass energy in 2012. The delivered luminosity accounts for the luminosity delivered from the start of stable beams until the LHC requests ATLAS to put the detector in a safe standby mode to allow a beam dump or beam studies. The recorded luminosity reflects the DAQ inefficiency, as well as the inefficiency of the so-called "warm start": when the stable beam flag is raised, the tracking detectors undergo a ramp of the high-voltage and, for the pixel system, turning on the preamplifiers. The data quality assessment shown corresponds to the All Good efficiency shown in the 2012 DQ table. The luminosity shown represents the preliminary 8 TeV luminosity calibration. Data quality has been assessed after reprocessing.

4.2 Double Hadronic Tau Triggers

Among all the collected events, only the events that "fire" the double-hadronic-tau trigger `EF_tau29Ti_medium1_tau20Ti_medium1` are considered. These triggered events have two reconstructed hadronic tau trigger objects with transverse momenta of at least 29 and 20 GeV and that pass the *medium* threshold of the tau trigger object identification at the event filter (EF) level of the ATLAS trigger system. The EF level tau identification uses boosted decision trees during the 2012 data taking period that are based on the boosted decision trees used for the offline tau identification. Both trigger objects must also be reconstructed with one to three associated tracks. This trigger configuration is optimized according to allotted bandwidth limits and is the lowest-threshold unprescaled trigger over the entire 8 TeV dataset, although it does have a significant impact on the overall signal acceptance. 84% of the expected VBF-produced Higgs events are already lost by the trigger and the offline p_{T} thresholds on the tau candidates used in this analysis are set where this trigger becomes efficient, as shown in Figure 4.2.

4.3 Object Definitions

This section presents the requirements on the reconstructed objects used in both the cut-based and multivariate analyses. All requirements are summarized in Table 4.1.

4.3.1 Electrons and Muons

The presence of reconstructed electrons or muons passing quality criteria results in the removal of such events in the analysis of the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel. Although electrons or muons are not directly relevant to the signal regions of the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ analysis, the quality criteria on these objects are consistent with what is used in the analyses of the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ channels to ensure orthogonality of the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ signal regions with the other channels.

Reconstructed electron candidates are selected if they have a transverse momentum greater than 15 GeV and are within the region of $|\eta| < 2.47$ but not in the detector "crack region" of 1.37 $< |\eta| <$ 1.52 at the transition between the barrel and end-cap calorimeters. Electrons must also pass the `medium` electron identification. Additional quality criteria are also applied [36].

Reconstructed muon candidates are selected if they pass the STACO `loose` muon identification, are within the region of $|\eta| <$ 2.5, and have a transverse momentum greater than 10 GeV. Muons must have either a track in both the muon spectrometer and inner detector or only in the inner detector. Quality criteria on the inner detector track associated to the muon are also applied [37].

4.3.2 Jets

Only jets within the region of $|\eta| <$ 4.5 and a transverse momentum greater than 30 GeV are considered.

A jet-vertex fraction (JVF) is used to reduce the number of jets in the event due to pile-up activity in the same or nearby bunch-crossings for each event. The JVF is defined as the ratio between the sum of the transverse momentum of the tracks in the jet associated to the primary vertex and the sum of the transverse momentum of the tracks in the jet associated to any vertex in the event. Conventionally, JVF = -1 is assigned to jets with no associated tracks. To suppress contributions from the increased pile-up conditions in the 8 TeV run, jets with $|\eta| <$ 2.4 and $p_{\mathrm{T}} <$ 50 GeV are required to have $|\mathrm{JVF}| >$ 0.5.

The lower threshold on the jet transverse momentum is increased by 5 GeV for jets beyond the tracking system ($|\eta| >$ 2.4) to reduce contributions from event pile-up and the underlying event.

4.3.3 Hadronic Tau Leptons

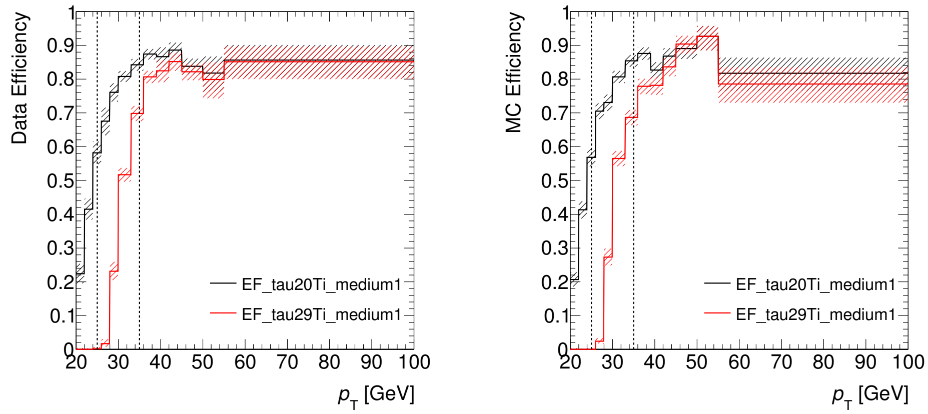
Tau candidates are required to have at least one associated track. Additional requirements on the charge and number of tracks are applied in the event categorization discussed in Section 4.4 and 4.5. The calorimeter-based tau direction and the direction of the leading track must satisfy $|\eta| <$ 2.47. The leading track must also not lie within the transition region between the barrel and end-cap calorimeters (1.37 $< |\eta| <$ 1.52) to avoid true electrons not reconstructed as electrons but misidentified as tau candidates.

The boosted decision tree (BDT) tau identification method [30] is used to identify tau candidates, by requiring that candidates pass the `medium` identification threshold, corresponding to approximately 55–60% efficiency. The standard muon veto and the `loose` electron veto using a BDT-based algorithm are applied [38] on 1-prong tau candidates only.

The two hadronic tau candidates with the highest transverse momentum are then selected while requiring that the leading tau candidate has a transverse momentum of at least 35 GeV and the subleading candidate has a transverse momentum of at least 25 GeV. These thresholds are determined by where the double hadronic tau trigger becomes efficient, as shown in Figure 4.2.

4.3.4 Object Overlap Removal

When different objects selected according to the above criteria overlap each other geometrically within a cone of $\Delta R < 0.2$, only one of them is considered for further analysis. The overlap is resolved by selecting muon, electron, hadronic tau candidate, and jet, in this order of priority. This priority is consistent across the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$, $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$, and $\tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ channel analyses. After the selection of the hadronic tau candidates above, only two candidates remain, and since events are rejected if they contain electron or muon candidates as in Section 4.3.1, the overlap removal in the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel analysis simply consists of the removal of any jets matching either of the selected tau candidates within a cone of $\Delta R < 0.2$.



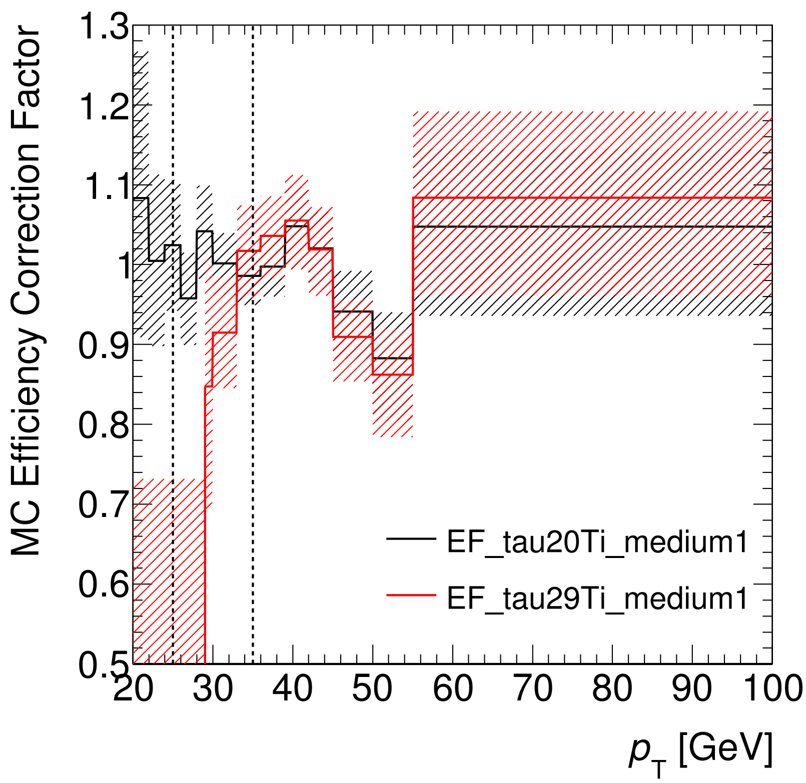


Figure 4.2: The figures above show the trigger efficiencies in data and simulation, and the factor correcting simulation to data. Trigger efficiencies are parametrized by p_{T}, η, number of tau tracks, tau identification level, and data period. Curves for medium 1-prong taus in the barrel region for periods B to D are shown. Vertical dashed lines show the analysis tau p_{T} thresholds of 25 and 35 GeV.

Object	Selection Cuts

Muons	$p_{\mathrm{T}} > 10$ GeV
	$\lvert\eta\rvert < 2.5$
	`Staco` *loose*
	no `expectedBLayerHit` or `nBLHits` > 0
	`nPixHits` + `nPixelDeadSensors` > 0
	`nSCTHits` + `nSCTlDeadSensors` > 4
	`nPixHoles` + `nSCTHoles` < 3
	if $\lvert\eta\rvert < 1.9$:
	`nTRTHits` + `nTRTOutliers` > 5 and `nTRTOutliers` < 0.9(`nTRTHits` + `nTRTOutliers`)
	if $\lvert\eta\rvert \geq 1.9$:
	`nTRTHits` + `nTRTOutliers` ≤ 5 or `nTRTOutliers` < 0.9(`nTRTHits` + `nTRTOutliers`)

Electrons	$p_{\mathrm{T}} > 15$ GeV $\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.47$ $\lvert q\rvert = 1$ `author` = 1 or 3 `mediumPP` OQ&1446=0
Taus	$p_{\mathrm{T}} > 20$ GeV $\lvert\eta\rvert < 2.5$ leading-p_{T} track in $\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.5$ `author` = 1 or 3 $\lvert q\rvert = 1$ i.e. $N_{\mathrm{tracks}} = 1$ or 3 $N_{\mathrm{tracks},\Delta R<0.6} = 1$ or 3 BDT `medium` = 1 Electron BDT `loose` = 0 if $N_{\mathrm{tracks}} = 1$ Muon veto = 0
Jets	$R = 0.4$ anti-k_t with LCW topological clusters $p_{\mathrm{T}} > 30$ GeV $\lvert\eta\rvert < 4.5$ if $\lvert\eta\rvert > 2.4$: $p_{\mathrm{T}} > 35$ GeV if $\lvert\eta\rvert < 2.4$ and $p_{\mathrm{T}} < 50$ GeV: $\lvert\mathrm{JVF}\rvert > 0.5$

Table 4.1: Reconstructed object requirements.

4.4 Event Selection

Following the selection of the tau candidates and jets, events are selected in two stages. The initial event selection, known as the preselection, serves to remove poorly reconstructed events and select events with two good quality tau candidates. Due to the complexity of the ATLAS event data model, the cumulative size of the data and simulation samples is too large to store on local resources. The preselection also provides suitable data reduction on the Worldwide LHC Computing Grid so that the remaining components of the analysis can be performed locally. Following the preselection, events are categorized by kinematic properties and event topologies consistent with the VBF and gluon-fusion Higgs production mechanisms. Categorization is discussed separately in Section 4.5. This section details the event preselection summarized in Table 4.2.

The preselection begins in data or embedded data by requiring that each event has passed the central ATLAS data quality assessment. ATLAS uses a document called a *Good Runs List* (GRL) containing the luminosity blocks that are free of identified data defects. The GRL documents are also used to determine the integrated luminosity collected by the experiment. Events are then required to fire the unprescaled double-hadronic-tau trigger as discussed in Section 4.2. To reject non-collision events such as those from cosmic rays, events are required to have at least one primary vertex with at least four associated tracks, where each track has $p_{\mathrm{T}} > 500$ MeV. Events are rejected if there were problems with the liquid argon or tile calorimeters, such as a high voltage trip. All selected jets must satisfy the "*looser*" jet quality criteria [39]. This requirement rejects events containing bad jets arising from hardware problems such as energy spikes or coherent noise in the calorimeters, effects from variability in the LHC beam conditions, or cosmic-ray showers.

The minimum transverse momenta of the leading and subleading tau candidates, 35 and 25 GeV, were mentioned in Section 4.3.3 since that is the first point in the event selection where the tau collection is reduced to only two candidates (after already selecting all tau candidates passing the `medium` identification). Then only the jets that overlap with those two tau candidates are excluded, as discussed in Section 4.3.4. To increase the rejection of fake tau candidates, at least one candidate is then required to pass the `tight` tau identification threshold. The tau candidates are also required to originate from the same primary vertex.

The cut $0.8 < \Delta R_{\tau\tau} < 2.4$ is applied to reject events where the two tau candidates are overlapping and poorly reconstructed, or where the tau candidates are back-to-back, a region dominated by fake tau candidates from QCD multijet events.

Finally, cuts on the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector are applied to enhance the purity of real $H \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$ events where a neutrino from each of the hadronic tau decays are typically emitted close to the tau candidate directions. In the decay of the Higgs boson, the taus are also boosted and the measured $E_{\mathrm{T}}^{\mathrm{miss}}$ vector will usually lie within the minor segment of the transverse plane subtended by the tau candidates. It is kinematically possible for the true $E_{\mathrm{T}}^{\mathrm{miss}}$ to point outside of the tau candidates and the imperfect reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution will also yield such events, therefore events where the $E_{\mathrm{T}}^{\mathrm{miss}}$ is close to a tau candidate, but not pointing between the tau candidates, are also accepted. These conditions are satisfied by the requirement $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV and that the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector is between the direction of the two tau candidates or within $\Delta\phi < \pi/4$ in the transverse plane from the nearest tau candidate: $\min\left\{\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{had1}}), \Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{had2}})\right\} < \pi/4$.

The cut $|\Delta\eta_{\tau\tau}| < 1.5$ removes non-resonant events and is repeated in each category instead of being placed at preselection so that a fake-enriched control region in each category can be constructed by inverting this cut.

Preselection
1. For data and embedded data each event must be contained in the GRL
2. Double hadronic tau trigger was fired: `EF_tau29(Ti)_medium1_tau20(Ti)_medium1`
3. Vertex with at least four associated tracks
4. No error in the liquid argon or tile calorimeters, such as a high voltage trip
5. All reconstructed jets must satisfy the "*looser*" jet quality requirements
6. No selected electrons or muons (see Table 4.1)
7. At least two selected tau candidates passing the *medium* BDT identification threshold
8. Select the leading tau and subleading tau candidates by p_{T} and require that the leading tau have $p_{\mathrm{T}} > 35$ GeV and the subleading tau have $p_{\mathrm{T}} > 25$ GeV
9. At least one of the selected tau candidates must the pass the BDT *tight* identification threshold
10. The tau candidates originate from the same primary vertex
11. $0.8 < \Delta R_{\tau\tau} < 2.4$
12. $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV
13. Either $E_{\mathrm{T}}^{\mathrm{miss}}$ vector is between the direction of the two tau candidates or $\min\left\{\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{had1}}), \Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \tau_{\mathrm{had2}})\right\} < \pi/4$

Table 4.2: Summary of the event preselection.

4.5 Event Categorization

Following the event preselection, events are then categorized according to the kinematics and topologies that separately enhance the purities of events produced by the vector boson fusion (VBF) and gluon-gluon fusion (ggF) Higgs production mechanisms. The categorizations used by the cut-based and multivariate analyses differ, as shown in Table 4.3 and graphically in Figure 4.4. While both analyses have categories catering to the VBF and ggF processes, the categorization philosophy is fundamentally different between them. The cut-based analysis takes advantage of combining more categories that cut harder into the feature space. Since the cut-based analysis only uses the ditau invariant mass to fit and test for the presence or absence of the Higgs signal, a more aggressive categorization is required to obtain a signal sensitivity that is comparable to the performance of the multivariate analysis. In the multivariate analysis, the categorization is kept as loose as possible to maintain a sufficiently high number of events in each category to train the boosted decision trees. The performance of the multivariate analysis is then achieved in each category by combining the information from multiple features into a final discriminant – the output of the boosted decision trees. In some sense, the cut-based analysis partitions the feature space in a similar way to what is automatically performed by the decision trees in the multivariate analysis.

Events are first passed through a series of selections that target the topology of VBF signal events. Events must contain at least two hard jets with a p_{T} of 50 and 30 GeV and that are separated longitudinally. These two jets will be referred to as the VBF-tagged jets. There is no cut on the maximum number of jets or veto of any additional jets that lie between the VBF-tagged jets to avoid the theoretical uncertainties associated with the suppression of events with real emissions from higher-order QCD corrections. The multivariate analysis VBF category requires that $|\Delta\eta(\mathrm{jet}_1, \mathrm{jet}_2)| > 2$ while the cut-based analysis cuts tighter with $|\Delta\eta(\mathrm{jet}_1, \mathrm{jet}_2)| > 2.6$. The cut on $|\Delta\eta(\mathrm{jet}_1, \mathrm{jet}_2)|$, shown in Figure 4.3, separates the VBF from ggF signal events and reduces the uncertainty on the signal content of the VBF category, which is larger for the ggF production mechanism. This cut also produces a VBF category that has a more direct constraint on the Higgs gauge coupling. This completes the definition of the VBF category in the multivariate analysis. Events are then tested for inclusion in the boosted category of the multivariate analysis.

The cut-based analysis continues with the VBF selection by requiring that the invariant mass of the VBF-tagged jets (M_{jj}) be above 250 GeV and that the tau candidates are emitted centrally within the longitudinal region bounded by the pseudo-rapidities of the VBF-tagged jets: $\min(\eta_{\text{jet},1}, \eta_{\text{jet},2}) < \eta_{\tau,1}, \eta_{\tau,2} < \max(\eta_{\text{jet},1}, \eta_{\text{jet},2})$. Events are then assigned to three separate VBF categories. The most sensitive, the *VBF High-p_T^H* category, isolates the high transverse momentum region of the Higgs events by cutting at $p_{\text{T}}^{\text{H}} > 140$ GeV and $\Delta R_{\tau\tau} < 1.5$. As shown by the contours of Figure A.1, this requirement rejects most of the multijet background and selects events in which the $m_{\tau\tau}^{\text{MMC}}$ invariant mass, the final discriminating variable, is estimated with improved resolution. The p_{T}^{H} is kinematically correlated with the $\Delta R_{\tau\tau}$ between the tau candidates, as shown in Figure A.3, so that the "High-p_{T}^{H}" region effectively selects events at low $\Delta R_{\tau\tau}$. Nevertheless, an explicit cut on $\Delta R_{\tau\tau}$ is needed for the suppression of background events at high-$\Delta R_{\tau\tau}$ that contaminate the sensitive $m_{\tau\tau}^{\text{MMC}}$ range.

Among the expected signal events in the VBF High-p_{T}^{H} category, roughly 72 % are produced via the VBF mechanism and the rest via the ggF mechanism. The other VBF-like events with low p_{T}^{H} are further categorised by the tighter selection on the VBF-tagged jets shown in the contour plot in Figure A.4. The diagonal cut in the $(|\Delta\eta_{jj}|, M_{jj})$ plane, $M_{jj}[\text{ GeV}] > -250|\Delta\eta_{jj}| + 1550$, identifies the tight region with a high signal to background ratio and the loose region with a low signal to background ratio. These two regions define the *VBF Low-p_T^H Tight* and the *VBF Low-p_T^H Loose* categories where the dominant source of background are multijet events and the fractions of VBF signal are 77 % and 56%, respectively. The use of the diagonal cut on the $(|\Delta\eta_{jj}|, M_{jj})$ plane is intended to explicitly take advantage of the correlation of these two variables in signal events, which allows the definition of regions with significantly different signal to background ratios. This completes the definition of the VBF categories used by the cut-based analysis. Events are then tested for inclusion in the boosted categories of the cut-based analysis.

The boosted category selection is intended to select events produced via the ggF mechanism where the Higgs boson, recoiling from a hard jet, has high transverse momentum. Boosted events are selected with $p_{\text{T}}^{\text{H}} > 100$ GeV, removing most of the multijet background as shown in Figure 4.3. This completes the definition of the boosted category in the multivariate analysis. The cut-based analysis continues by creating two separate boosted categories based on the same selection on $p_{\mathrm{T}}^{\mathrm{H}}$ and $\Delta R_{\tau\tau}$ used in the VBF High-$p_{\mathrm{T}}^{\mathrm{H}}$ category, since the correlations among these variables in the VBF-like and boosted events are very similar, as shown in the contour plots in Figures A.5-A.7.

In the multivariate analysis, events not contained in the VBF or boosted categories are included in the so-called rest category. These events are used to help further constrain the normalizations of the $Z \to \tau\tau$ and multijet backgrounds with a fit of the $|\Delta\eta_{\tau\tau}|$ distribution, as described in Section 5.4.

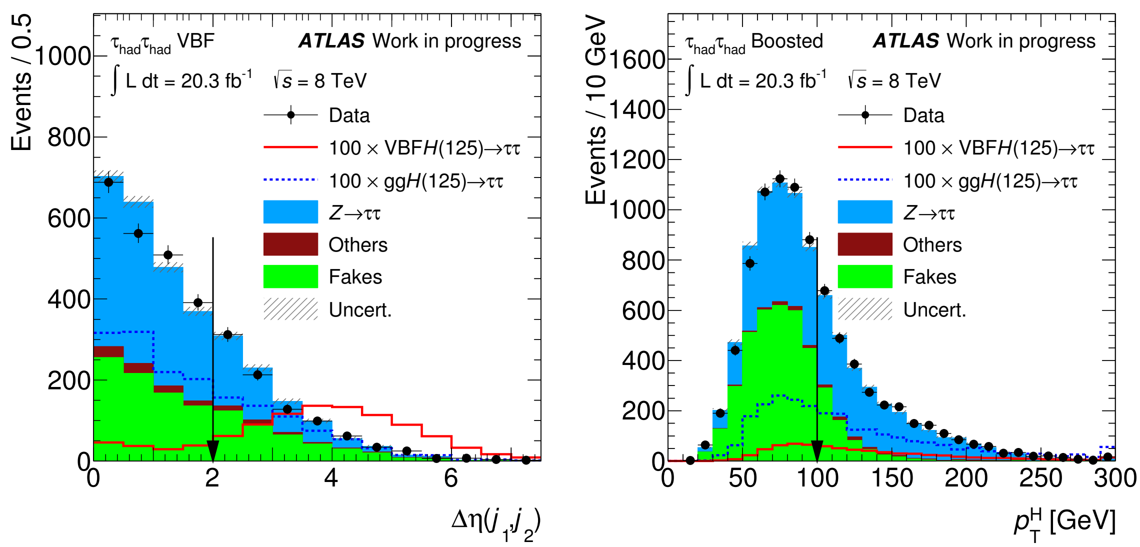


Figure 4.3: Left: The cut on $|\Delta\eta(\mathrm{jet}_1, \mathrm{jet}_2)|$ (shown here for the multivariate analysis) serves to separate the VBF from the ggF signal events. Right: The cut on $p_{\mathrm{T}}^{\mathrm{H}}$ in the boosted category removes most of the multijet background and the signal is dominated by the ggF process.

Category	Requirements
VBF	▽ $\lvert\Delta\eta_{\tau\tau}\rvert < 1.5$ ▽ Leading and subleading jets with $p_{\mathrm{T}} > 50, 30$ GeV respectively ▷ $\lvert\Delta\eta(\mathrm{jet}_1, \mathrm{jet}_2)\rvert > 2 \rightarrow$ MVA VBF ▽ $\lvert\Delta\eta(\mathrm{jet}_1, \mathrm{jet}_2)\rvert > 2.6$ and $M_{\mathrm{jj}} > 250$ GeV ▽ $\min(\eta_{\mathrm{jet},1}, \eta_{\mathrm{jet},2}) < \eta_{\tau,1}, \eta_{\tau,2} < \max(\eta_{\mathrm{jet},1}, \eta_{\mathrm{jet},2})$ ▷ Cut-based VBF Preselection (see contours in Appendix A.1) ▷ $\Delta R_{\tau\tau} < 1.5$ and $p_{\mathrm{T}}^{H} > 140$ GeV $\rightarrow$ Cut-based VBF High-p_{T}^{H} ▽ $\Delta R_{\tau\tau} > 1.5$ or $p_{\mathrm{T}}^{H} < 140$ GeV ▷ $M_{jj}[\text{ GeV}] > -250\lvert\Delta\eta_{jj}\rvert + 1550 \rightarrow$ Cut-based VBF Low-p_{T}^{H} Tight ▷ $M_{jj}[\text{ GeV}] < -250\lvert\Delta\eta_{jj}\rvert + 1550 \rightarrow$ Cut-based VBF Low-p_{T}^{H} Loose
Boosted	▽ $\lvert\Delta\eta_{\tau\tau}\rvert < 1.5$ ▷ $p_{\mathrm{T}}^{H} > 100$ GeV $\rightarrow$ MVA Boosted ▷ Cut-based Boosted Preselection (see contours in Appendix A.1) ▷ $\Delta R_{\tau\tau} < 1.5$ and $p_{\mathrm{T}}^{H} > 140$ GeV $\rightarrow$ Cut-based Boosted High-p_{T}^{H} ▷ $\Delta R_{\tau\tau} > 1.5$ or $p_{\mathrm{T}}^{H} < 140$ GeV $\rightarrow$ Cut-based Boosted Low-p_{T}^{H}
Rest	▽ $\lvert\Delta\eta_{\tau\tau}\rvert < 1.5$ ▷ Events used by the MVA Rest category of the fit model to constrain the $Z \rightarrow \tau\tau$ and multijet backgrounds using the shape of the $\Delta\eta_{\tau\tau}$ distribution.

Table 4.3: Summary of the event categorization in the multivariate and cut-based analyses. A flowchart of the categorization is also illustrated in Figure 4.4.

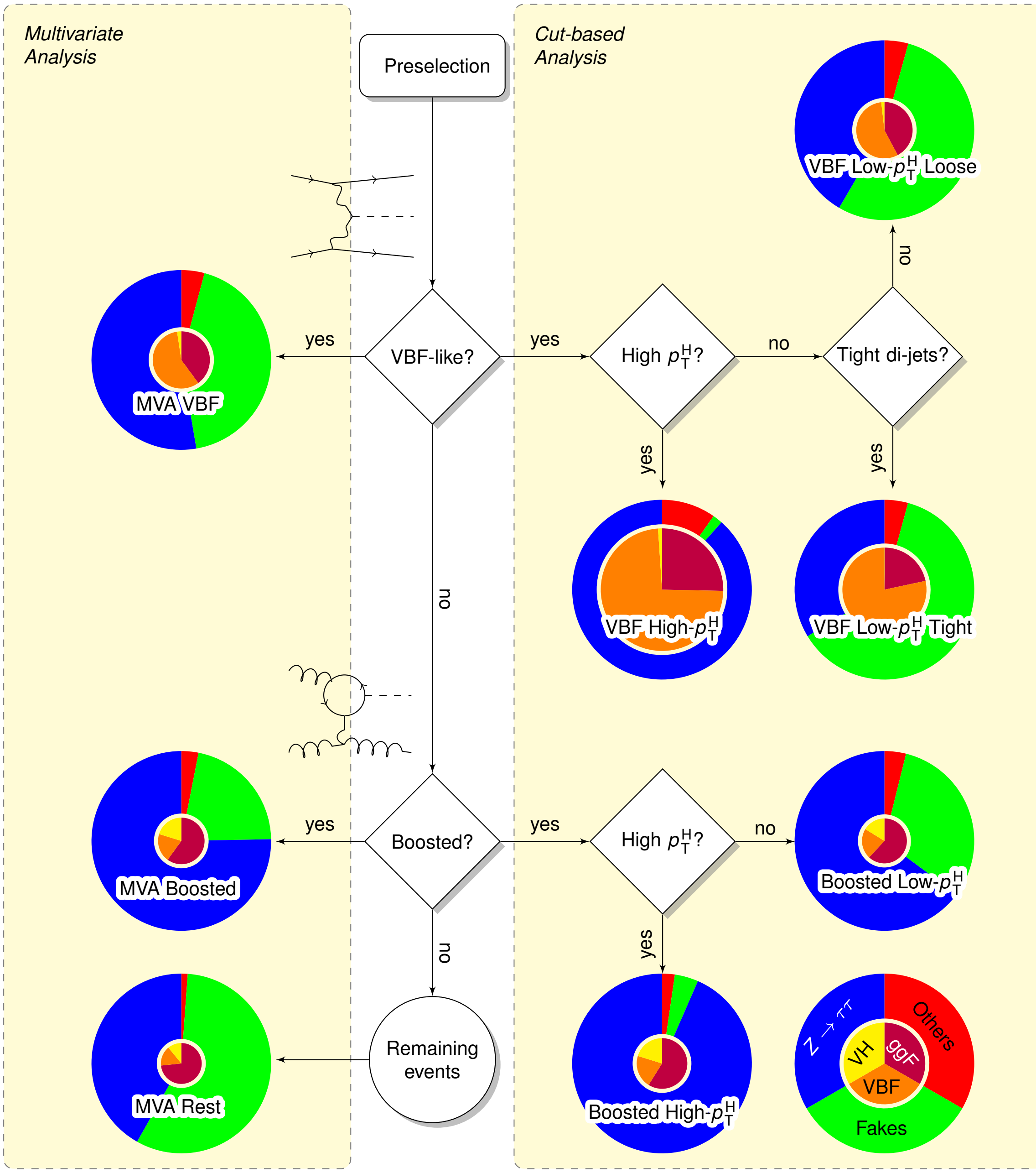


Figure 4.4: Flowchart of the event selection and categories in the multivariate and cut-based analyses. The inner pie charts represent the expected signal yield fractions from VBF, ggF and VH production modes for the SM Higgs boson with m_H = 125 GeV. The outer pie charts represent the yield fractions for the $Z \to \tau\tau$background, the fake tau background from multijet events, and the other minor backgrounds. The ratios of the inner and outer pie chart radii are proportional to the signal to background ratio in each category.

$\sqrt{s}$ = 8 TeV $\int$ Ldt = 20.3 fb^{-1} m$_H$ = 125 GeV

	VBF	**Boosted**	**Rest**
VH	0.29 ± 0.02	4.48 ± 0.07	1.61 ± 0.04
VBF	8.60 ± 0.07	4.48 ± 0.05	2.76 ± 0.04
ggF	6.0 ± 0.2	14.1 ± 0.3	12.2 ± 0.3
Total Signal	14.9 ± 0.2	23.1 ± 0.3	16.5 ± 0.3
Z$\rightarrow \tau\tau$	479.1 ± 11.8	2230.2 ± 26.2	2376.6 ± 23.6
Multijet	392.4 ± 12.2	640.0 ± 16.1	3242.6 ± 34.7
Others	37.9 ± 3.0	91.3 ± 5.3	65.0 ± 5.4
Total Background	909.4 ± 17.2	2961.5 ± 31.2	5684.3 ± 42.3
Data	892	3020	5648

Table 4.4: Summary of the observed and pre-fit expected event yields in each category of the multivariate analysis. The yields of $Z \rightarrow \tau\tau$ and the multijet backgrounds are estimated from data, as described in Chapter 5. Background and signal yields and uncertainties are before the global fit described in Chapter 6. Uncertainties are statistical only. Systematics will be introduced in Chapter 5.

Chapter 5

Event Model

Given the rarity of the Higgs signal process, accurate models of all backgrounds are essential. This analysis relies on data-driven background models where possible to minimize the dependence on simulation and associated systematic uncertainties. In the $\tau_{\text{had}}\tau_{\text{had}}$ final state, the dominant backgrounds are the mainly irreducible $Z \to \tau\tau$ background with two real hadronically decaying taus and a topology very similar to the signal processes, and the reducible background from multijet production in which two jets fake hadronically decaying taus. The $Z \to \tau\tau$ model is derived from selected $Z \to \mu\mu$ data events where the muons have been replaced by simulated tau leptons. Two multijet background models are derived from data, where one is used as the nominal model and the other provides an estimate of the systematic uncertainty on the shape and normalization. Minor backgrounds from W, diboson and top quark production are modelled by Monte Carlo (MC) simulation and are dominated by $W \to \tau\nu$ + *jets*, in which the second tau candidate is faked by a jet. This chapter presents the details of all background and signal models as well as their associated systematic uncertainties.

The target region in all categories is the isolated opposite-sign (OS) region. Events are defined as OS if the product of the tau charges is -1 ($q_1 \times q_2 = -1$) and if each tau has either one or three tracks. A track-based tau isolation is implemented by counting additional tracks within the exterior annulus of $0.2 < \Delta R < 0.6$ around the tau axis according to the procedure outlined in Section 3.5.2. A di-tau event is then classified as isolated if both taus have no additional tracks within the exterior annulus. Conversely, a di-tau event is classified as non-isolated if at least one of the taus contains an additional track in this exterior annulus.

$Z \rightarrow \tau\tau$ is modelled using embedded $Z \rightarrow \mu\mu$ data events directly in the isolated OS region, as described in Section 5.1. A data-driven technique is also used to estimate the fake di-tau background from multijet processes but from a control region that is orthogonal to the isolated OS region, as described in Section 5.2.

5.1 Embedded $Z \rightarrow \tau\tau$ Model

The $Z \rightarrow \tau\tau$ background would ideally be modelled by selecting these events from data. Identifying hadronic tau decays, however, is significantly more difficult than identifying muons or electrons, making it impossible to obtain a pure and sufficiently large sample in ATLAS data. It would also be impossible to completely exclude signal events from such a sample. On the contrary, a large and pure sample of $Z \rightarrow \mu\mu$ events can be selected from data and these events are kinematically identical to $Z \rightarrow \tau\tau$ events, except for effects due to the difference between the muon and tau mass. The Higgs coupling to muons is also much smaller than taus, leading to negligible signal contamination in such a sample. $Z \rightarrow \mu\mu$ events in data are therefore used to estimate the $Z \rightarrow \tau\tau$ background by replacing the tracks and associated calorimeter cells of the identified muons with simulated hadronically decaying tau leptons. The two tau leptons are simulated by `TAUOLA` [40] while being fixed to the kinematics of the data muons they replace and accounting for the $\mu - \tau$ mass difference. This procedure is known as embedding [41] and the hybrid sample it produces will be referred to as embedded $Z \rightarrow \tau\tau$.

In the embedded $Z \rightarrow \tau\tau$ sample, only the tau decays and corresponding detector response are modelled by simulation and all other properties are taken from data. These properties include the underlying event kinematics, pile-up effects, the kinematics and topology of additional jets, an improved model of the $E_{\mathrm{T}}^{\mathrm{miss}}$, and the mixture of electroweak and QCD Z production. Using the additional jets from data benefits the analysis when exploiting VBF and boosted event topologies, as it is no longer sensitive to associated systematic effects from simulation and the topologies are exactly as they are in data.

The description of the embedding method is documented in Ref. [42]. $Z \rightarrow \mu\mu$ events are first selected in data by single and di-muon triggers. These events are required to have exactly two well-reconstructed muons with leading and subleading transverse momenta of $p_{\mathrm{T}} > 20$ GeV and $p_{\mathrm{T}} > 15$ GeV. A track isolation of $\Sigma_{\mathrm{tracks}}^{\Delta R<0.2} p_{\mathrm{T,track}}/p_{\mathrm{T}}^{\mu} < 0.2$ is applied and the muons must have a common primary vertex. The reconstructed di-muon invariant mass must be at least 40 GeV. The embedding procedure of the di-muon events starts with the reconstruction of the Z boson and the simulation with `TAUOLA` and `PHOTOS` [43] of its decay into tau leptons accounting for the mass difference between m_{τ} and m_{μ}. The simulated final state is then processed through the simulation of the ATLAS detector (top right of Figure 5.1). In the original $Z \to \mu\mu$ data event (top left of Figure 5.1), the two muons are removed together with all the associated tracks and energy deposits in the calorimeter. The subtraction of the energy deposits is based on the amount of energy left by the muons in simulated $Z \to \mu\mu$ events. The event without muons is then merged with the simulated $Z \to \tau_{\mathrm{had}}\tau_{\mathrm{had}}$ and the resulting embedded event (bottom of Figure 5.1) is processed through the full event reconstruction so that all objects, including $E_{\mathrm{T}}^{\mathrm{miss}}$, are recreated based on the modified energy deposits and tracks.

The embedded $Z \to \tau\tau$ sample must undergo a series of corrections to ensure that the kinematic spectrum of the embedded tau leptons accurately matches the spectrum that would be given by true $Z \to \tau\tau$ decays. An initial set of corrections cancels the bias from the muon trigger and reconstruction efficiencies. Reciprocals of these efficiencies as a function of p_{T} and η are applied as weights to each event. Additionally, the tau trigger is not simulated in the embedded $Z \to \tau\tau$ sample and so the product of the trigger efficiencies for each embedded tau lepton is applied as an event weight.

The tau polarization depends on the initial state quark configuration, which is of course not known in data and therefore not directly taken into account in the original embedded $Z \to \tau\tau$ sample. Although the effect on the analysis is expected to be small, the tau polarization and spin correlations can be emulated with the `TauSpinner` [44] package. This produces and a weight that is applied to each event and achieves the correct polarization and spin correlations.

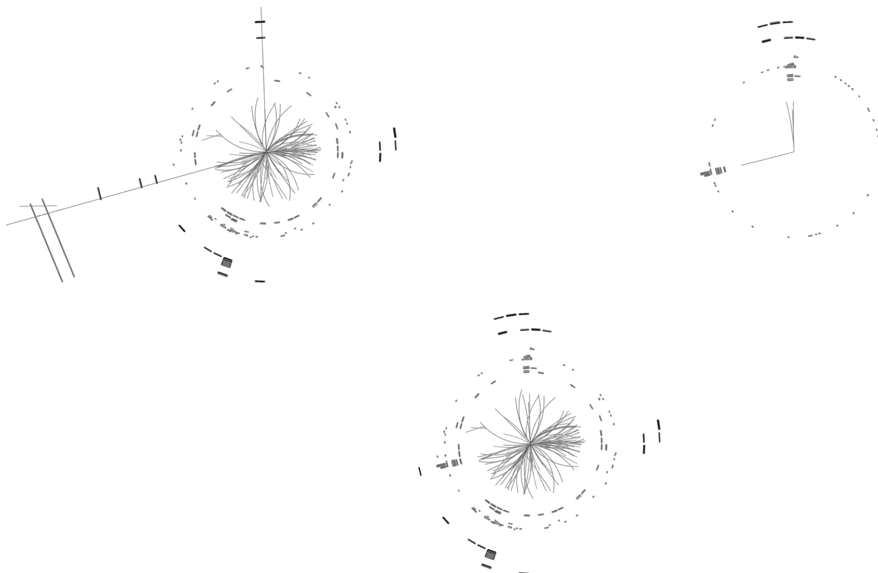


Figure 5.1: Event displays of a di-muon event selected in data (top left), a simulated $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ event (top right) and the resulting embedded $Z \rightarrow \tau_{\text{had}}\tau_{\text{had}}$event (bottom).

5.2 Multijet Background Model

The fake di-tau background from multijet processes is estimated from a control region that is orthogonal to the isolated OS region. This orthogonal control region can be constructed by inverting the OS charge requirement, inverting the isolation requirement, or inverting both.

When inverting the OS requirement, two options are considered. One option is using the same-sign (SS) requirement defined as $q_1 \times q_2 = +1$ where each tau is required to have either one or three tracks. A second option is simply inverting the equality in the OS requirement: $q_1 \times q_2 \neq -1$. This second option is called the not-opposite-sign (nOS) region and tau candidates are allowed to have any non-zero number of tracks. Note that the SS requirement selects a subset of the nOS events.

With the above options for inverting the target region requirements, the following control regions in data are considered:

- SS (with no isolation requirement)
- isolated SS
- non-isolated SS
- nOS (with no isolation requirement)
- isolated nOS
- non-isolated nOS
- non-isolated (with no charge requirement)
- non-isolated OS

For each of the above control regions (m), the multijet background shape (Fakes_m) is estimated at preselection or in any category as:

$$\mathrm{Fakes}_m = \mathrm{Data}_m - b_m \times (Z \to \tau\tau)_m - \mathrm{Others}_m \tag{5.1}$$

Contributions from $Z \to \tau\tau$ and the other minor backgrounds (described in the next section) in this control region are subtracted to remove double-counting of such events. The free parameter b_m is from the normalization of $Z \to \tau\tau$ in the isolated OS region and will be discussed in Section 5.4. The selection of the nominal multijet model and the systematic uncertainty will be presented in Section 5.5 after introducing the background normalizations.

5.3 Other Backgrounds

Minor backgrounds from W/Z+jets, top production, and diboson processes are estimated by MC simulation, where at least one real hadronically decaying tau must be included in the final state. Since these sources of background events contribute only 1-4% of the total expected background yield, as shown in Table 4.4, they have been combined and treated as one sample called “Others” in the background model. The dominant contribution is from $W \to \tau\nu$+jets where reconstructed tau is real and another is faked by an additional jet. The MC generators used for each of these minor backgrounds are described below in Section 5.6. The multijet background model subtracts contamination from this Others sample in Equation 5.1 so background events with two fake tau candidates are naturally included by the data-driven multijet background model.

To properly model the simulated electroweak background events, the probability of mis-identifying a jet as a tau candidate is corrected to match data in a $Z \rightarrow \mu\mu$ control sample. These correction scale factors were determined specifically for this analysis, unlike the tau efficiency scale factors described in Section 3.5.2, as they depend on the specific background composition. A $Z \rightarrow \mu\mu$ control sample is obtained by requiring two isolated muons with an invariant mass between 70$<$ $m_{\mu\mu}$ $<$110 GeV. An additional jet is then required that must satisfy the tau candidate selection. The dominant uncertainty on this measurement is statistical ($\sim$20%). The mis-identification probability is parameterized according to the offline tau identification criteria for one and three-track candidates, the trigger, and data period. This correction is applied in the electroweak MC simulation samples to any reconstructed tau that does not match a generator-level true tau object within ΔR$<$0.2. In principle, quark or gluon-initiated jets result in different mis-identification probabilities and the difference should be taken into account. In this analysis, however, the $Z \rightarrow \mu\mu$ control sample provides a similar quark/gluon ratio to our electroweak backgrounds. Any difference in quark/gluon-initiated jet fractions in different electroweak backgrounds is smaller than the current statistical uncertainty on this measurement.

5.4 Background Normalization

Following the definition of the multijet background in Equation 5.1, the full background model in the isolated OS region is then estimated as:

$$\text{Background}_{\text{Isol. OS}} = a_m \times \text{Fakes}_m + b_m \times (Z \rightarrow \tau\tau)_{\text{Isol. OS}} + \text{Others}_{\text{Isol. OS}} \tag{5.2}$$

The normalizations of the top, diboson, and other electroweak backgrounds, combined in the term $\text{Others}_{\text{Isol. OS}}$, are fixed by their cross sections and observed luminosity in the data. The normalizations of the $Z \rightarrow \tau\tau$ and multijet backgrounds (a_m and b_m) are free parameters and must be constrained by a fit of some variable where these backgrounds have different shapes. This variable should also carry minimal discrimination between

$Z \to \tau\tau$ and $H \to \tau\tau$ and be performed in a region with a minimal expected contribution from $H \to \tau\tau$. The signal contamination at preselection and in the rest category is 0.56% and 0.29%, respectively, assuming a SM Higgs boson.

The $\Delta\eta(\tau,\tau)$ is a simple quantity that provides sufficient discrimination between the resonant $Z \to \tau\tau$ and non-resonant multijet background, as shown in Figure 5.2. Unlike other potential distributions such as the number of tau tracks or tau-id BDT score, the $\Delta\eta(\tau,\tau)$ distribution does not depend on the track composition of each tau and provides more flexibility in the track composition of the multijet model, as is required to use the nOS model. The two-dimensional recounted tracks distributions of the leading and subleading taus at 7 TeV and 8 TeV are shown in Figure 5.3. A bias toward 1-prong tau candidates at the trigger level in 8 TeV reduced the quality of a fit of this distribution when normalizing the $Z \to \tau\tau$ and multijet backgrounds. This also contributed to the decision to instead use a fit of $\Delta\eta(\tau,\tau)$.

For the purpose of initializing the background model and creating validation plots, a maximum likelihood fit of $\Delta\eta(\tau,\tau)$ is performed at preselection. A signal template is not included in this fit. As in Equation 5.1, the subtracted component from $Z \to \tau\tau$ in the control region depends on the normalization of $Z \to \tau\tau$ through the b_m parameter. Due to software limitations, the subtraction in the multijet background template is fixed (b_m only in Equation 5.1) during the fit. So as the normalizations a_m and b_m in Equation 5.2 are determined after a fit, b_m in Equation 5.1 is then updated and the fit is repeated (updating b_m in Equation 5.1 each time) until the values of a_m and b_m converge. Convergence within a threshold of 10^{-5} typically occurs within 3 to 5 iterations. Before the fitting begins, a_m and b_m are initialized such that $Z \to \tau\tau$ and the multijet backgrounds are each normalized to 50% of the data. An example of a post-fit $\Delta\eta(\tau,\tau)$ distribution is shown in Figure 5.4.

This fit at preselection acts only as a starting point for the normalizations since they are again set free in the final global fit containing the BDT distributions in the boosted and VBF categories and the $\Delta\eta(\tau,\tau)$ distribution in the rest category.

The $\Delta\eta(\tau,\tau)$ fit has also been tested in each category, yielding consistent results, as shown in Figure 5.5. Alternate distributions have also been fitted and compared with the $\Delta\eta(\tau,\tau)$ fit, as shown in Figure 5.6.

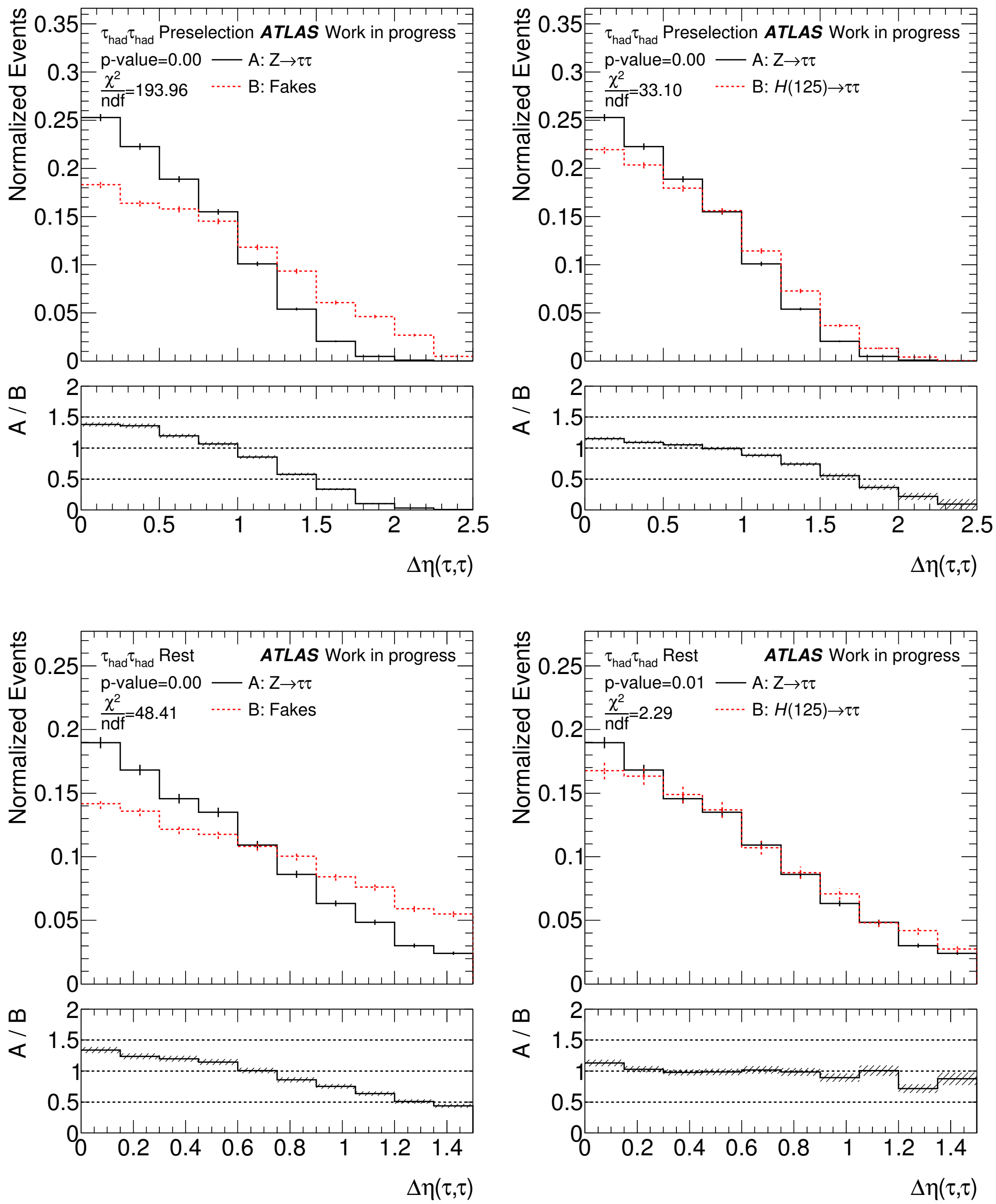
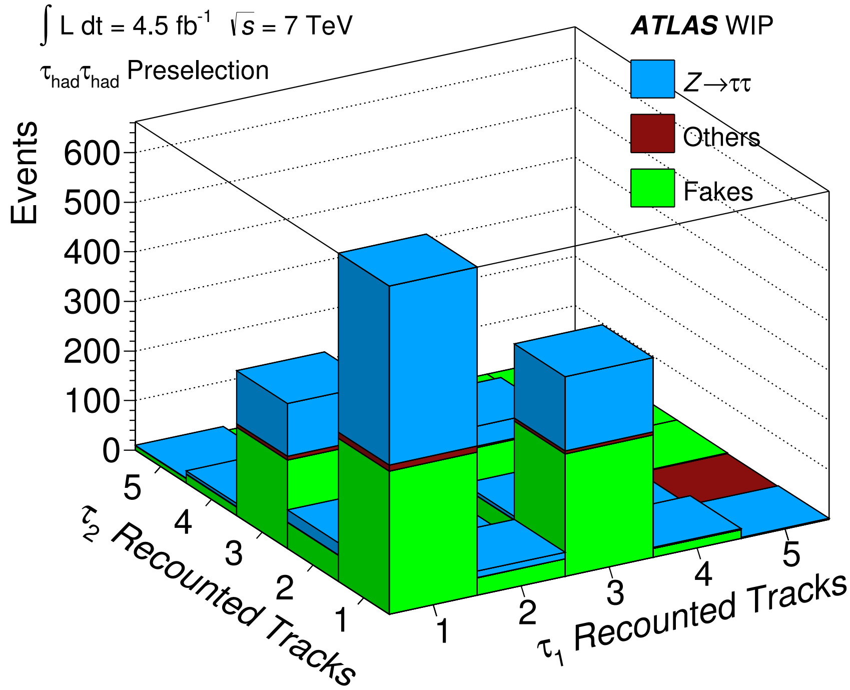


Figure 5.2: Comparisons of the $Z \to \tau\tau$, $H \to \tau\tau$, and multijet $\Delta\eta(\tau,\tau)$ shapes at preselection (top) and in the rest category (bottom).



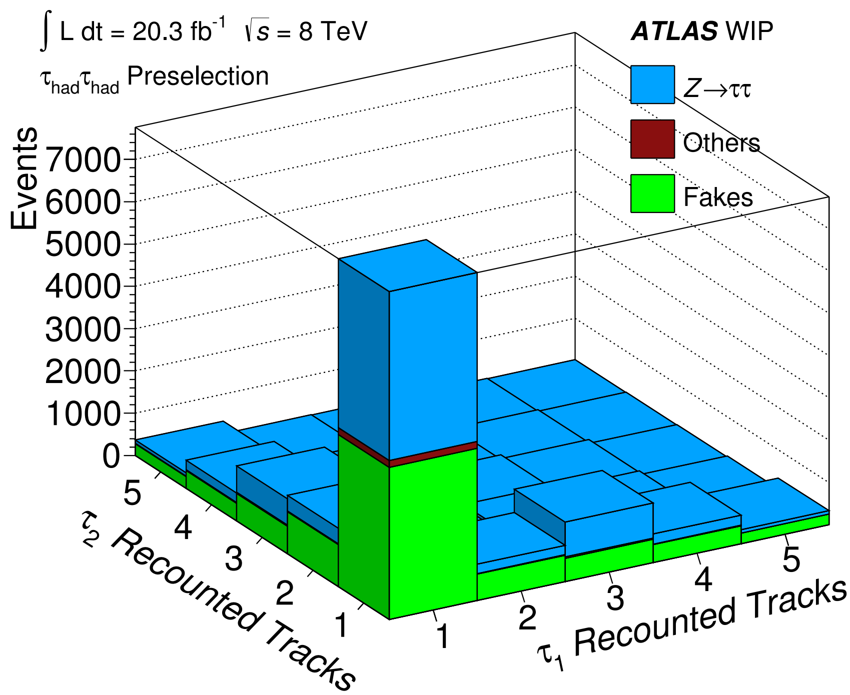


Figure 5.3: The two-dimensional recounted tracks distributions of the leading and subleading taus at 7 TeV (top) and 8 TeV (bottom).

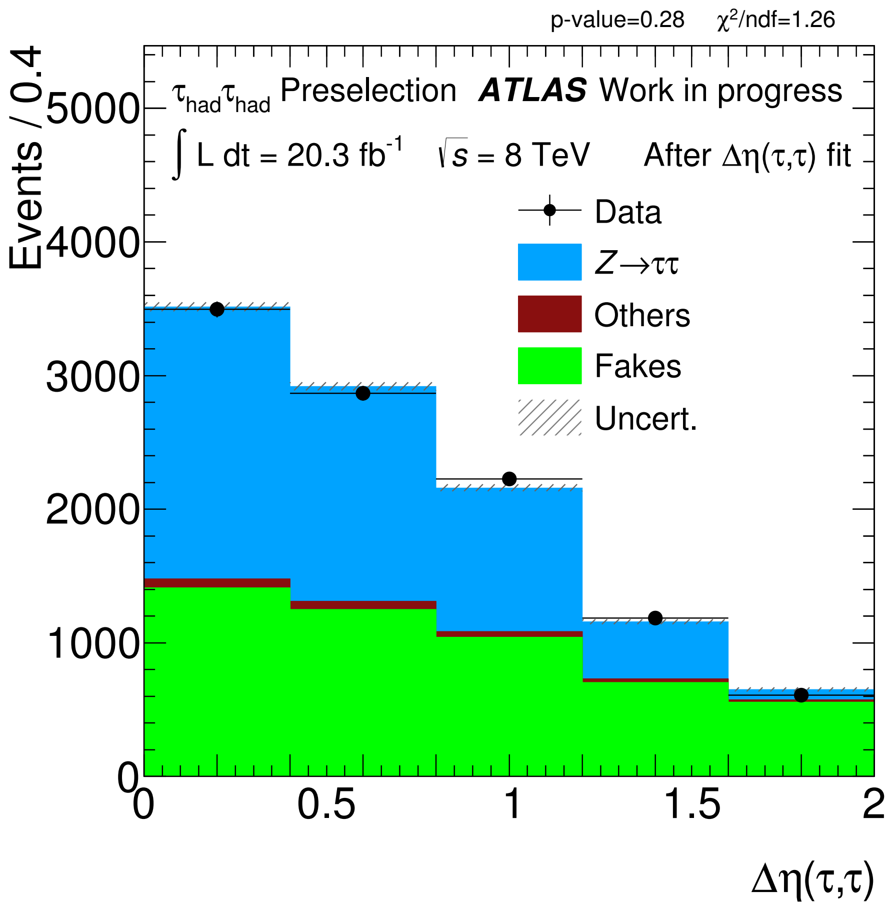


Figure 5.4: Post-fit $\Delta\eta(\tau,\tau)$ distribution at preselection.

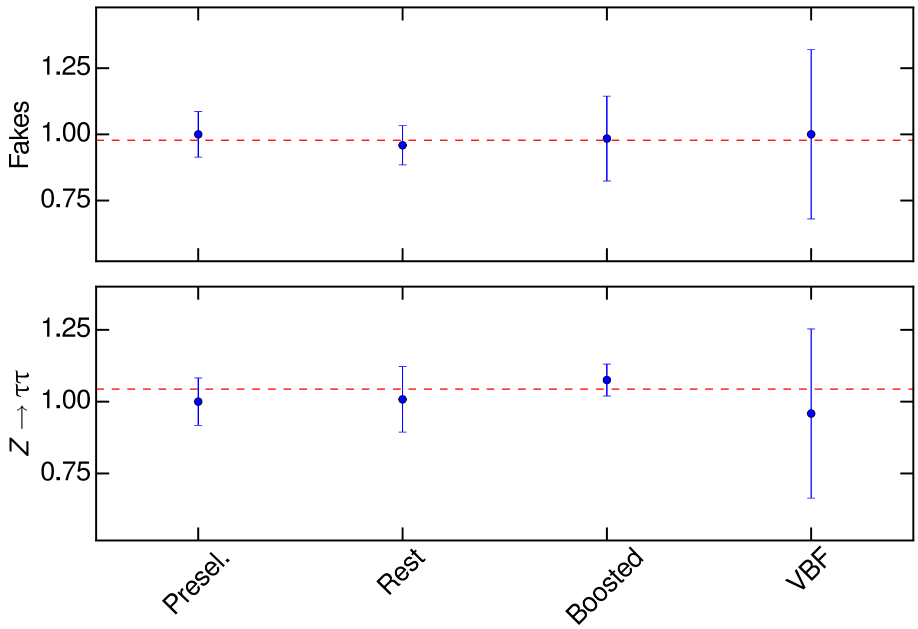


Figure 5.5: The $\Delta\eta(\tau,\tau)$ fit is performed in each category to show the consistency of the best-fit normalizations.

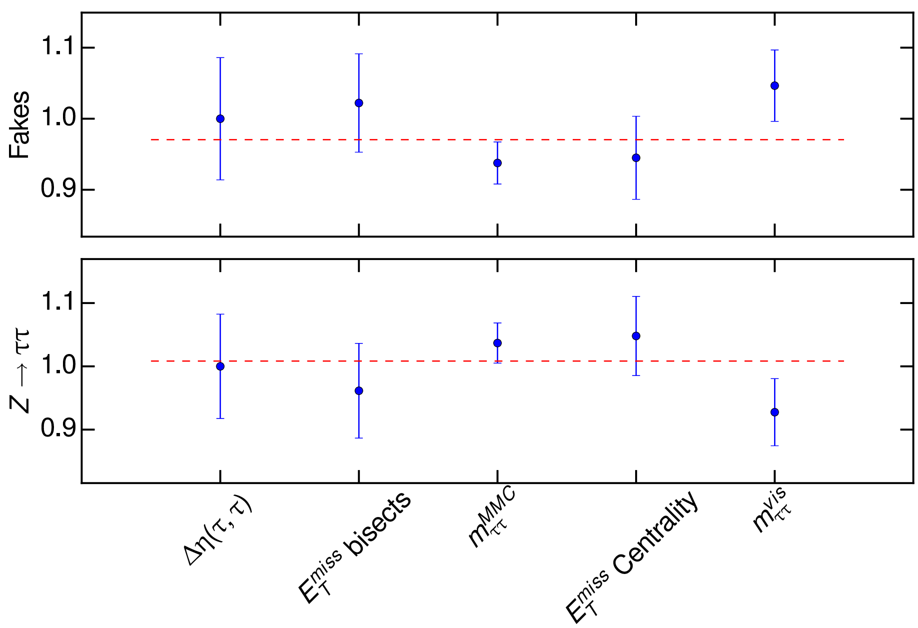


Figure 5.6: Alternate distributions are fit at preselection and compared with the $\Delta\eta(\tau,\tau)$ fit.

5.5 Selection of the Multijet Background Model

Of all models listed in Section 5.2, one multijet background model is selected as the nominal model and another is selected to represent a systematic uncertainty on both the yield and shape of the multijet background in each category. The nominal model should ideally have the best agreement with data while containing a sufficiently large number of unweighted events. The signal contamination in the nominal should also be minimal. The unweighted number of events and the signal contamination in each model is shown in Table 5.1 and Table 5.2, respectively. Figure 5.8 demonstrates the consistency of the $\Delta\eta(\tau,\tau)$ fit for each multijet background model.

The higher signal contamination in the non-isolated OS and non-isolated models (of which non-isolated OS is a subset) prevents them from being a sensible choice as a nominal multijet background model, although the non-isolated model is used to train the BDTs since it contains the highest number of events. As shown in Figure 5.7, the SS models typically have the worst agreement at low $m_{\tau\tau}^{\mathrm{MMC}}$. The nOS models contain significantly more events and provide a better model of $m_{\tau\tau}^{\mathrm{MMC}}$ and $\Delta R(\tau\tau)$. The non-isolated nOS model is used as the nominal multijet background model because of the generally better agreement with data at low $m_{\tau\tau}^{\mathrm{MMC}}$ and across $\Delta R(\tau\tau)$. The isolation requirement is then inverted with the isolated nOS model to provide an estimate of the systematic uncertainty.

The shape and normalization systematic uncertainty on the non-isolated nOS model is determined by the symmetrized difference between the two models after setting the normalization of the isolated nOS model equal to the normalization of the non-isolated nOS model at preselection. The direction of the difference in normalization between the models can then differ as each model is propagated into each category. This is expressed as an overall normalization and shape systematic on the multijet background model in each category.

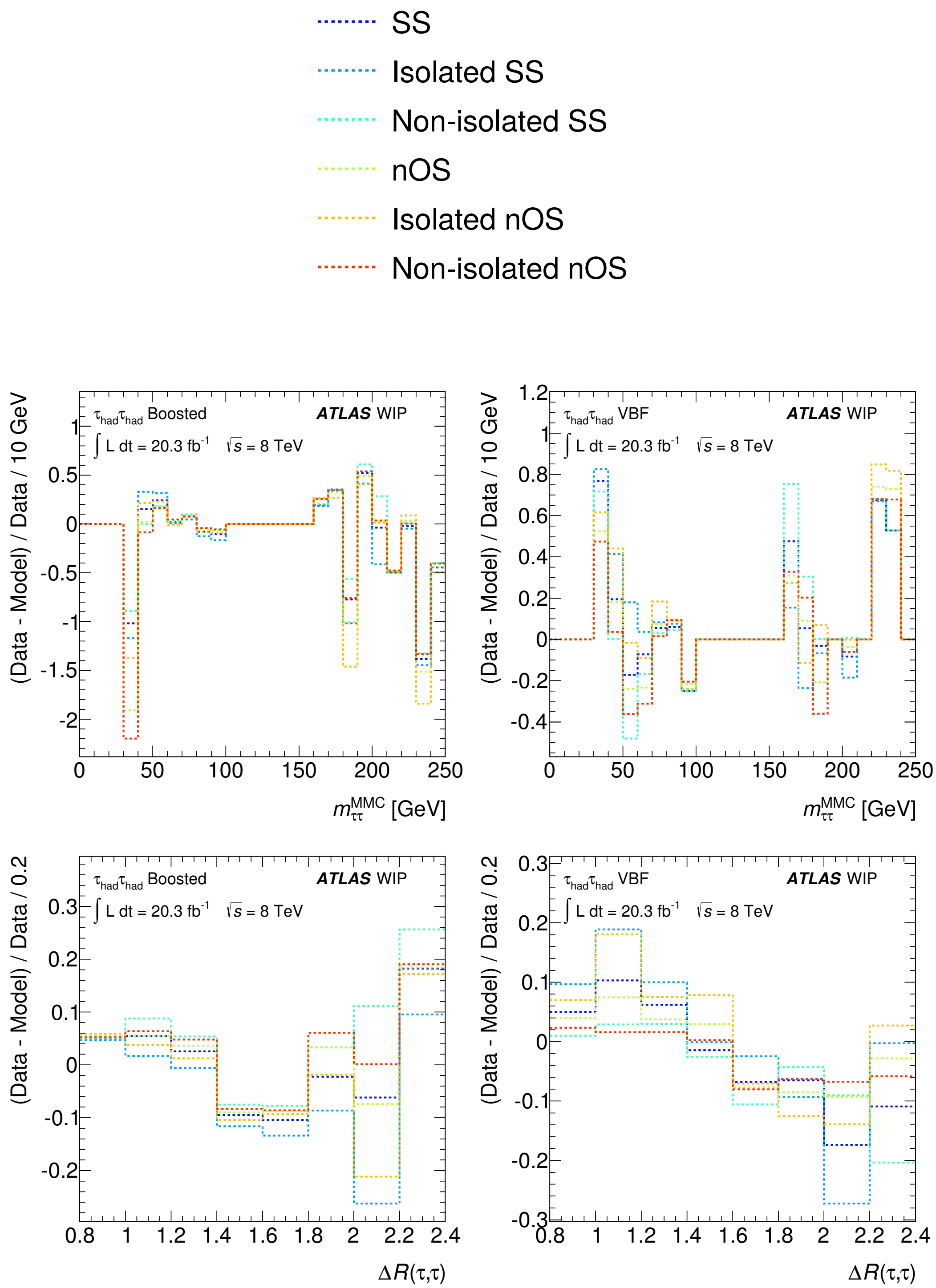


Figure 5.7: Comparison of the multijet background models using the $\Delta R(\tau\tau)$ (top) and $m_{\tau\tau}^{\mathrm{MMC}}$ (bottom) distributions in the boosted (left) and VBF (right) categories.

Model	Presel.	Rest	Boosted	VBF
SS	7995	4995	1214	658
Isolated SS	3630	2229	600	275
Non-isolated SS	4365	2766	614	383
nOS	29474	17854	5600	2536
Isolated nOS	11646	6755	2560	1013
Non-isolated nOS	17828	11099	3040	1523
Non-isolated	35688	21024	7941	3236
Non-isolated OS	17829	9907	4892	1710

Table 5.1: Unweighted number of events in each multijet background model.

Model	Presel.	Rest	Boosted	VBF
SS	0.1 ± 0.0	0.1 ± 0.0	0.3 ± 0.0	0.3 ± 0.0
Isolated SS	0.2 ± 0.0	0.1 ± 0.0	0.5 ± 0.0	0.6 ± 0.1
Non-isolated SS	0.1 ± 0.0	0.0 ± 0.0	0.2 ± 0.0	0.1 ± 0.0
nOS	0.4 ± 0.0	0.2 ± 0.0	1.1 ± 0.0	1.3 ± 0.0
Isolated nOS	0.7 ± 0.0	0.3 ± 0.0	2.0 ± 0.1	2.5 ± 0.1
Non-isolated nOS	0.3 ± 0.0	0.1 ± 0.0	0.7 ± 0.0	0.8 ± 0.0
Non-isolated	1.1 ± 0.0	0.6 ± 0.0	3.1 ± 0.1	3.3 ± 0.1
Non-isolated OS	3.5 ± 0.1	1.9 ± 0.0	9.4 ± 0.6	10.3 ± 0.7

Table 5.2: Signal contamination [%] in each multijet background model assuming a SM Higgs boson.

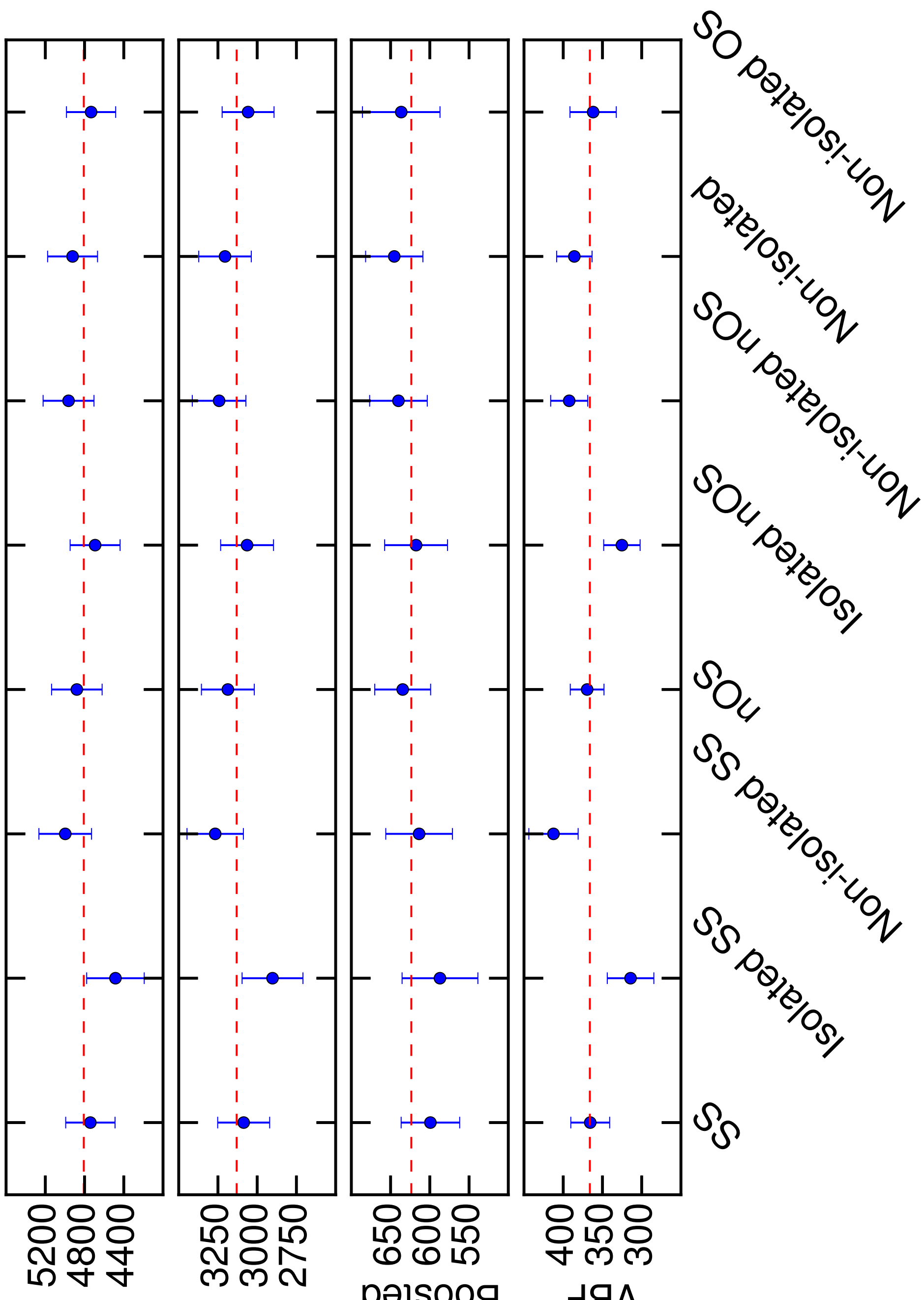


Figure 5.8: Consistency of the weighted number of multijet events using a $\Delta\eta(\tau,\tau)$ fit of each multijet background model at preselection.

5.6 Monte Carlo Simulation

Signal events and the minor background events in the *Others* sample were simulated using various Monte Carlo (MC) generators, as summarised in Table 5.3. The generators used for the simulation of the hard scattering process and the model used for the simulation of the parton shower, of the hadronisation and of the underlying event activity are listed. In addition, the cross-section values to which the simulation is normalised and the perturbative order in QCD of the respective calculations are given.

The gluon fusion and the *VBF* Higgs production are simulated with POWHEG [45–48] interfaced to PYTHIA8 [49] providing the parton shower. In the POWHEG event generator the CT10 [50] parametrisation of the PDFs is used. The overall normalisation of the *ggF* process is taken from a calculation at next-to-next-to-leading order (NNLO) [51–56] in QCD, including soft-gluon resummation up to the order of next-to-next-to-leading logarithm (NNLL) [57]. Next-to-leading order (NLO) electroweak (EW) corrections are also included [58, 59]. The *VBF* production is normalised to a cross section calculated with full NLO QCD and EW corrections [60–62] with an approximate NNLO QCD correction applied [63]. The associated *VH* production process is simulated with PYTHIA8. The CTEQ6L1 [64] parametrisation of PDFs is used for the PYTHIA8 event generator. The predictions for *VH* production are normalised to cross sections calculated at NNLO in QCD [65], with NLO EW radiative corrections [66] applied.

Additional corrections to the shape of the generated p_{T} distribution of Higgs bosons produced via gluon fusion are applied to match the distribution from a calculation at NNLO including the NNLL corrections provided by the HRES2.1 [67] program. In this calculation, the effects of finite masses of the top and bottom quarks [67, 68] are included and dynamical renormalisation and factorisation scales, $\mu_{\mathrm{R}}, \mu_{\mathrm{F}} = \sqrt{m_H^2 + p_{\mathrm{T}}^2}$, are used. A reweighting is performed separately for events with less than or equal to one jet at particle level and for events with two or more jets. In the latter case, the Higgs boson p_{T} spectrum is reweighted to match the MINLO HJJ predictions [69]. The reweighting is derived such that the inclusive Higgs boson p_{T} spectrum and the p_{T} spectrum of events with at least two jets matches the HRES2.1 and MINLO HJJ predictions, respectively, and that the jet multiplicities are in agreement with (N)NLO calculations from JETVHETO [70–72].

The NLO EW corrections for the *VBF* production depend on the p_{T} of the Higgs boson, varying from a few percent at low p_{T} to $\sim$ 20% at p_{T} = 300 GeV [73]. The *VBF*-produced Higgs boson p_{T} spectrum is therefore reweighted, based on the difference between the POWHEG+PYTHIA and the HAWK [60, 61] calculation, which includes these corrections.

Other background processes are simulated using different generators, each interfaced to PYTHIA [49, 74] or HERWIG [75] to provide the parton shower, hadronisation and the modelling of the underlying event, as indicated in table 5.3. For the HERWIG samples, tau decays are simulated by TAUOLA [40]. PHOTOS [43] provides photon radiation from charged leptons for all samples. The samples for W/Z+jets production are generated with ALPGEN [76], employing the MLM matching scheme [77] between the hard process (calculated with LO matrix elements for up to five jets) and the parton shower. For *WW* production the loop-induced $gg \to WW$ process is also generated using the GG2WW [78] program. In the ACERMC [79], ALPGEN, and HERWIG event generators the CTEQ6L1 parametrisation of the PDFs is used, while the CT10 parametrisation is used for the generation of events with GG2WW.

For all samples, a full simulation of the ATLAS detector response [80] using the GEANT4 program [81] was performed. In addition, events from minimum bias interactions were simulated using the AU2 [82] tuning of PYTHIA8. They are overlaid on the signal and background simulated events according to the luminosity profile of the recorded data. The contributions from these pile-up interactions are simulated both within the same bunch crossing as the hard-scattering process and in neighbouring bunch crossings. Finally, the resulting simulated events are processed through the same reconstruction programs as the data. All MC simulation samples are then corrected to better reproduce the data, using official ATLAS recommendations. The events in these samples are produced with different pileup conditions that reproduce each period of data taking in 2012. In order to have a proper simulation of the pileup conditions, each period in the MC events are weighted to the integrated luminosity recorded in the corresponding period in data.

Signal (m_H = 125 GeV)	MC generator	$\sigma \times \mathcal{B}$ [pb] $\sqrt{s}$ = 8 TeV		
ggF, $H \to \tau\tau$	POWHEG [45–48] + PYTHIA8 [49]	1.22	NNLO+NNLL	[4, 51–56]
VBF, $H \to \tau\tau$	POWHEG + PYTHIA8	0.100	(N)NLO	[4, 60–62]
WH, $H \to \tau\tau$	PYTHIA8	0.0445	NNLO	[4, 65]
ZH, $H \to \tau\tau$	PYTHIA8	0.0262	NNLO	[4, 65]
Background	**MC generator**	**$\sigma \times \mathcal{B}$ [pb] $\sqrt{s}$ = 8 TeV**		
$W(\to \ell\nu)$, (ℓ = e, μ, τ)	ALPGEN [76]+PYTHIA8	36800	NNLO	[83, 84]
$Z/\gamma^*(\to \ell\ell)$, 60 GeV< $m_{\ell\ell}$ < 2 TeV	ALPGEN+PYTHIA8	3910	NNLO	[83, 84]
$Z/\gamma^*(\to \ell\ell)$, 10 GeV< $m_{\ell\ell}$ < 60 GeV	ALPGEN+HERWIG [75]	13000	NNLO	[83, 84]
VBF $Z/\gamma^*(\to \ell\ell)$	SHERPA [85]	1.1	LO	[85]
$t\bar{t}$	POWHEG + PYTHIA8	$253^\dagger$	NNLO+NNLL	[86–91]
Single top : Wt	POWHEG + PYTHIA8	$22^\dagger$	NNLO	[92]
Single top : s-channel	POWHEG + PYTHIA8	$5.6^\dagger$	NNLO	[93]
Single top : t-channel	AcerMC [79]+PYTHIA6 [74]	$87.8^\dagger$	NNLO	[94]
$q\bar{q} \to WW$	ALPGEN+HERWIG	$54^\dagger$	NLO	[95]
$gg \to WW$	gg2WW [78]+HERWIG	$1.4^\dagger$	NLO	[78]
WZ, ZZ	HERWIG	$30^\dagger$	NLO	[95]

Table 5.3: Monte Carlo generators used to model the signal and background processes at $\sqrt{s}$ = 8 TeV. The cross sections times branching fractions ($\sigma \times \mathcal{B}$) used for the normalisation of some processes (many of these are subsequently normalised to data) are included in the last column together with the QCD perturbative order of the calculation. For the signal processes the $H \to \tau\tau$ branching ratio is included, and for the W and Z/γ^* background processes the branching ratios for leptonic decays (ℓ = e, μ, τ) of the bosons are included. For all other background processes inclusive cross sections are quoted (marked with a $\dagger$).

5.7 Systematic Uncertainties

The signal and background models are affected by various systematic uncertainties that influence the estimation of compatibility between the signal and background models with the observed data. Sources of systematic uncertainties are discussed below, grouped into three categories: experimental uncertainties, background modelling uncertainties, and theoretical uncertainties. The effects on both the total signal and background yields and on the shape of any distribution distribution are evaluated for all uncertainties. A summary of the systematic uncertainties and their impact on the number of expected events for the signal and the total background is given in Table 5.4. In this table also the dominant sources that affect the shape of the BDT output distribution are marked. The inclusion of the uncertainties in the profile likelihood fit is described in Section 6.3. The effect on overall yields and distribution shapes from the various uncertainty components has been studied and only the relevant ones have been retained. This simplification, described in Section 6.5.1, is intended to avoid introducing noise in the final statistical analysis.

5.7.1 Experimental Uncertainties

The major experimental systematic uncertainties result from uncertainties on efficiencies for triggering, object reconstruction and identification, as well as from uncertainties on the energy scale and resolution of jets and hadronically decaying taus. In addition, uncertainties on the luminosity affect the number of expected simulated signal and background events.

Luminosity and Pile-up

The uncertainty on the 2012 integrated luminosity of 20.3 fb^{-1} is $\pm 2.8\%$. It was determined from a calibration of the luminosity scale described in Ref. [96].

The simulated samples are generated in advance of data-taking and therefore the exact profile and time evolution of the pile-up distribution is not known. After the data was recorded, the pile-up distribution in simulation was weighted to match data within an uncertainty of approximately 1%.

Tau Identification Efficiency

As described in Ref. [28] and Section 3.5.2, the MC simulation is corrected so that the tau identification efficiency matches what has been measured in the data. The tau identification efficiency is then varied within the uncertainties of $\pm$(2–3)% for 1-prong and $\pm$(3–5)% for 3-prong tau decays. These correction factors and the corresponding variations are applied only on truth-matched hadronic taus. The statistical uncertainty of the tag and probe measurement is treated as an independent parameter in this analysis.

The correction factors on the rate of misidentification of jets as hadronic tau candidates are also varied within their uncertainties as described in Section 5.3. These correction factors and the corresponding variations are applied only on hadronic tau candidates that do not match true tau decays.

Tau Trigger Efficiency

Trigger efficiencies are measured by comparing a selection of data and simulated events similar to that used to measure the tau identification efficiency as described in Section 3.5.2. Scale factors correcting the efficiency in simulation to what is observed in data are determined and are applied to truth-matched hadronic taus in simulated events. The tau trigger is not emulated in the embedded $Z \to \tau\tau$ sample and the trigger efficiency measured in data is applied as an event weight instead. The correction factors or efficiencies are then varied within their uncertainties. The statistical uncertainty of the tag and probe measurement is treated as an independent parameter.

Tau Energy Scale (TES)

The TES calculation and its uncertainty are described in Ref. [33] and in Section 3.5.3. The TES was calculated by using a mix of in-situ TES corrections (for the range $p_T <$ 50 GeV), obtained by fitting the reconstructed visible mass for $Z \to \tau\tau$ events in data, and a decomposition method (for $p_T >$ 70 GeV), with an interpolation performed in between. The TES is measured with a precision of $\pm 2 - 4\%$ [33]. Variations in the TES are also propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement and contribute to systematic effects there. The TES uncertainty is divided into four uncorrelated parts:

- *I*n-situ interpolation on true taus: uncertainty of the in-situ energy scale, relevant mainly for $p_T < 70$ GeV.
- *S*ingle particle interpolation on true taus: uncertainty of the particle decomposition component, relevant mainly for $p_T > 70$ GeV.
- *M*odelling of true taus: sum of many components related to modelling of pile-up, underlying event, and detector modelling.
- A single systematic uncertainty representing the total TES uncertainty is assigned for fake taus, and treated uncorrelated to the separate TES uncertainties on true taus.

The systematic uncertainty of the tau energy resolution was assessed both by smearing true monte carlo taus and by changing the tau hadronic parton shower model. In both cases the effect on the resolution was near 1% and this was found to have a negligible effect on the final result.

Jet Energy Scale (JES)

Since the multijet and $Z \to \tau\tau$ backgrounds are evaluated from data, the JES [23] uncertainty enters through the signal and minor backgrounds that are determined from MC simulation. The JES uncertainties arise from several independent sources:

- *In-situ jet energy corrections uncertainty:* These components account for bin-to-bin correlations in jet calibration and corrections derived from in-situ techniques.
- η *intercalibration uncertainty:* The uncertainty on the intercalibration in different detector pseudorapidity regions, containing a modelling and a statistical component.
- *Flavor composition and response uncertainties:* These components arise from the fact that knowledge of the quark-gluon composition is limited, while quark-initiated and gluon-initiated jets have different calorimeter responses. These components apply to light jets only (excluding *b*-jets). The composition is conservatively taken to be (50 ± 50)%. Since it is known that certain background and signal components can be more quark or more gluon dominated, two nuisance parameters are considered.

- *Uncertainties due to pile-up.* These components represent uncertainties due to in-time and out-of-time pileup (parametrised in terms of μ, the average number of interactions per bunch crossing and NPV, the number of primary vertices). Two components are considered. One component accounts for residual P_T dependence of the correction as a function of NPV and μ, but it is found to be unimportant in this analysis. The other component accounts for residual dependence on the underlying event of the jet energy scale following jet-area based pile-up correction.

Jet Energy Resolution (JER)

The systematic uncertainty due to the jet energy resolution (JER) is obtained by smearing every jet by the uncertainty in the resolution as determined by the in-situ measurement described in Ref. [97]. A one-sigma upwards variation in the resolution is obtained by smearing every jet with a smearing factor accounting for the uncertainty in the resolution in-situ measurement. The final effect of the variation is symmetrised in order to have a two-sided uncertainty in the fit. Similar to the JES, the JER uncertainty enters mainly through the minor backgrounds that are determined from MC simulation.

$E_{\mathrm{T}}^{\mathrm{miss}}$ Uncertainties

Systematic uncertainties on the energy scales of all objects affect the $E_{\mathrm{T}}^{\mathrm{miss}}$ and so it is recalculated after each of these variations is applied. In addition, the scale of the soft $E_{\mathrm{T}}^{\mathrm{miss}}$ term for energy outside reconstructed objects and the resolution uncertainties are considered independently [27].

5.7.2 Background Modelling Uncertainties

Systematic uncertainties on the data-driven $Z \to \tau\tau$ and multijet background estimation techniques are described below.

Embedding Method

The systematic uncertainties of the embedding method are related to the selection of $Z \to \mu\mu$ events in data and to the subtraction of the muon energy depositions in the calorimeters. A calorimeter isolation variable $I(E_{\mathrm{T}}, \Delta R)$ is defined as the sum of the total transverse energy in the calorimeter in a given cone of size ΔR around the the muon track, divided by the p_{T} of the muon. A track-based isolation $I(p_{\mathrm{T}}, \Delta R)$ is defined as the sum of the transverse momenta of tracks within a cone of ΔR around the muon track, divided by the muon p_{T}. The selection uncertainties of the embedding method are then estimated by varying the muon isolation criteria in the selection from the nominal value of $I(p_{\mathrm{T}}, 0.2) < 0.2$ to tighter ($I(p_{\mathrm{T}}, 0.4) < 0.06$ and $I(E_{\mathrm{T}}, 0.2) < 0.04$) and looser (no isolation requirements) values. The muon-related cell energies to be subtracted are varied within $\pm$20%.

Multijet Background

The default multijet template, derived from a sample in data where the tau candidates fail the isolation and opposite-sign charge requirements, is compared with an alternative template derived from a sample where the tau candidates fail just the opposite-sign charge requirement. The normalisation of the alternative template is fixed to that of the default template at preselection after the $\Delta\eta(\tau_{\mathrm{had}}, \tau_{\mathrm{had}})$ fit; the difference in how it propagates into the various categories gives a difference in yields, which along with the difference in shape between the two templates constitutes the systematic uncertainties on the background estimate. This leads to an overall multijet yield variation of 12 % (3 %) in the *VBF* (*Boosted*) category. However, there is a very strong shape dependence, such that the uncertainties on the BDT output are much larger at higher output values.

5.7.3 Theoretical Uncertainties

Theoretical systematic uncertainties are estimated for the signal and for all background contributions modelled by simulation. Since the major background contributions, from $Z \to \tau\tau$ and misidentification of hadronically decaying tau, are estimated using data-driven methods, they are not affected by these uncertainties. Uncertainties on the signal cross sections are assigned from missing higher-order corrections, from uncertainties in the PDFs, and from uncertainties in the modelling of the underlying event.

For the *VBF* and *VH* Higgs boson production cross sections the uncertainties due to missing higher order QCD corrections are estimated by varying the factorisation and renormalisation scales by factors of two around the nominal scale m_W, as prescribed by the LHC Higgs Cross Section Working Group [98]. The resulting uncertainties range from $\pm 2\%$ to $\pm 4\%$, depending on the process and the category-specific selection considered. In addition, a 2% uncertainty related to the inclusion of the NLO EWK corrections is assigned. It is based on the difference between the POWHEG and the HAWK [60, 61] calculation. The largest variation is found to be of the order of $\pm 3\%$ in the bin with the highest BDT output in the *VBF* category.

For the Higgs boson production via gluon fusion, the uncertainties on the cross sections associated with the analysis categories are estimated by varying the renormalisation and factorisation scales around the central values $\mu_{\mathrm{R}}, \mu_{\mathrm{F}} = \sqrt{m_H^2 + p_{\mathrm{T}}^2}$ in the NLO cross-section calculations of the $H + 1$ jet and $H + 2$ jet production. In the calculations appropriate cuts of $p_{\mathrm{T}}^H > 100$ GeV and on the jet kinematics ($\Delta\eta, p_{\mathrm{T}}$) are applied at parton level for the *Boosted* and *VBF* categories, respectively. The resulting uncertainties on the *ggF* contributions are found to be about $\pm 24\%$ in the *Boosted* and $\pm 23\%$ in the *VBF* categories. Whereas the *ggF* contribution is dominant in the *Boosted* category, it only contributes of the order of 20% to the signal in the *VBF* category. Since the two categories are exclusive, their anti-correlation is taken into account following the prescription of Ref. [99].

Although no explicit veto on additional jets is applied in the *VBF* selection, enough kinematical information is provided as input to the BDT so that the high BDT-output region corresponds to a more exclusive region, where the probability of finding a third-jet is reduced. Since the cross section of gluon-fusion events produced with a third jet is only known at LO, this could introduce a large uncertainty on the gluon-fusion contamination in the highest (and most sensitive) BDT-output bins. The uncertainty on the BDT shape of the *ggF* contribution has been evaluated using the MCFM Monte Carlo program [95], which calculates $H + 3$*jets* at LO. Scale variations induce changes of the *ggF* contribution in the highest BDT bin of about $\pm 30\%$. They have been taken into account in the final fit.

Uncertainties related to the simulation of the underlying event and parton shower are estimated by comparing the acceptance from POWHEG+PYTHIA to POWHEG+HERWIG for both *VBF* and *ggF* Higgs production modes. Differences in the signal yields range from $\pm$1% to $\pm$8% for the *VBF* and from $\pm$1% to $\pm$9% for *ggF* production, depending on the channel and category. The BDT score distribution of the POWHEG+PYTHIA and POWHEG+HERWIG samples are compatible with each other within statistical uncertainties.

The PDF uncertainties are estimated by studying the change in the acceptance when using different PDF sets or varying the CT10 PDF set within its uncertainties. The standard *VBF* POWHEG sample and a MC@NLO [100] *ggF* sample, both generated with the CT10 PDFs, are reweighted to the MSTW2008NLO [101], NNPDF [102] and the CT10 eigen-tunes parametrisation. The largest variation in acceptance for each category is used as a flat PDF uncertainty; it varies between approximately $\pm$4.5% and $\pm$6% for *ggF* production and between about $\pm$0.8% and $\pm$1.0% for *VBF* production. A shape uncertainty is also included to cover any difference between the BDT score in the default sample, and the reweighted ones. The uncertainty on the total cross section for *VBF*, VH and *ggF* production modes due to the PDFs is also considered.

Variations in the acceptance for different MC generators are also included, comparing POWHEG+HERWIG samples to MC@NLO+HERWIG for *ggF* and AMC@NLO+HERWIG [103] for *VBF*. The generator modelling uncertainty is around 2% for *ggF* and 4% for *VBF* productions modes.

Finally the uncertainty on the decay branching ratio, BR($H \rightarrow \tau\tau$), of $\pm$5.7% [73] affects the signal rates.

The theoretical systematic uncertainties on the background predictions taken from the simulation are evaluated by applying the same procedures as used for the signal samples. The estimated uncertainties resulting from the choices of the QCD scales, the PDF parametrisation and the underlying-event model are reported in Table 5.4.

Source	*VBF*		*Boosted*	
	S	*B*	*S*	*B*
Experimental				
Luminosity	±2.8	±0.1	±2.8	±0.1
Tau trigger	$^{+7.7}_{-8.8}$	< 0.1	$^{+7.8}_{-8.9}$	< 0.1
Tau identification	±6.6	±3.8	±6.6	±5.1
Tau energy scale†	±2.9	±2.5	±2.9	±2.5
Jet energy scale and resolution†	$^{+10.1}_{-8.0}$	±0.3	$^{+5.1}_{-6.2}$	±0.2
E_T^{miss} soft scale and resolution	±0.5	±0.2	±0.1	< 0.1
Background Model				
Multijet modelling†	–	±5.2	–	±0.6
Embedding method†	–	±2.2	–	±3.3
Theoretical				
Higher-order QCD corrections †	$^{+10.7}_{-8.2}$	< 0.1	$^{+20.3}_{-15.4}$	< 0.1
UE/PS	±4.6	< 0.1	±3.8	< 0.1
Generator modelling	±2.4	< 0.1	±1.2	< 0.1
EW corrections	±1.1	< 0.1	±0.4	< 0.1
PDF †	$^{+4.3}_{-4.0}$	$\pm$ 0.2	$^{+6.3}_{-5.8}$	$\pm$ 0.1
BR ($H \to \tau\tau$)	$\pm$ 5.7	–	±5.7	–

Table 5.4: Percent impact of systematic uncertainties on the total signal, *S*, (sum of all production modes) and on the sum of all background estimates, *B*, for each category. Uncertainties that affect the shape of the BDT-output distribution in a non-negligible way are marked with a †.

Chapter 6

Signal Extraction

This chapter documents the signal extraction procedure implemented in the multivariate analysis. While the cut-based analysis performs a fit of the $m_{\tau\tau}^{\mathrm{MMC}}$, the multivariate analysis trains boosted decision trees (BDTs) to separate signal from background and a fit is performed on the BDT output distributions. Several discriminating features are selected in each category and BDTs are trained. The binning of the BDT output distribution is optimized to obtain the best expected signal significance and the statistical model is constructed for the hypothesis testing. This model is validated in a $m_{\tau\tau}^{\mathrm{MMC}}$ sideband excluding the signal region. This chapter concludes with the expected sensitivity and the next chapter presents the unblinded results.

6.1 Discriminating Features

At a very early stage in the development of this analysis, a set of discriminating features was selected to train the BDTs. The final list of features was the result of an iterative process, involving the removal of weak features while checking the impact on the expected sensitivity, and a harmonization with the parallel searches performed in the $H \to \tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $H \to \tau_{\mathrm{lep}}\tau_{\mathrm{lep}}$ channels.

The selected features exploit differences in the event kinematics and topologies between the Higgs signal and $Z \to \tau\tau$ and multijet backgrounds. Some features are specific to the VBF category and describe the typical VBF topology of two central taus surrounded by two well-separated forward jets. Features common to the VBF and Boosted categories depend mainly on the tau candidates and $E_{\mathrm{T}}^{\mathrm{miss}}$, discriminating between resonant and non-resonant kinematics and favouring events where the $E_{\mathrm{T}}^{\mathrm{miss}}$ bisects the taus in the transverse plane. Features are defined and discussed below. Plots of all features are shown in the VBF and boosted categories in Figure 6.1-6.3 and are summarized in Table 6.1.

$m_{\tau\tau}^{\mathrm{MMC}}$

The $m_{\tau\tau}^{\mathrm{MMC}}$ di-tau mass, defined in Section 3.6, is the most powerful feature for discriminating against the $Z \to \tau\tau$ and multijet backgrounds in both the VBF and Boosted categories. Inclusion of the $m_{\tau\tau}^{\mathrm{MMC}}$ makes the BDTs explicitly dependent on a specific Higgs mass region if trained with signal samples for a single m_H. One proposed solution was to train a set of BDTs for each available m_H from 100 to 150 GeV in steps of 5 GeV and then to perform a hypothesis test using the corresponding BDT for each m_H. However, the scope of this analysis was finally limited to a search for a Higgs boson consistent with the one already observed with $m_H \approx 125$ GeV due to the smaller simulated signal samples at the other m_H values and the foreseen increase in the analysis complexity by performing the same statistical model validation at 10 additional m_H hypothesis. The cut-based analysis complements the multivariate analysis here by providing a simple probe of m_H since it directly profiles the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution. Plots of the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions are shown in Figure 6.1a for the VBF category and Figure 6.3a for the Boosted category.

$\Delta R_{\tau\tau}$

Combined with the $m_{\tau\tau}^{\mathrm{MMC}}$, the $\Delta R_{\tau\tau}$ helps discriminate between the resonant $Z \to \tau\tau$ and $H \to \tau\tau$ and non-resonant multijet background. This is clearly shown in the contours of Figure A.2. The $\Delta R_{\tau\tau}$ distributions are shown in Figure 6.1b for the VBF category and Figure 6.3b for the Boosted category. The lower $\Delta R_{\tau\tau}$ region is more populated in the Boosted category as the taus are typically closer together when the resonance is boosted.

$p_{\mathrm{T}}^{\mathrm{Total}}$

The visible components of the tau decay products, the two leading jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$ should approximately balance in the transverse plane for a well-reconstructed VBF event. p_T^{Total} is defined as:

$$p_T^{\mathrm{Total}} = |\vec{p}_{\mathrm{T}}^{\tau_1} + \vec{p}_{\mathrm{T}}^{\tau_2} + \vec{p}_{\mathrm{T}}^{\mathrm{j1}} + \vec{p}_{\mathrm{T}}^{\mathrm{j2}} + \vec{E}_{\mathrm{T}}^{\mathrm{miss}}| \tag{6.1}$$

and will tend to be closer to zero than background events. p_T^{Total} is only used in the VBF category and distributions are shown in Figure 6.2c.

$\sum p_{\mathrm{T}}$

The scalar sum of the p_{T} of the visible components of the tau decay products and of the jets will tend to be higher for well-reconstructed boosted signal events. $\sum p_{\mathrm{T}}$ is defined as:

$$\sum p_{\mathrm{T}} = p_{\mathrm{T}}(\tau_1) + p_{\mathrm{T}}(\tau_2) + \sum_{\mathrm{jets}} p_{\mathrm{T}} \tag{6.2}$$

and is only used in the boosted category. Distributions are shown in Figure 6.3d.

$p_{\mathrm{T}}(\tau_1)/p_{\mathrm{T}}(\tau_2)$

The tau p_{T} ratio is expected to be closer to one for well-reconstructed boosted di-tau events and help reject multijet background events. Distributions are shown in Figure 6.3e. This variable is only used in the Boosted category.

$E_{\mathrm{T}}^{\mathrm{miss}}\phi$ centrality

The $E_{\mathrm{T}}^{\mathrm{miss}}\phi$ centrality is a variable that quantifies the relative angular position of the $E_{\mathrm{T}}^{\mathrm{miss}}$ with respect to the tau decay products in the transverse plane. The transverse plane is transformed such that the direction of the tau decay products are orthogonal, and that the smaller ϕ angle between the tau decay products defines the positive quadrant of the transformed plane. $E_{\mathrm{T}}^{\mathrm{miss}}\phi$ centrality is defined as the sum of the x' and y' components of the $E_{\mathrm{T}}^{\mathrm{miss}}$ unit vector in this transformed plane:

$$C_{\phi_{\tau 1},\phi_{\tau 2}}(\phi_{E_{\mathrm{T}}^{\mathrm{miss}}}) = \frac{x' + y'}{\sqrt{x'^2 + y'^2}} \tag{6.3}$$

where

$$x' = \frac{\sin(\phi_{E_{\mathrm{T}}^{\mathrm{miss}}} - \phi_{\tau 1})}{\sin(\phi_{\tau 2} - \phi_{\tau 1})} \qquad y' = \frac{\sin(\phi_{\tau 2} - \phi_{E_{\mathrm{T}}^{\mathrm{miss}}})}{\sin(\phi_{\tau 2} - \phi_{\tau 1})} \tag{6.4}$$

$C_{\phi_{\tau 1},\phi_{\tau 2}}(\phi_{E_{\mathrm{T}}^{\mathrm{miss}}})$ is bounded between $\sqrt{2}$ when the $E_{\mathrm{T}}^{\mathrm{miss}}$perfectly bisects the smaller ϕ angle between the taus and $-\sqrt{2}$ when the $E_{\mathrm{T}}^{\mathrm{miss}}$is pointing in the opposite direction. If the $E_{\mathrm{T}}^{\mathrm{miss}}$ is aligned with either of the taus then $C_{\phi_{\tau 1},\phi_{\tau 2}}(\phi_{E_{\mathrm{T}}^{\mathrm{miss}}})$ has a value of 1. This feature is used in the VBF and Boosted categories. Distributions of the positive $C_{\phi_{\tau 1},\phi_{\tau 2}}(\phi_{E_{\mathrm{T}}^{\mathrm{miss}}})$ region are shown in Figure 6.1c for the VBF category and Figure 6.3c for the Boosted category. In well-modelled $H \to \tau\tau$ and $Z \to \tau\tau$ events the $E_{\mathrm{T}}^{\mathrm{miss}}$ bisects the taus more frequently than in the multijet background, leading to higher values of $C_{\phi_{\tau 1},\phi_{\tau 2}}(\phi_{E_{\mathrm{T}}^{\mathrm{miss}}})$.

Tau η centrality

Similar to the $E_{\mathrm{T}}^{\mathrm{miss}}\phi$ centrality, a features is constructed that describes the centrality of each tau candidate in the longitudinal plane with respect to the two selected jets in the VBF category. The tau η centrality follows the mathematical definition of a Gaussian distribution centered at the midpoint in η between the two jets and with width proportional to the η separation between the two jets. Beginning with the definition of the center and width:

$$\mu = \frac{\eta_{j1} + \eta_{j2}}{2} \qquad \sigma = \eta_{j1} - \mu \tag{6.5}$$

the η centrality $C_{\eta_{j1},\eta_{j2}}(\eta_\tau)$ is defined as:

$$C_{\eta_{j1},\eta_{j2}}(\eta_\tau) = \exp\left[-\frac{(\eta_\tau - \mu)^2}{\sigma^2}\right] \tag{6.6}$$

leading to:

$$C_{\eta_{j1},\eta_{j2}}(\eta_\tau) = \exp\left[\frac{-4}{(\eta_{j1} - \eta_{j2})^2}\left(\eta_\tau - \frac{\eta_{j1} + \eta_{j2}}{2}\right)^2\right], \tag{6.7}$$

where η_τ, η_{j1} and η_{j2} are the pseudorapidities of the tau and the two leading jets, respectively. This variable has a value of 1 when the tau is halfway in η between the two jets, $1/e$ when the tau is aligned with one of the jets, and $< 1/e$ when the object is outside the jets. This variable is only used in the VBF category and distributions are shown for the two taus in Figure 6.2a and Figure 6.2b. Well-reconstructed VBF signal events will typically have η centrality values closer to one than the backgrounds.

$\Delta\eta(j_1, j_2)$

The absolute pseudorapidity difference between the VBF jets will tend to be larger for VBF signal events than any of the backgrounds. Distributions are shown in Figure 6.1e.

$\eta_{j1} \times \eta_{j2}$

In the VBF category, the two leading jets will tend to be in opposite hemispheres of the ATLAS detector for VBF Higgs production. The product of the jet pseudorapidities will therefore tend to be negative. This feature is of course highly correlated with $\Delta\eta(j_1, j_2)$ but adds information in the events where the absolute pseudorapidity difference is large but both jets are in the same hemisphere. Distributions of $\eta_{j1} \times \eta_{j2}$ are shown in Figure 6.1f.

$m_{j1,j2}$

The invariant mass of the VBF di-jet system is also correlated with $\Delta\eta(j_1, j_2)$ but this feature accounts for the energy of the jets, which is expected to be higher for the VBF signal. Distributions of $m_{j1,j2}$ are shown for the VBF category in Figure 6.1d.

Variable	Definition	VBF	Boosted
$m_{\tau\tau}^{MMC}$	MMC di-tau mass	•	•
$\Delta R_{\tau\tau}$	ΔR separation of the two taus	•	•
$p_{\mathrm{T}}(\tau_1)/p_{\mathrm{T}}(\tau_2)$	Ratio of leading to subleading tau p_{T}		•
$\sum p_T$	Sum of p_{T} of taus and jets		•
p_T^{Total}	$\lvert \vec{p}_{\mathrm{T}}^{\tau_1} + \vec{p}_{\mathrm{T}}^{\tau_2} + \vec{p}_{\mathrm{T}}^{\mathrm{j1}} + \vec{p}_{\mathrm{T}}^{\mathrm{j2}} + \vec{E}_{\mathrm{T}}^{\mathrm{miss}} \rvert$	•	
$m_{j1,j2}$	Invariant mass of the two leading jets	•	
$\eta_{j1} \times \eta_{j2}$	η product of the two leading jets	•	
$\Delta\eta(j_1, j_2)$	η separation of the two leading jet	•	
$E_T^{miss}\phi$ centrality	See text	•	•
τ_1 η centrality	See text	•	
τ_2 η centrality	See text	•	

Table 6.1: Discriminating variables used to train BDTs in the VBF and Boosted categories.

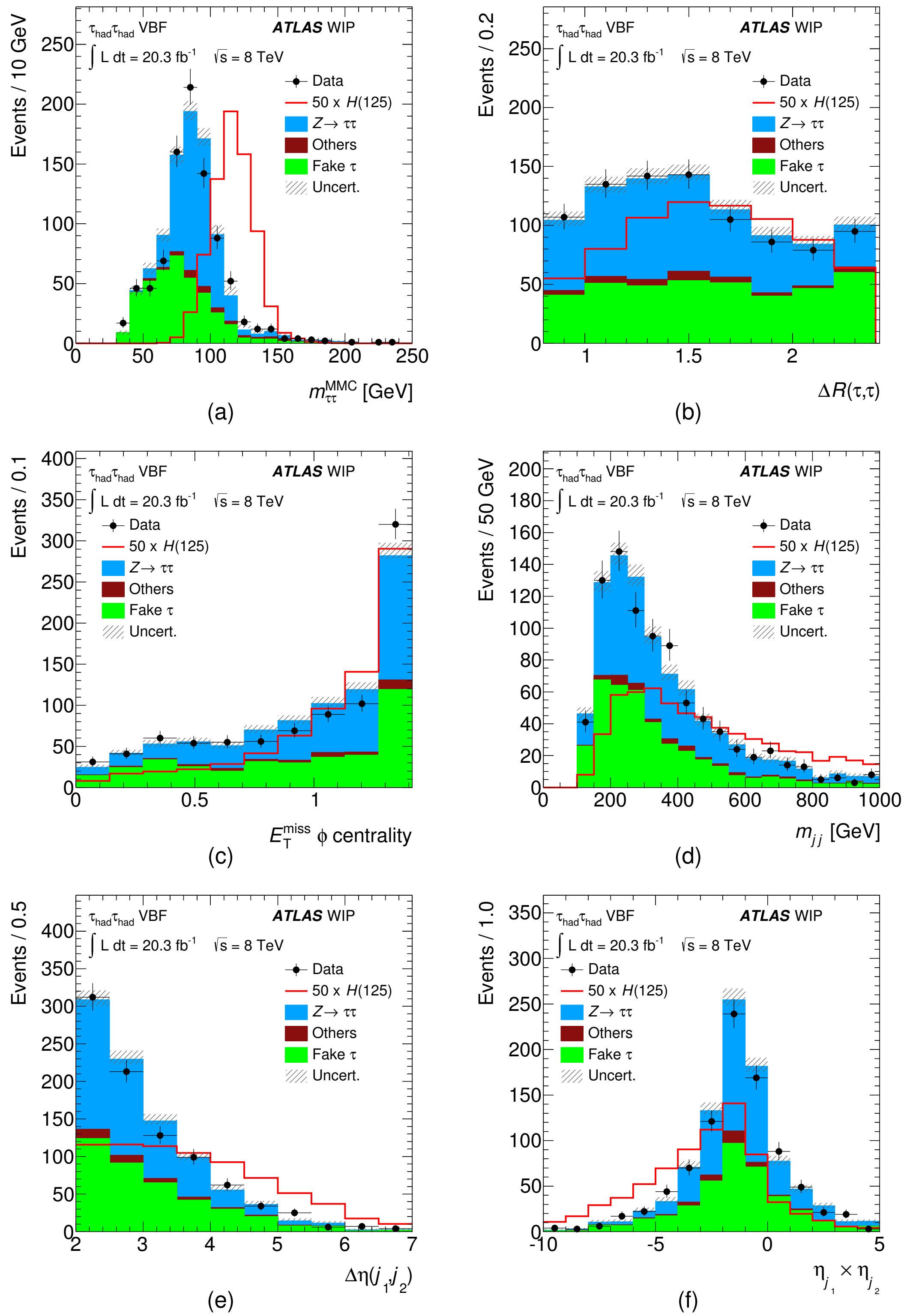


Figure 6.1: Discriminating features used in the VBF category (see Section 6.1).

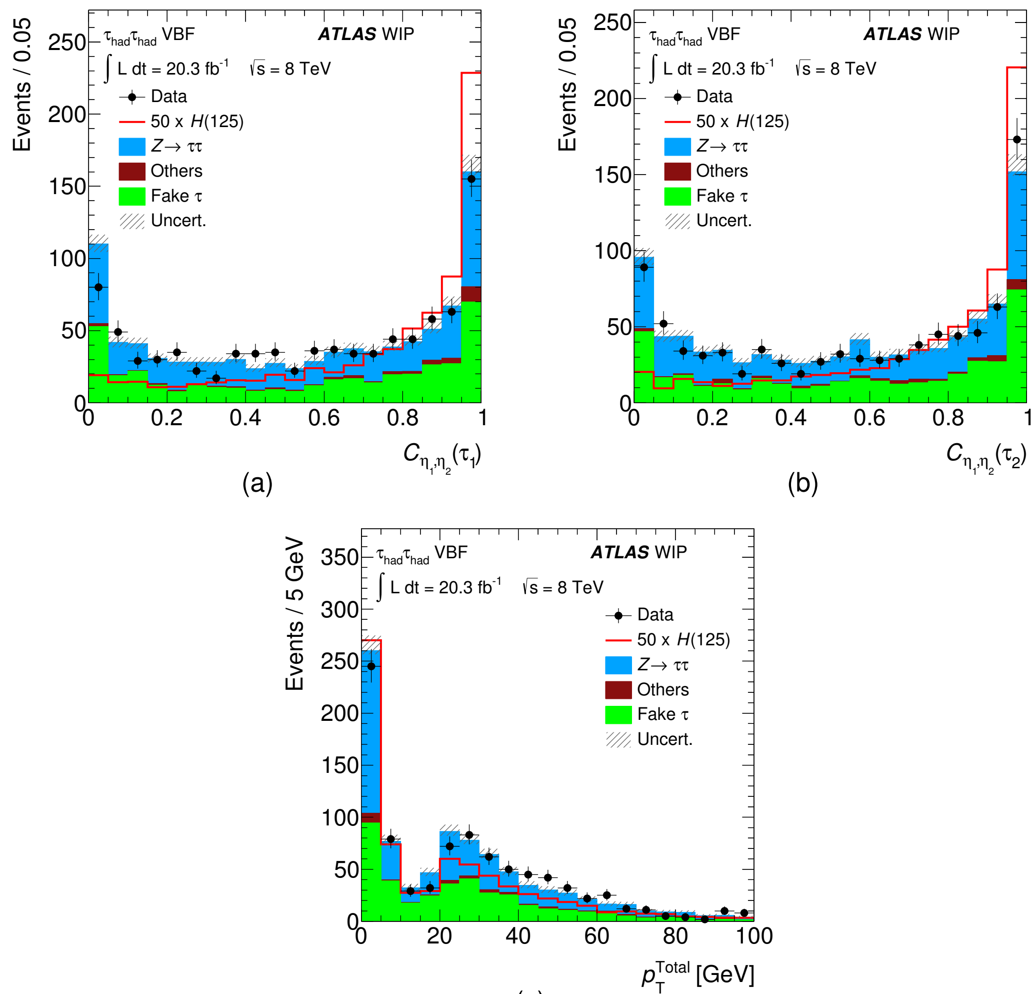


Figure 6.2: Discriminating features used in the VBF category (see Section 6.1).

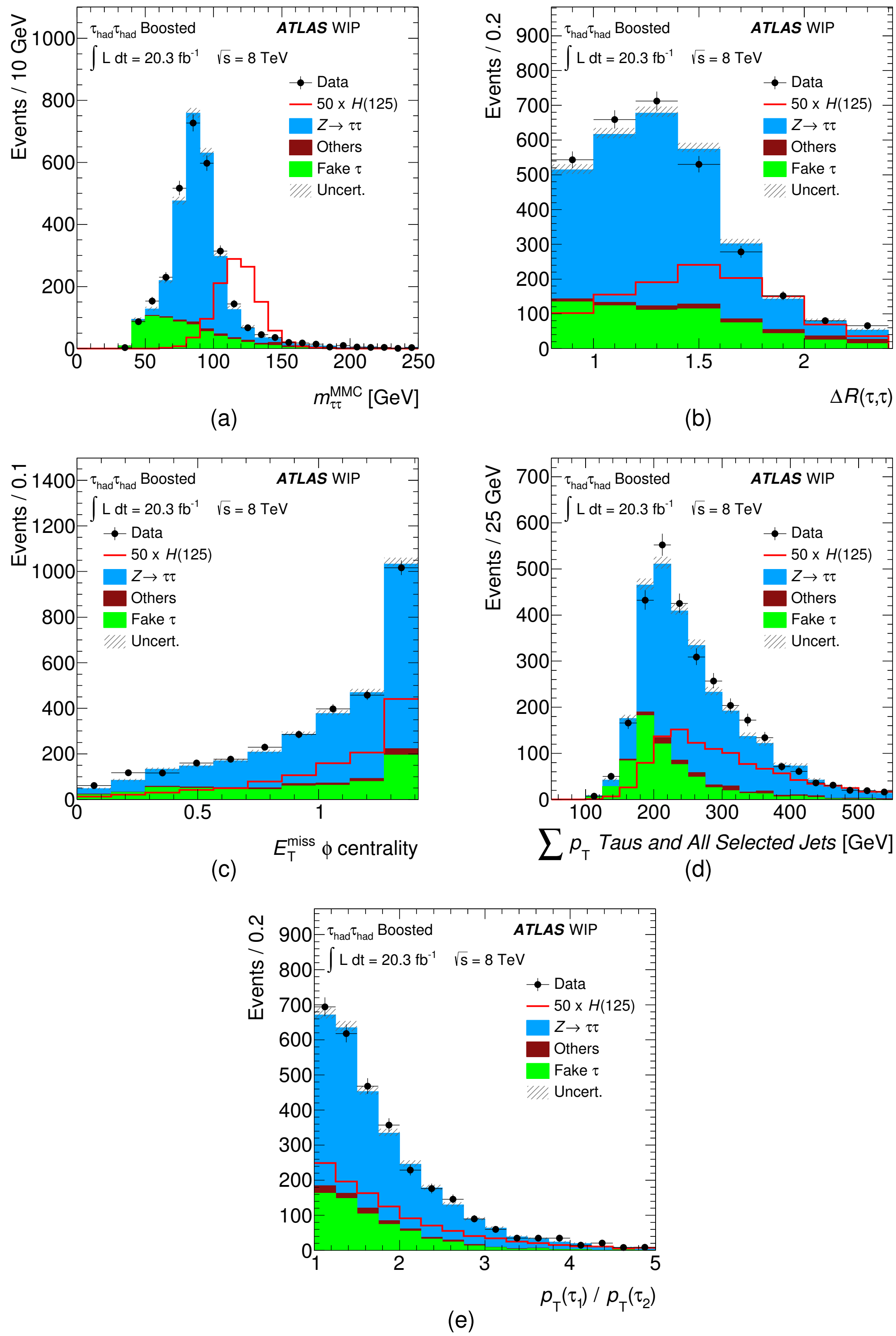


Figure 6.3: Discriminating features used in the Boosted category (see Section 6.1).

6.2 Classification with Boosted Decision Trees

The main component of this search is the task of labelling events in data as either signal or background based on what can be learned about typical events in the signal and background models. This is formally known as binary classification with supervised learning, a form of machine learning where a sample of events with known signal and background labels (the *training* sample) is used to infer labels for independent events (the *test* sample) where labels may or may not be known a priori. There exist many algorithms to perform this task, from the very simple approach of placing a single cut on one feature, to placing cuts (even oblique cuts) on several features as done in the cut-based analysis presented in this thesis, to progressively more complicated algorithms such as "nearest neighbour", decision trees, support vector machines, or artificial neural networks.

No matter how simple or complex the algorithm is, the goal is always to define *optimal* decision boundaries between regions where signal is enhanced and regions that are dominated by background processes. Optimal boundaries minimize the classification error on the test sample and maximize the sensitivity of the search to the presence of the Higgs signal in the data. Depending on the dimensionality of the feature space and the complexity of the relationships between features in the signal and background distributions, the optimal decision boundaries may be simple and linear or highly multivariate, nonlinear, and may define many disjoint signal regions. This complexity is what sets apart the various classification algorithms. A simple algorithm may be sufficient to classify events that are for the most part linearly separable, but a more complex algorithm is often justified by the fact that it can learn more complicated decision boundaries and achieve a lower error on difficult classification problems.

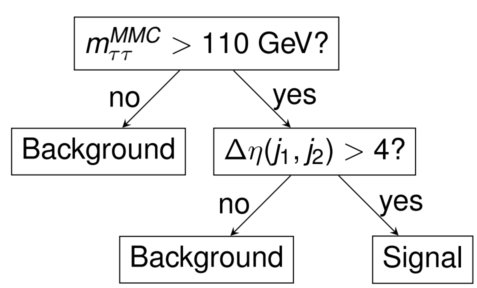


Figure 6.4: A simple decision tree.

Decision trees offer a favourable balance between simplicity and exceptional performance on a wide range of classification problems. This combination has largely influenced their widespread adoption by the High Energy Physics community. The decision tree algorithm will be outlined in detail below, but in simple terms they recursively partition the feature space into multiple rectangular regions where signal or background purities are enhanced. The sequence of cuts (decisions) are laid out in a binary tree data structure. Events begin at the root node of the binary tree and are passed down through the tree to right or left daughter nodes depending on whether they pass or fail a cut on a single feature. Events end up at one of the leaf nodes where they can be classified as either signal or background depending on which label was dominant at the leaf node in the training sample, or as is the case in this analysis, assigned a continuous output value such as the signal purity of the leaf node.

Decision trees are relatively easy to understand and interpret since the trees can be visualized and the classification of each event can be explained by a series of Boolean conditions. In contrast, neural networks combine responses of many nodes in multiple "hidden" layers in complex ways that are difficult to interpret and yet have been known to offer little or no gain in performance over more simple learning models. Decision tree learning algorithms, however, are known to grow overly-complex trees that do not generalize well to unseen events in the test sample if the algorithm is not properly controlled. This condition is called *overfitting* and occurs when the model learns too many details about the particular sample of events in the training sample, going beyond the general underlying relationships between features and learning statistical noise. This is illustrated by the extreme case where a decision tree is allowed to continue splitting the training sample into ever smaller regions until the leaf nodes each contain a single training event. This decision tree would perfectly classify all events in the training sample but would have a very high classification error on the independent test sample. After selecting an appropriate machine learning algorithm, it is most critical to determine the values of the learning model parameters that give the highest level of generalization and performance on the independent test sample. A *cross-validated grid-search* (described below) is used to determine the best learning parameters for this analysis.

Decision trees can also be unstable as small variations in which events are used in the training sample can lead to completely different tree structures. The problem of learning the optimal tree is also NP-complete, meaning it cannot be solved in polynomial time in any known way. Practical algorithms learn tree structures in a "top-down" manner by determining the cuts that locally optimize the separation between signal and background at each node, even though a suboptimal cut at an intermediate node could have resulted in more optimal cuts at deeper nodes, possibly resulting in a decision tree with a better overall classification. These are reasons why decision trees are usually learned as part of an ensemble where the responses of many decision trees are averaged. The averaging of many simple decision trees can also model complex nonlinear decision boundaries even though the few leaf nodes of each tree enclose rectangular regions in the feature space. This is illustrated by the toy classification problem in Figure 6.5. This analysis uses a variant of the AdaBoost [104] algorithm that *adaptively boosts* the weights of events misclassified by a decision tree before training the next decision tree. Subsequent decision trees in the ensemble effectively focus more on correcting erroneous classification by previous decision trees and common patterns learned by multiple trees are reinforced. The ensemble of decision trees is called a boosted decision tree and the output is an average of the outputs from each decision tree. Figure 6.6 demonstrates how AdaBoost improves the classification performance beyond what is achieved by a single decision tree.

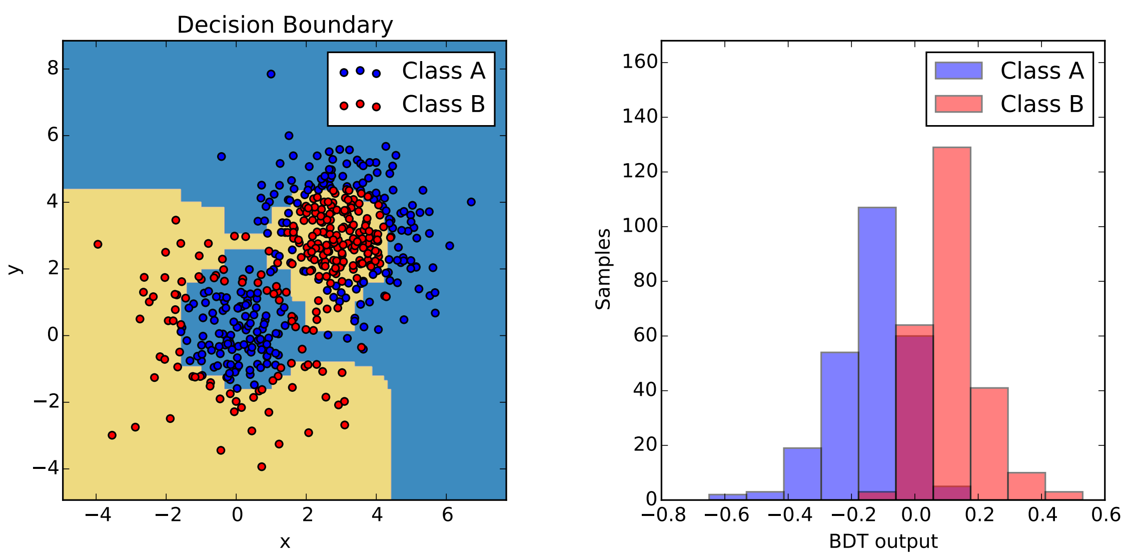


Figure 6.5: A toy two-class dataset and the nonlinear decision boundary learned by a BDT is shown on the left and the corresponding BDT output distributions are shown on the right.

The mathematical formulation of the decision tree and AdaBoost algorithms is as follows. The decision tree node splitting algorithm begins at the root node with a vector of N training events each described by n features $x_i \in R^n, i \in 1, ..., N$ and a corresponding vector of class labels $y_i \in \{0, 1\}$ (background = 0 and signal = 1) and event weights $w_i \in R$.

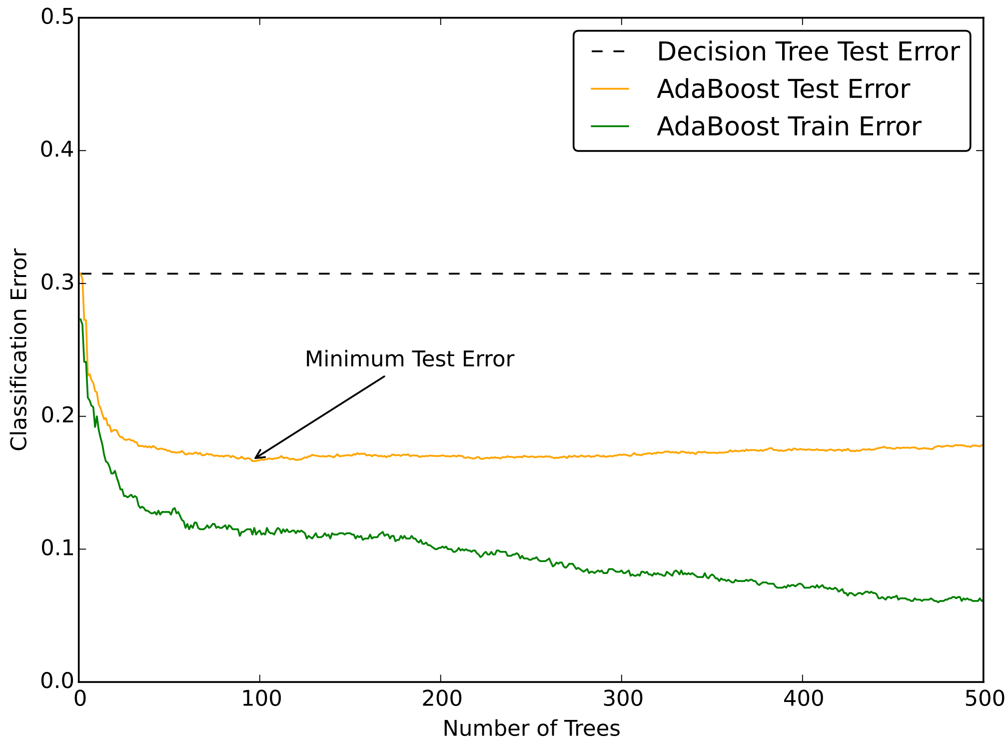


Figure 6.6: Demonstration of how the test and train error depend on the number of trees in the AdaBoost ensemble with a toy dataset. AdaBoost quickly improves the classification error beyond what a single decision tree achieves. The test error reaches a minimum at around 100 trees before the model begins to overfit the training sample.

Let the subset of the training events, labels, and weights at node m be represented by Q. For each candidate cut $\theta = (j, t_m)$ consisting of a feature j and a cut threshold t_m, Q can be partitioned into two disjoint subsets $Q_{\text{left}}(\theta)$ and $Q_{\text{right}}(\theta)$:

$$Q_{\text{left}}(\theta) = (x, y, w)|x_j \leq t_m \qquad Q_{\text{right}}(\theta) = Q \setminus Q_{\text{left}}(\theta) \tag{6.8}$$

The *impurity* at node m is computing using the Gini index

$$G(Q) = p_m(1 - p_m) \tag{6.9}$$

where p_m is the signal purity at node m

$$p_m = \frac{\sum_{i \in Q} w_i \times I(y_i = 1)}{\sum_{i \in Q} w_i} \tag{6.10}$$

I is the indicator function with a value of 1 if $y_i = 1$ (signal) and 0 otherwise. The best cut on the best feature that minimizes the combined impurity of the left and right partitions θ^* is then determined

$$H(Q, \theta) = \sum_{i \in Q_{\text{left}}(\theta)} w_i G(Q_{\text{left}}(\theta)) + \sum_{i \in Q_{\text{right}}(\theta)} w_i G(Q_{\text{right}}(\theta)) \tag{6.11}$$

$$\theta^* = \operatorname*{argmin}_{\theta} H(Q, \theta) \tag{6.12}$$

The same procedure is then applied on the left and right partitions $Q_{\text{left}}(\theta^*)$ and $Q_{\text{right}}(\theta^*)$, and left and right daughter nodes are created below the current node. This repeats recursively in daughter nodes until a *stopping criterion* is satisfied, such as when a minimum number of events is reached at a node or the tree grows to a maximum depth. For this analysis, where there is a large difference between the number of events in the signal and background samples, the stopping criterion that gave the best results is requiring a minimum fraction of the overall training sample weight in each leaf node. This criterion is also easier to interpret since it does not depend on the total number of events in the training sample.

There are many machine learning software libraries that provide various implementations of boosted decision trees. `TMVA` [32] (Toolkit for Multivariate Data Analysis) is the most popular library used in conjunction with `ROOT` [105] but it has several shortcomings, including the lack of built-in support for training classifiers in parallel on multi-core processors and the lack of built-in tools for cross-validation and grid-searching training parameters. The `scikit-learn` [106] library provides all of this and an optimized version of the CART [107] decision tree algorithm. After implementing AdaBoost and support for weighted events, `scikit-learn` could be leveraged as a complete `TMVA` replacement for this analysis.

`scikit-learn` provides the SAMME and SAMME.R [108] AdaBoost extensions (Stage-wise Additive Modeling using a Multi-class Exponential loss function) that, although is unnecessary for this analysis, support multi-class problems. For binary classification the SAMME variants behave similar to the classic AdaBoost algorithm. The SAMME.R variant uses real-valued predictions such as weighted probability estimates (leaf node signal purity responses from each decision tree in this case) and was seen to produce the best results. After normalizing the event weights such that they sum to unity, a decision tree is constructed by the algorithm above and the signal purity (class probability) response is obtained for each event:

$$p_k^{(m)}(x) = \mathrm{Prob}(c = k|x), k \in 0, 1 \tag{6.13}$$

The response of the current decision tree m within the ensemble is then set to:

$$h_k^{(m)} \leftarrow \log p_k^{(m)}(x) - \frac{1}{2}\sum_{k'} \log p_{k'}^{(m)}(x), k \in 0, 1 \tag{6.14}$$

and the weights are altered according to:

$$w_i \leftarrow w_i \times \exp\left(-\beta \frac{1}{2} \boldsymbol{y}_i^T \log \boldsymbol{p}^{(m)}(x_i)\right), i \in 1, ..., N \tag{6.15}$$

where β is the *learning rate* and $\boldsymbol{y}$ is a two-vector with the components:

$$y_k = \begin{cases} 1 & \text{if } c = k \text{ (correct classification)} \\ -1 & \text{if } c \neq k \text{ (incorrect classification)} \end{cases} \tag{6.16}$$

If a background (signal) event lands at a leaf node with a signal purity above (below) 50% then it is marked as misclassified. The event weights are re-normalized and another decision tree is constructed. The weights of correctly classified events are suppressed and the weights of misclassified events are enhanced so the subsequent decision tree effectively focusses more difficult events. Boosting and training additional decision trees continues until a certain number M of trees is reached that is set in advance. The learning rate β controls the rate at which the event weights are boosted in subsequent trees. Small values of β lead to a more gradual learning and often better performance.

The output of the ensemble of decision trees for a single event is the average of all decision tree responses:

$$C_k(x_i) = \frac{1}{M}\sum_{m=1}^{M} h_k^{(m)}(x_i) \tag{6.17}$$

The $k = 0$ component $C_0(x_i)$ (background "probability" estimate) is multiplied by -1 and summed with the $k = 1$ signal component. If the resulting value is greater (less) than zero the event is classified as signal-like (background-like). For convenience, and to harmonize the output score distributions across all three analysis channels, the output t is transformed with the logistic function:

$$t' = \frac{2}{1 + \exp(-\alpha t)} - 1 \tag{6.18}$$

that restricts the score output to the domain $(-1, 1)$. An appropriate value for α is determined empirically based on how the transformation affects shape of the BDT output distribution. The value $\alpha = 2M\beta/3$ produces reasonable output distributions in both VBF and Boosted categories.

A BDT should not be applied on the same events used to train it. This is because, as discussed above, a classifier typically has some level of bias toward the events in the training sample and will perform better on those events than events in an independent sample. To avoid discarding events in the training sample, a simple method has been devised to allow the use of all events in the final result while never applying a BDT on the same events that trained it. The full signal and background samples are first split in half into even and odd samples using the parity of the event number and a BDT is trained on each half. Each BDT is then applied on the opposite half. Although OS data is not used in the training for either half, the same approach of applying the even-trained (odd-trained) BDT on the odd (even) events is followed.

The BDT trained on the even half is first optimized using a cross-validated grid-search over the number of trees and minimum weighted fraction of the training sample allowed at a leaf node (minimum leaf fraction). These parameters are among the most important for tuning the performance of a BDT, affecting the number of times boosting occurs and the complexity of each tree. The minimum leaf fraction is searched with 100 steps from 0.001 to 0.3 in the VBF category and to 0.04 in the Boosted category. The number of trees is searched from 1 to 200 in both categories. This grid is composed of 20,000 points in the parameter space. A larger grid was initially used before zooming in on the region yielding the best performance in each category. The training in the VBF category uses a learning rate of 0.1 and a slower rate of 0.01 was seen to provide better results in the Boosted category. To estimate the performance at each grid point, a 10-fold cross-validation is performed. Cross-validation is a common way to estimate the generalization (performance on independent events) of the learning model by using only the events present in the training sample. The training sample is split into 10 chunks of approximately equal size and a BDT is trained for each combination of 9 chunks and then tested on the chunk left out. The cross-validation is *stratified*, meaning, as the training sample is split into 10 chunks, the relative fraction of signal and background in each chunk is approximately equal. A performance metric is evaluated on the test chunk for each iteration and then all 10 values of the metric are averaged to give the estimate of the general performance of a BDT trained at the grid point.

The metric used in this analysis is the area under the receiver operating characteristic (ROC) curve (AUC). This curve is defined by the signal efficiency versus the background efficiency as a cut passes from left to right over the full range of the BDT output. The area under this curve is 1 if the classes are perfectly separable, and 0.5 if the signal and background distributions are completely overlapping, as would be achieved by a classifier performing no better than random guessing.

The cross-validated AUC over the grid-search is shown in Figure 6.8 for the VBF and Boosted categories. These figures show the location of the optimal parameters and summarize the training of 200,000 BDTs in each category (10 cross-validation folds over 20,000 grid points). The optimal parameters and training sample sizes in each category are summarized below in Table 6.2. The same optimal parameter values are used when training BDTs on the odd half of the samples in each category.

Features can be ranked by their relative importance in constructing each BDT. Stronger features will be selected more often when splitting nodes and will bring a greater overall reduction in the impurity $H(Q, \theta)$ (Equation 6.11). The importance of a feature is computed as the normalized total reduction in $H(Q, \theta)$ averaged over all trees in the boosted ensemble. The ranking of the features by importance is shown in Table 6.3a for the VBF category and Table 6.3b for the Boosted category. $m_{\tau\tau}^{MMC}$ and $\Delta R(\tau, \tau)$ are the highest ranked features in both categories. The first tree of the VBF BDT trained on the even sample is shown in

	VBF	Boosted
Signal Events	10,734	10,663
Background Events	2,433	6,773
Number of Trees	69	66
Minimum Leaf Fraction	0.103	0.012
Learning Rate	0.1	0.01

Table 6.2: The training sample sizes and optimal training parameter values in the VBF and Boosted categories.

Figure 6.9. The first cut of the decision tree is on $m_{\tau\tau}^{MMC}$ at the root node.

BDT output distributions in the VBF and Boosted categories are shown in Figure 6.7 where the Higgs signal is scaled by a factor of 50. The data is blinded in bins where the signal efficiency is greater than 50%. Appendix A.1 shows signal and background distribution contours in important two-dimensional planes exploited by the cut-based analysis, as well as how the average BDT score behaves when projected onto the same planes.

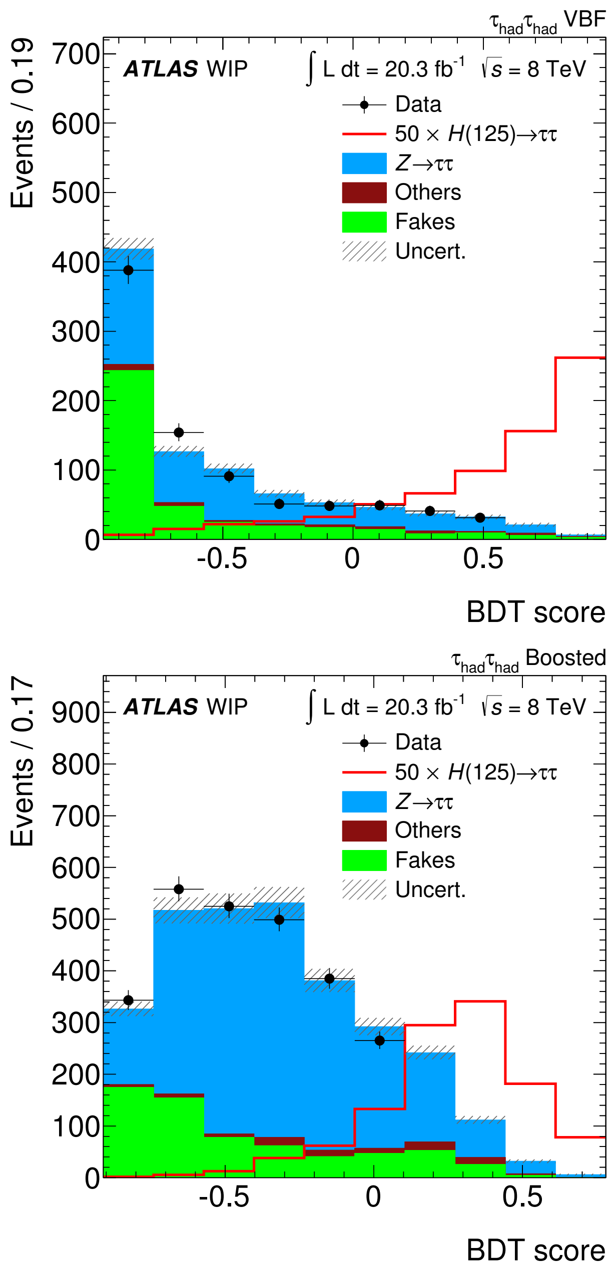


Figure 6.7: BDT output distributions in the VBF (top) and Boosted (bottom) categories. The data in the bins with an expected signal efficiency of at least 50% are blinded. The Higgs signal is scaled by a factor of 50 to illustrate the difference between the signal and background distributions.

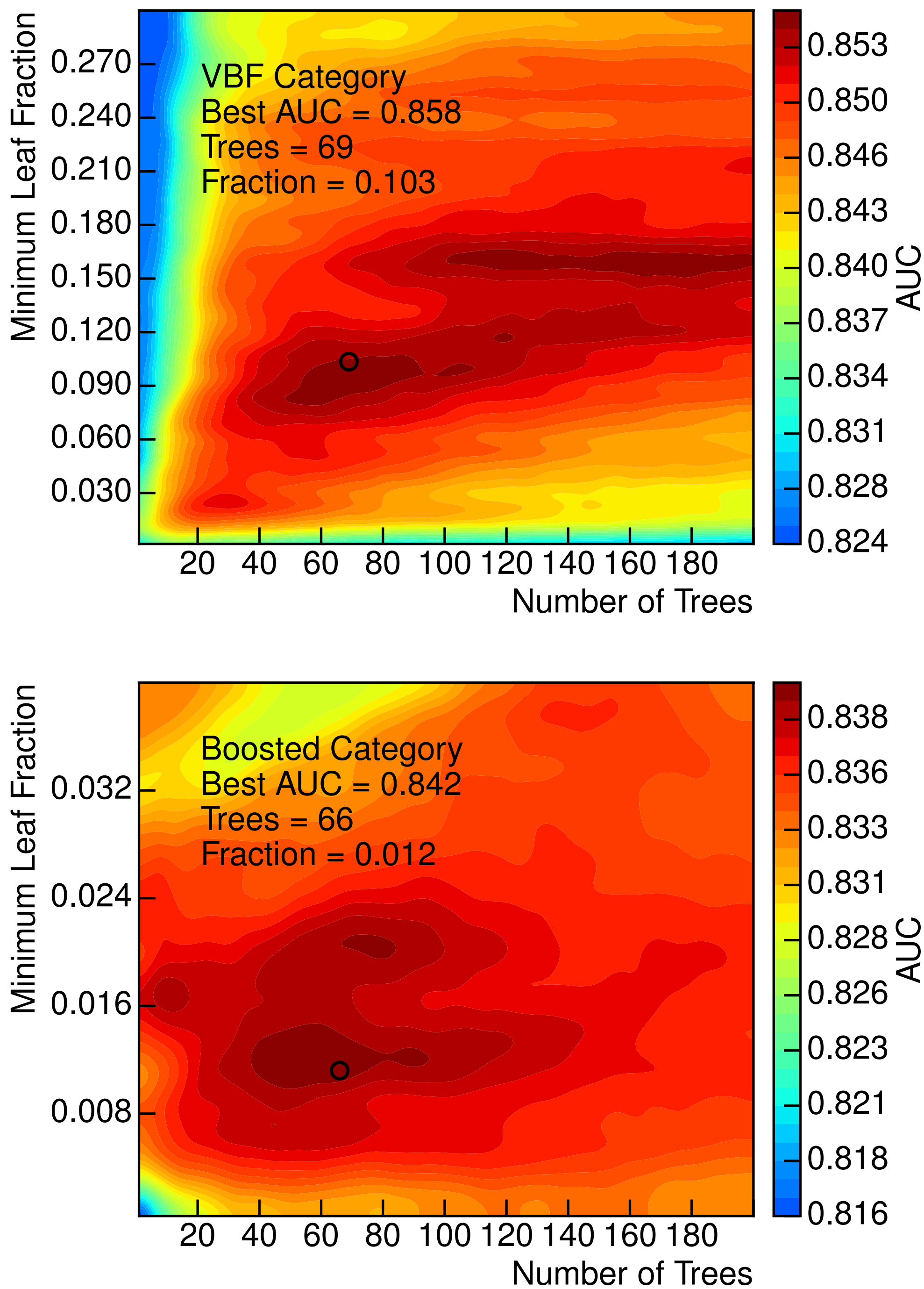


Figure 6.8: The cross-validated area under the ROC curve (AUC) grid-searched over the number of trees and minimum leaf fraction in the VBF (top) and Boosted (bottom) categories. The optimal BDT training parameters are marked with open circles. A total of 200,000 BDTs were trained in each category for the grid shown above after initially searching coarser grids extending higher along each axis.

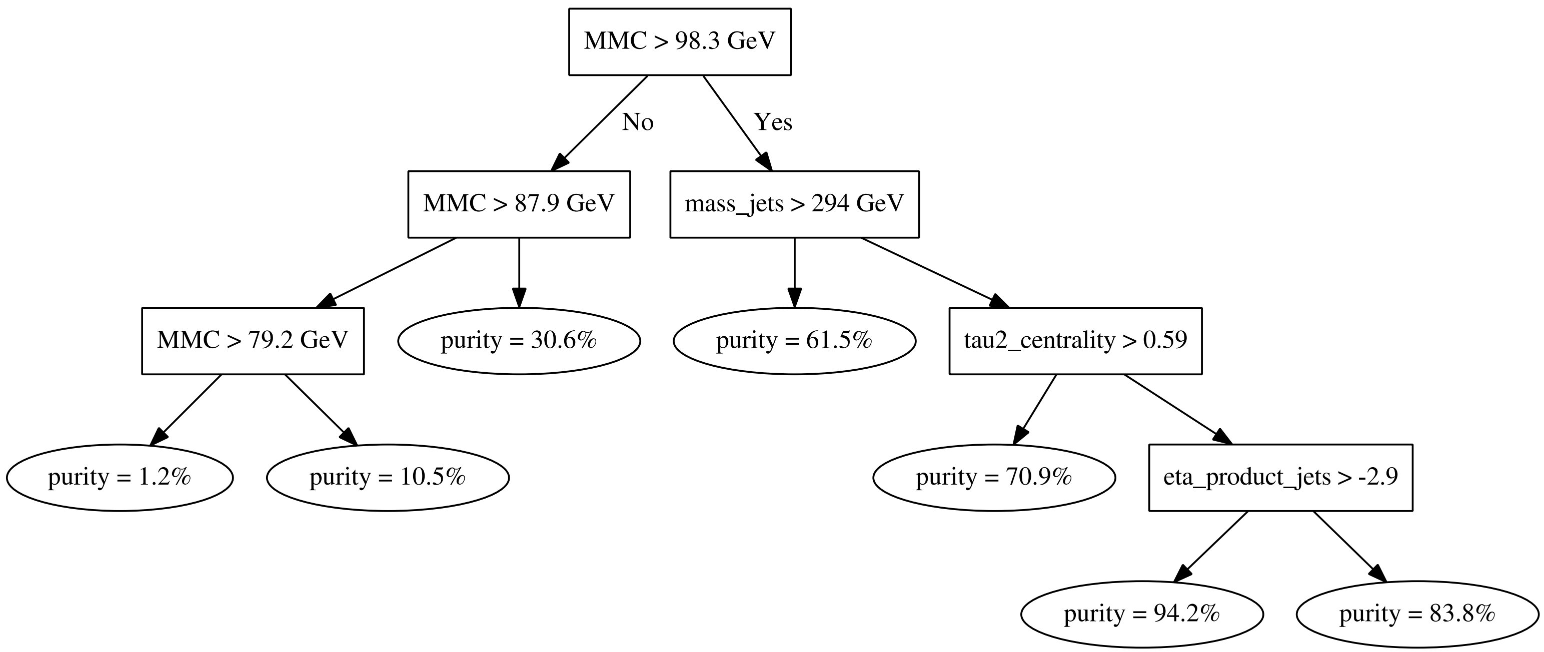


Figure 6.9: The first tree of the BDT trained in the VBF category. The leaf node signal purity is evaluated after the signal sample is normalized to the same weight as the background sample.

Rank	Feature	Gini Importance
1	$m_{\tau\tau}^{MMC}$	0.360
2	$\Delta R(\tau,\tau)$	0.127
3	$\sum \vec{p}_T$	0.122
4	τ_2 Centrality	0.095
5	E_T^{miss} Centrality	0.088
6	$\Delta\eta(jet_1, jet_2)$	0.065
7	jet$_1$ $\eta\times$ jet$_2$ η	0.061
8	$M(jet_1, jet_2)$	0.045
9	τ_1 Centrality	0.036

(a)

Rank	Feature	Gini Importance
1	$m_{\tau\tau}^{MMC}$	0.893
2	$\Delta R(\tau,\tau)$	0.033
3	$\tau_1 p_T/\tau_2 p_T$	0.027
4	E_T^{miss} Centrality	0.025
5	$\sum p_T$	0.022

(b)

Table 6.3: Training features ranked by Gini importance in the VBF (a) and Boosted (b) categories.

6.3 The Statistical Model

Particle physics experiments are often in search for a process or particle that has been predicted and a robust statistical formalism is required to test the compatibility of the observed data with the background-only hypothesis. If the disagreement between the data and background model is significant enough, then the background-only hypothesis is rejected in favour of the signal-plus-background hypothesis. If the signal sought by the experiment is not observed at any significant level, then an upper limit is placed on the rate at which this new process can occur, possibly rejecting the signal-plus-background hypothesis altogether. The ATLAS Higgs group uses a selection of frequentist methods involving a likelihood function to quote the expected and observed significance of a signal or to set exclusion limits [109]. All parameters have been included in the likelihood function following the recommendations of the LHC Higgs Combination Group [110]. This section outlines the construction of the likelihood function and how the expected and observed significance of a deviation from the background-only hypothesis is determined.

The Likelihood Function

The expected number of entries n_i in the i^{th} bin of a histogram is determined by the expected background and predicted signal scaled by the signal strength parameter μ:

$$E[n_i] = \mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}) \tag{6.19}$$

where the mean signal and background contributions are determined by integrating their respective probability density functions in bin i:

$$s_i = s_{\text{tot}} \int_{\text{bin } i} f_s(x; \boldsymbol{\theta}_s)\, dx \qquad b_i = b_{\text{tot}} \int_{\text{bin } i} f_b(x; \boldsymbol{\theta}_b)\, dx \tag{6.20}$$

$\boldsymbol{\theta}$ are nuisance parameters that encode uncertainties affecting the shape and overall normalization of the histograms that enter the model. Since the contents of each histogram bin follow Poisson statistics, the binned likelihood function is a product of Poisson probabilities over all bins:

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j(\boldsymbol{\theta}) + b_j(\boldsymbol{\theta}))^{n_j}}{n_j!} e^{-(\mu s_j(\boldsymbol{\theta}) + b_j(\boldsymbol{\theta}))} \tag{6.21}$$

The full parametrized probability density function used to model the signal and background distributions in this analysis is:

$$P(n_{cb}, a_p | \phi_p, \alpha_p, \gamma_b) = \prod_{c \in \text{categories}} \prod_{b \in \text{bins}} \text{Pois}(n_{cb}|\nu_{cb}) \times \prod_{p \in uncert.} G(a_p|\alpha_p, \sigma_p) \tag{6.22}$$

where $\text{Pois}(n_{cb}|\nu_{cb})$ is the Poisson probability of observing n_{cb} events in bin b of category c given the expected number of events:

$$\nu_{cb}(\phi_p, \alpha_p, \gamma_b) = \sum_{s \in \text{samples}} \lambda_{cs} \gamma_{cb} \phi_{cs}(\boldsymbol{\alpha}) \eta_{cs}(\boldsymbol{\alpha}) \sigma_{csb}(\boldsymbol{\alpha}) \tag{6.23}$$

λ_{cs} is the measured integrated luminosity, used to normalize the event yields in the simulated samples. γ_{cb} is the statistical uncertainty in bin b of category c. ϕ_{cs} is the product of unconstrained normalization factors for a given sample within a given category. In this analysis, the normalization of the $Z \to \tau\tau$ and multijet backgrounds are free ϕ_{cs} parameters that are constrained by the fit to the data. $\eta_{cs}(\boldsymbol{\alpha})$ are parametrized normalization uncertainties for a given sample in a given category (a factor around 1) and $\sigma_{csb}(\boldsymbol{\alpha})$ are parametrized shape uncertainties. $G(a_p|\alpha_p, \sigma_p)$ is a Gaussian constraint term describing an auxiliary measurement a_p (such as the measurement of the uncertainty on the tau energy scale) that constrains the nuisance parameter α_p (such as the tau energy scale):

$$G(a_p|\alpha_p, \sigma_p) = \frac{1}{\sqrt{2\pi\sigma_p^2}} \exp\left[-\frac{(a_p - \alpha_p)^2}{2\sigma_p^2}\right] \tag{6.24}$$

where $a_p = 1$ corresponds to the nominal value and the α_p parameter is the allowed to float relative to the nominal value in the fit to data. σ_p is the uncertainty on the auxiliary measurement. α_p and a_p are scaled for each parameter such that the Gaussian has unit variance. Statistical uncertainties are modelled by gamma distributions, a generalization of the Poisson distribution allowing non-integer inputs.

Statistical Significance

To test a hypothesized value of μ, the *profile likelihood ratio* is constructed:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \ . \tag{6.25}$$

where $\hat{\hat{\boldsymbol{\theta}}}$ in the numerator denotes the components of $\boldsymbol{\theta}$ that maximizes L for the specified μ (the conditional maximum-likelihood estimator of $\boldsymbol{\theta}$). $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ in the denominator represent the values that unconditionally maximize the likelihood function. The presence of nuisance parameters, represented by $\boldsymbol{\theta}$, broadens the profile likelihood as a function of μ. As the size of the systematic variations are increased, the compatibility of the model with the observed data will increase, and the sensitivity of the experiment to a potential signal will decrease. The profile likelihood ratio $\lambda(\mu)$ satisfies $0 \leq \lambda(\mu) \leq 1$ with values closer to one implying better agreement between the observed data and expected background and signal models for a hypothesized value of μ. The signal strength *mu* is in general assumed to be non-negative, but in the asymptotic approximation $\hat{\mu}$ is modelled as a Gaussian and is allowed to take on negative values.

In testing the compatibility of the background-only hypothesis with the observed data, we want to quantify the probability of obtaining a result at least as signal-like as observed in data while assuming the background-only hypothesis. If this probability is sufficiently small, then the background-only hypothesis is rejected and we claim the discovery of the new signal. For this task, the following test statistic is used [109]:

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \, , \\ 0 & \hat{\mu} < 0 \, , \end{cases} \tag{6.26}$$

which will only consider a lack of agreement between the data and background-only hypothesis if $\hat{\mu} > 0$. As the observed yield in data increases above the yield expected by the background model, $\hat{\mu}$ increases and q_0 also increases. An increase in q_0 corresponds to an increase in the incompatibility between the observed data and background-only hypothesis. This incompatibility is quantified by the *p-value*, p_0, defined as the integral of the probability density function of the test statistic q_0 from the observed value to infinity:

$$p_0 = \int_{q_{0,\mathrm{obs}}}^{\infty} f(q_0|0)\, dq_0 \tag{6.27}$$

A graphical illustration of how the p-value is defined is shown in Figure 6.10.

The p-value is often converted into an equivalent significance, Z, defined such that a Gaussian distributed variable found Z standard deviations above the mean has the upper-tail probability equal to p_0:

$$Z = \Phi^{-1}(1 - p_0) \tag{6.28}$$

where Φ^{-1} is the inverse of the cumulative distribution of the standard Gaussian. The standard for claiming discovery within the particle physics community is a significance of at least $Z \geq 5\,\sigma$, corresponding to the p-value $p = 2.87 \times 10^{-7}$. One may also claim *evidence* for a new signal if the significance is at least $3\,\sigma$, corresponding to the p-value $p = 1.3 \times 10^{-3}$.

The sensitivity of an experiment is characterized by the median expected significance, and not by the actual significance it observes in data. To compute the median significance, the observed data is replaced by a representative data set, referred to as the Asimov data set [109]. The Asimov data set is derived by estimating each parameter of $\boldsymbol{\theta}$ to be the value that maximizes the likelihood function with respect to that parameter. Different signal strengths can be injected into the Asimov data set to test different hypotheses. The median and observed p-values are illustrated in Figure 6.10. The median and observed p-values are not necessarily equal because of statistical fluctuations in the observed data, or because the hypothesized signal strength does not accurately reflect the true amount of signal in the data, or because of systematic effects.

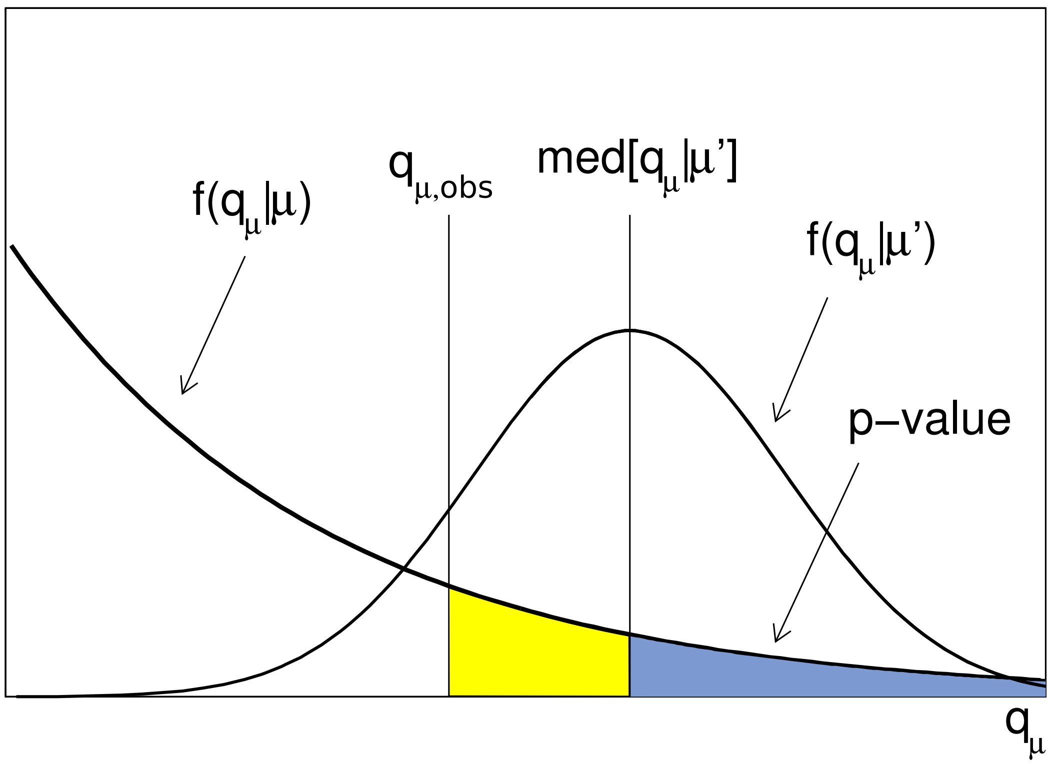


Figure 6.10: Illustration of the probability density functions of the test statistic q_μ for two values of μ. In the case of testing the compatibility of the background-only hypothesis versus the signal-plus-background hypothesis, μ = 0 and μ' = 1. The median expected and observed p-values are drawn as the filled regions under the probability density function $f(q_\mu|\mu)$ extending to infinity from the median value of the test statistic assuming the signal-plus-background hypothesis (the expected sensitivity) or the observed value, respectively. This figure has been adapted from Ref. [109].

ATLAS uses a framework of several software packages to perform the hypothesis testing discussed above. The `HistFactory` [111] package converts the histogram-based background and signal models into the probability density function represented by constructs defined in the `RooStats` [112] software library. `RooFit` [113] performs the maximum likelihood fit of the likelihood function to the observed data or the Asimov data set and uses the `MINUIT` [114] routines for the function minimization and error estimation.

6.4 Binning Optimization

The choice of binning for the BDT distributions in the VBF and Boosted categories plays an important role in optimizing the sensitivity of the analysis. The discrimination between the signal and background strengthens as progressively finer binning is selected. This increases up to a limit where the statistical uncertainty on the background estimation in each bin begins to dominate. Using too many fine bins can therefore reduce the robustness of the background estimation and unnecessarily increase the complexity of the fit since the likelihood function is a product over more bins. Too many bins might also result in bins containing no background prediction only due to the limited number of events available to model them, resulting in an overly optimistic estimate of the sensitivity.

An algorithm has been designed to determine a binning that optimizes the expected significance while respecting certain minimum background requirements in each bin. The algorithm begins with one bin from the minimum to the maximum BDT score and then determines the best insertion point for a new bin edge that maximizes the overall signal significance. Candidate bin edge locations are only considered if the new bins that would be created to the left and right of the edge would have at least five unweighted and one weighted event(s) from the $Z \to \tau\tau$ and multijet backgrounds. The algorithm repeats the same procedure of inserting an optimal bin edge between the minimum BDT score and the position of the previously inserted bin edge. This continues until the expected signal sensitivity improves by less than 1% or the background requirement cannot be satisfied by any new bin edge position. The final binning is generally sparse in regions dominated by background and more dense in the direction of the higher BDT scores. The overall signal sensitivity is mainly driven by the width of the highest bin, while the lower bins provide decreasing improvements. The expected significance using the Asimov data set and including the effects of all systematic uncertainties is 1.8 σ.

A graphical illustration of the bin edges selected by this algorithm is shown in Figure 6.11. A comparison with the expected significance achieved by using up to 20 fixed-width bins is also shown. This algorithm achieves a better expected significance with only five bins in each category. The pre-fit expected yields of all signal and background samples in each of the optimized bins are shown in Table 6.4 for the VBF category and Table 6.5 for the Boosted category.

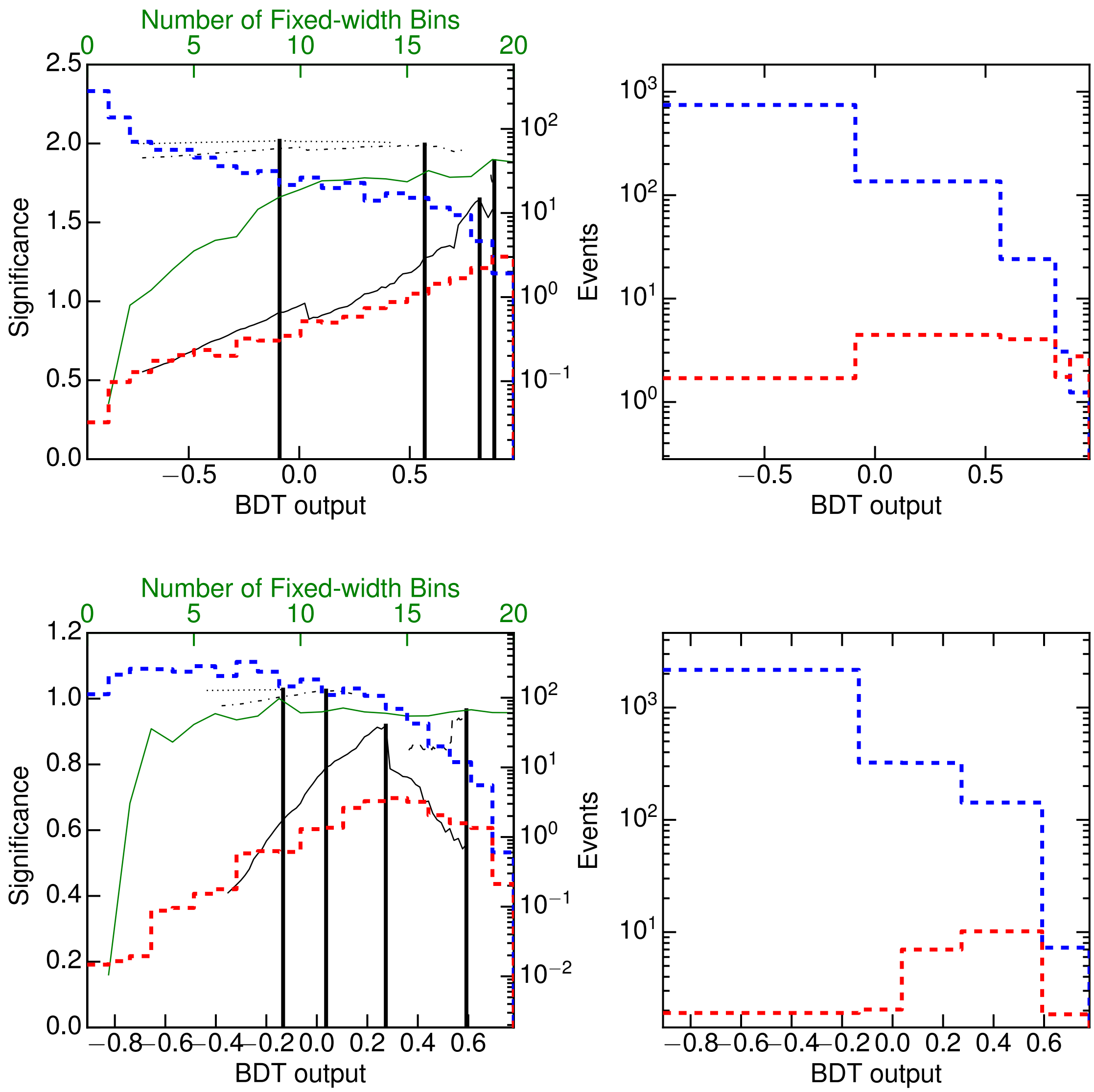


Figure 6.11: An illustration of the bin edges selected by the binning optimization algorithm in the VBF (top row) and Boosted (bottom row) categories. The blue (red) histograms are the background (signal). The thick vertical lines in the left plots are the optimized bin edge locations and the right plots show the BDT distributions histogrammed according to the optimized binning. The expected significance for *N* fixed-width bins is shown for comparison with the green line using the top x-axis in the left plots. This binning algorithm achieves a better expected significance with only 5 bins than the simple approach of using up to 20 fixed-width bins.

Sample / Bin	1	2	3	4	5
Signal Z 125					
weighted	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
unweighted	71.0 ± 8.4	100.0 ± 10.0	74.0 ± 8.6	10.0 ± 3.2	5.0 ± 2.2
Signal W 125					
weighted	0.0 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
unweighted	60.0 ± 7.7	107.0 ± 10.3	59.0 ± 7.7	10.0 ± 3.2	3.0 ± 1.7
Signal gg 125					
weighted	1.1 ± 0.1	2.4 ± 0.1	1.6 ± 0.1	0.5 ± 0.1	0.4 ± 0.1
unweighted	225.0 ± 15.0	479.0 ± 21.9	338.0 ± 18.4	102.0 ± 10.1	79.0 ± 8.9
Signal VBF 125					
weighted	0.5 ± 0.0	2.0 ± 0.0	2.3 ± 0.0	1.2 ± 0.0	2.3 ± 0.0
unweighted	1368.0 ± 37.0	4586.0 ± 67.7	5450.0 ± 73.8	2811.0 ± 53.0	5411.0 ± 73.6
Total Signal					
weighted	1.7 ± 0.1	4.5 ± 0.1	4.0 ± 0.1	1.7 ± 0.1	2.8 ± 0.1
unweighted	1724.0 ± 41.5	5272.0 ± 72.6	5921.0 ± 76.9	2933.0 ± 54.2	5498.0 ± 74.1
Ztautau					
weighted	380.2 ± 10.4	83.7 ± 4.9	13.7 ± 2.3	0.7 ± 0.4	0.8 ± 0.4
unweighted	1661.0 ± 40.8	360.0 ± 19.0	58.0 ± 7.6	5.0 ± 2.2	4.0 ± 2.0
Others					
weighted	20.5 ± 2.2	12.5 ± 1.8	4.2 ± 1.0	0.7 ± 0.2	0.1 ± 0.0
unweighted	685.0 ± 26.2	255.0 ± 16.0	74.0 ± 8.6	23.0 ± 4.8	3.0 ± 1.7
Fakes					
weighted	344.3 ± 11.4	39.8 ± 3.9	6.3 ± 1.6	1.6 ± 0.9	0.4 ± 0.4
unweighted	1302.0 ± 36.1	175.0 ± 13.2	33.0 ± 5.7	12.0 ± 3.5	1.0 ± 1.0
Total Background					
weighted	745.1 ± 15.6	136.0 ± 6.5	24.1 ± 3.0	3.1 ± 1.0	1.2 ± 0.6
unweighted	3648.0 ± 60.4	790.0 ± 28.1	165.0 ± 12.8	40.0 ± 6.3	8.0 ± 2.8
S/B	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.0	0.6 ± 0.2	2.2 ± 1.0

Table 6.4: Pre-fit signal and background content in the VBF category for each bin after the binning optimization terminates.

Sample / Bin	1	2	3	4	5
Signal Z 125					
weighted	0.1 ± 0.0	0.2 ± 0.0	0.4 ± 0.0	0.8 ± 0.0	0.2 ± 0.0
unweighted	277.0 ± 16.6	393.0 ± 19.8	946.0 ± 30.8	1707.0 ± 41.3	348.0 ± 18.7
Signal W 125					
weighted	0.2 ± 0.0	0.3 ± 0.0	0.7 ± 0.0	1.4 ± 0.0	0.3 ± 0.0
unweighted	270.0 ± 16.4	368.0 ± 19.2	908.0 ± 30.1	1709.0 ± 41.3	369.0 ± 19.2
Signal gg 125					
weighted	1.2 ± 0.1	1.8 ± 0.1	4.0 ± 0.2	6.1 ± 0.2	1.0 ± 0.1
unweighted	310.0 ± 17.6	499.0 ± 22.3	1068.0 ± 32.7	1626.0 ± 40.3	272.0 ± 16.5
Signal VBF 125					
weighted	0.2 ± 0.0	0.5 ± 0.0	1.3 ± 0.0	2.0 ± 0.0	0.4 ± 0.0
unweighted	565.0 ± 23.8	1185.0 ± 34.4	3089.0 ± 55.6	4721.0 ± 68.7	857.0 ± 29.3
Total Signal					
weighted	1.7 ± 0.1	2.8 ± 0.1	6.5 ± 0.2	10.2 ± 0.2	1.9 ± 0.1
unweighted	1422.0 ± 37.7	2445.0 ± 49.4	6011.0 ± 77.5	9763.0 ± 98.8	1846.0 ± 43.0
Ztautau					
weighted	1589.4 ± 22.5	340.8 ± 9.7	197.3 ± 7.3	96.6 ± 5.5	6.1 ± 1.5
unweighted	6014.0 ± 77.5	1552.0 ± 39.4	933.0 ± 30.5	389.0 ± 19.7	19.0 ± 4.4
Others					
weighted	43.3 ± 3.5	11.9 ± 1.9	18.8 ± 2.4	16.3 ± 2.4	0.9 ± 0.5
unweighted	608.0 ± 24.7	139.0 ± 11.8	153.0 ± 12.4	100.0 ± 10.0	11.0 ± 3.3
Fakes					
weighted	489.7 ± 14.0	60.6 ± 5.0	60.0 ± 5.0	29.5 ± 3.5	0.3 ± 0.4
unweighted	2196.0 ± 46.9	348.0 ± 18.7	338.0 ± 18.4	152.0 ± 12.3	6.0 ± 2.4
Total Background					
weighted	2122.3 ± 26.7	413.4 ± 11.1	276.1 ± 9.2	142.5 ± 6.9	7.3 ± 1.6
unweighted	8818.0 ± 93.9	2039.0 ± 45.2	1424.0 ± 37.7	641.0 ± 25.3	36.0 ± 6.0
S/B	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.0	0.3 ± 0.1

Table 6.5: Pre-fit signal and background content in the boosted category for each bin after the binning optimization terminates.

6.5 Statistical Model Validation

Before looking at the observed p-value or preferred value of μ in the data, a series of tests are performed to validate the behaviour of the fit. The model is first simplified by removing negligible shape and normalization systematics that do not affect the expected sensitivity. A fit is then performed in a control region with a very small signal contamination to confirm that the best-fit signal strength is consistent with zero. Finally, all nuisance parameters are validated by a fit to the observed data in the full signal region of each category without looking at the signal-sensitive quantities.

6.5.1 Statistical Model Simplification

An important component in ensuring that the binned likelihood fit is well-behaved involves the removal (pruning) of certain shape or normalization systematic uncertainties that have insignificant deviations from the nominal templates and the symmetrization of shape systematic uncertainties that contain one-sided variations. The removal of negligible systematic uncertainties reduces the complexity of the model without affecting the expected significance. The symmetrization procedure removes double minima or discontinuities in the likelihood function when profiled over each nuisance parameter. The model simplification is also important within the context of the full combined fit of all Higgs decay channels where the complexity of the fit is severely affected by the large number of parameters.

Systematics with shape components are first symmetrized to remove all one-sided variations. In each bin of the $\pm 1\,\sigma$ variations, if both sides of the variation are above or below the nominal template, then the size of the larger deviation is reflected with respect to the nominal template. Shape systematics are then compared with the nominal templates with two methods and removed if deemed negligible. The first method prunes shape variations that are at least 90% compatible with the nominal template according to a χ^2 test. The second method removes shape components that differ by no more than 10% of the size of the statistical uncertainty on the background model in each bin. Figure 6.12 shows the change in expected significance as the threshold on the χ^2 probability is lowered. A threshold of 90% has been selected.

Normalization components of systematics where the $\pm 1\,\sigma$ variations do not change the nominal templates by more than 0.5% are removed. This does not affect the expected significance.

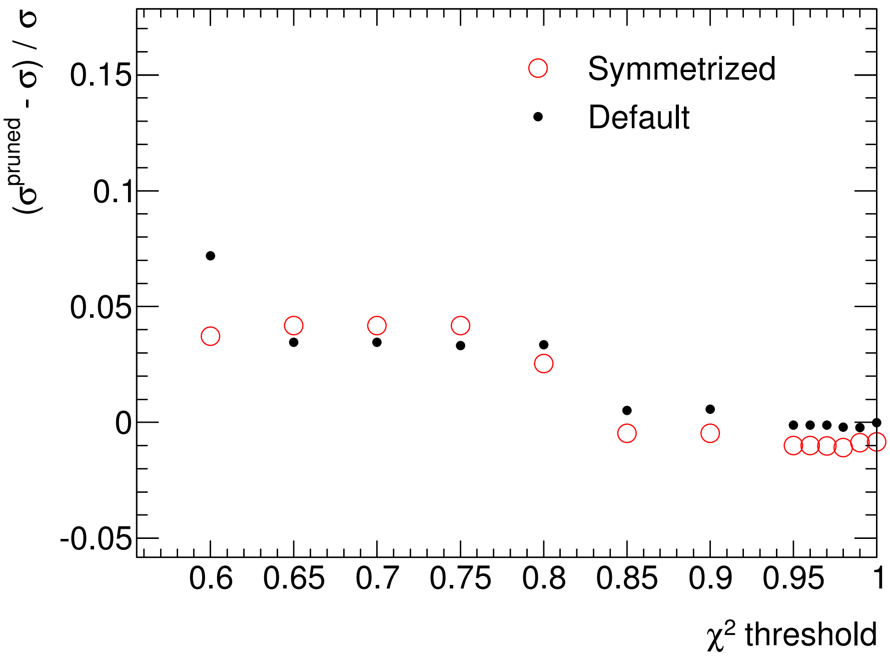


Figure 6.12: The change in expected significance as the χ^2 threshold is lowered and systematic shape variations are more aggressively pruned away. The default and symmetrized shape variations are compared.

6.5.2 Fitting in a Control Region

It is important to validate the modelling of the BDT output in a control region where little signal is expected. A model has been constructed using only events outside of the signal-sensitive region $100 < m_{\tau\tau}^{\mathrm{MMC}} < 150$ GeV and a fit to the observed data is performed. Pre-fit distributions of the BDT output are shown in Figure 6.13. The signal strength preferred by the data in this control region is consistent with zero. Figure 6.14 shows the BDT distributions in a QCD-rich control region constructed by inverting the $\Delta\eta(\tau, \tau) < 1.5$ cut in each category.

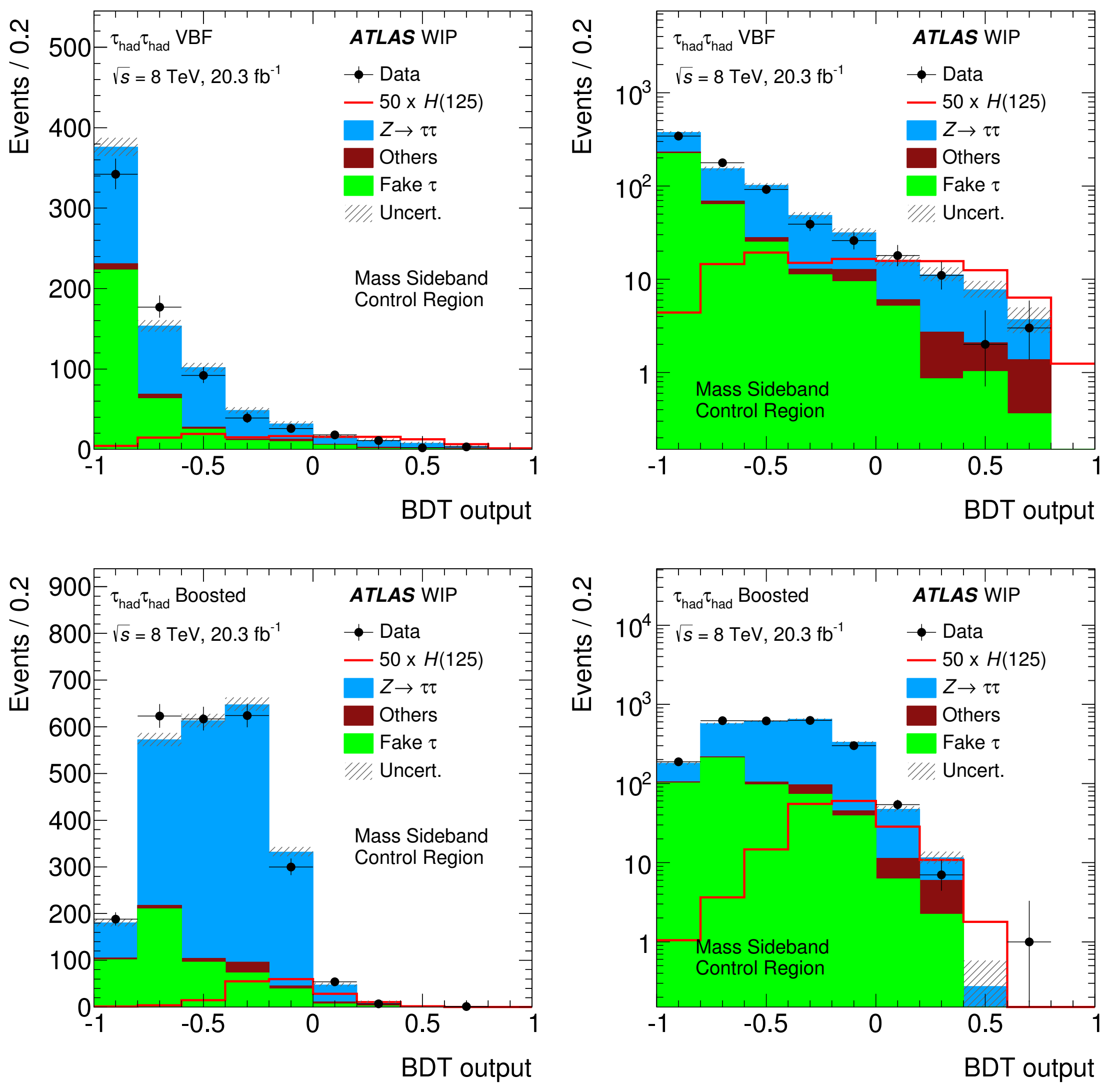


Figure 6.13: Distributions of the BDT output in an invariant mass control region that excludes the signal-sensitive region in $100 < m_{\tau\tau}^{\mathrm{MMC}} < 150$ GeV. The distributions are shown for the *VBF* (top) and *Boosted* (bottom) categories using linear (left) and logarithmic (right) y-axis scales. The contributions from a Standard Model Higgs boson with m_H = 125 GeV are superimposed, multiplied by a factor of 50. The error band includes statistical and pre-fit systematic uncertainties.

6.5.3 Fit Behaviour

An unconditional fit is performed to the observed data and the post-fit normalizations of the $Z \to \tau\tau$ and multijet backgrounds are 1.01 ± 0.11 and 1.00 ± 0.04, respectively. This confirms that the pre-fit normalizations derived from the fit of $\Delta\eta(\tau, \tau)$ at preselection still holds in the final fit of the BDT distributions in the VBF and Boosted categories along with the $\Delta\eta(\tau, \tau)$ distribution in the Rest category.

The deviations of all nuisance parameters from their nominal values after the fit is summarized by the black markers in Figure 6.15. A few of the tau energy scale parameters are pulled, but not more than $0.5\,\sigma$. The size of the black error bars represents the post-fit uncertainties on each nuisance parameter. A nuisance parameter is constrained by the observed data in the fit if the size of the black error bar is less than the size of the yellow bands, which correspond to the pre-fit $\pm 1\,\sigma$ uncertainties. The only nuisance parameter that is significantly constrained is the shape uncertainty on the multijet background estimate `ANA_HH_2012_QCD`. This constraint is expected since the pre-fit shape uncertainty is not a priori defined by a physical $\pm 1\,\sigma$ variation. The nuisance parameters in Figure 6.15 are also ranked top to bottom by decreasing impact on the uncertainty of $\hat{\mu}$, as explained in the caption. The most important nuisance parameters include the jet η modelling component of the jet energy scale (important for the jet-related features in the VBF category, the most sensitive category), the shape uncertainty on the multijet background, and a component of the tau energy scale.

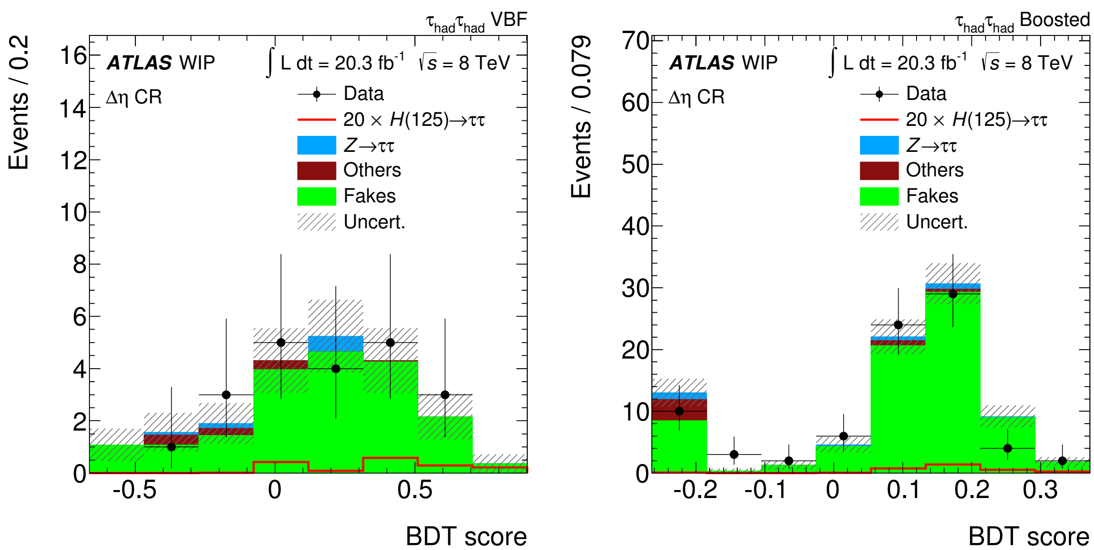


Figure 6.14: Distributions of the BDT output in a QCD-rich control region where the $\Delta\eta(\tau, \tau) < 1.5$ cut in each category is inverted. The distributions are shown for the *VBF* (left) and *Boosted* (right) categories. The contributions from a Standard Model Higgs boson with m_H = 125 GeV are superimposed, multiplied by a factor of 20.

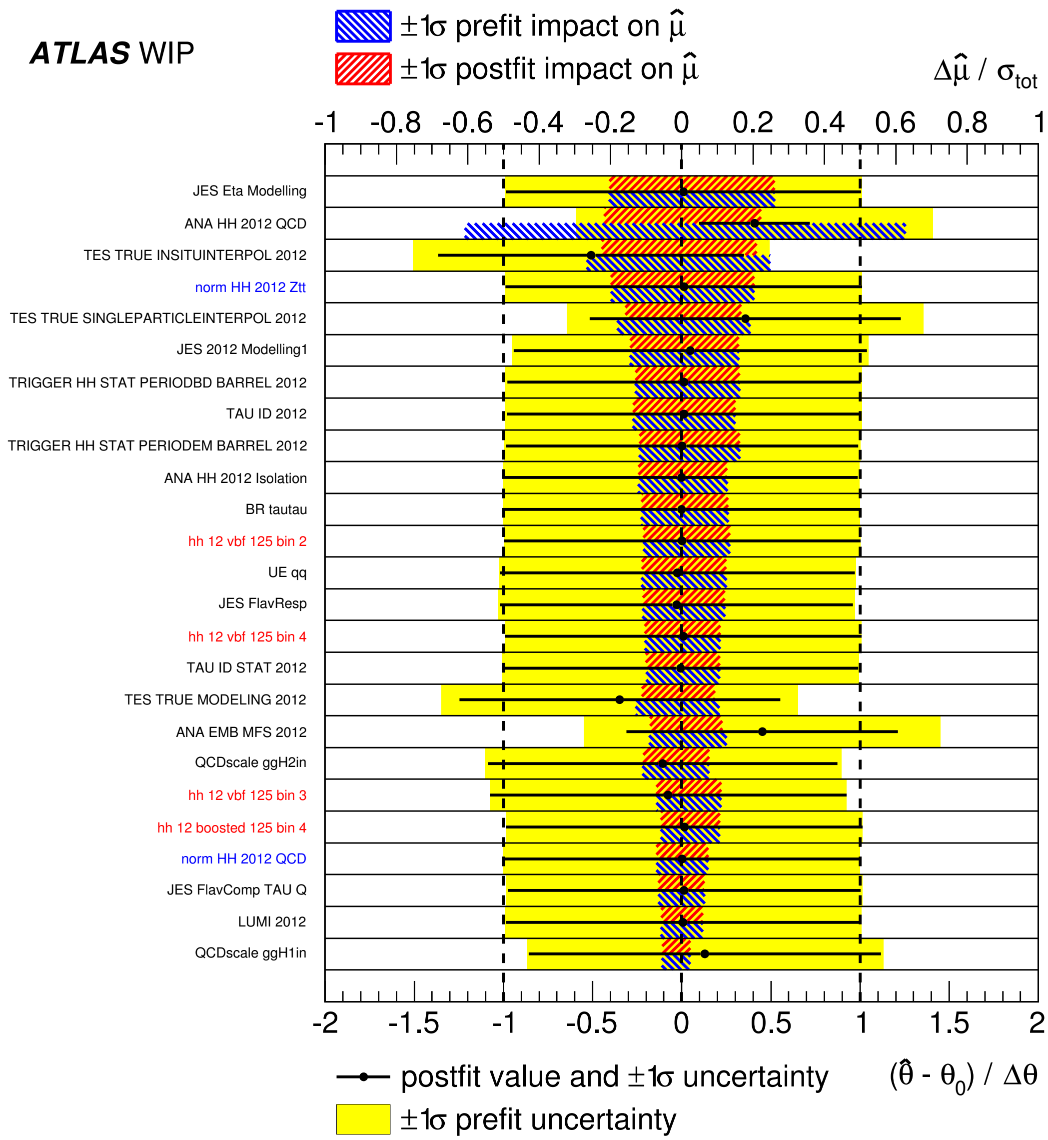


Figure 6.15: Impact of systematic uncertainties on the fitted signal-strength parameter $\hat{\mu}$. The systematic uncertainties are listed in decreasing order of their impact on $\hat{\mu}$ on the *y*-axis and only the top 25 parameters are shown. The hatched blue and red bands show the variations of $\hat{\mu}$ with respect to the total error on μ, σ_{tot}, referring to the top *x*-axis, when fixing the corresponding individual nuisance parameter θ to its post-fit value $\hat{\theta}$ modified upwards or downwards by its post-fit uncertainty, and repeating the fit. The filled circles, referring to the bottom *x*-axis, show the pulls of the fitted nuisance parameters, i.e. the deviations of the fitted parameters $\hat{\theta}$ from their nominal values θ_0, normalised to their nominal uncertainties $\Delta\theta$. The black lines show the post-fit uncertainties of the nuisance parameters, relative to their nominal uncertainties, which are indicated by the yellow band.

Chapter 7

Results

This chapter presents the results of the $H \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ search as well as the combination with the searches in the semi-leptonic and fully leptonic $H \rightarrow \tau\tau$ channels. Now that the statistical model has been validated, the data is unblinded and the real test of compatibility between the observed data and background-only hypothesis can be performed. The final Run I ATLAS result also includes the same multivariate analyses applied on the 7 TeV data. The inclusion of the 7 TeV data does not significantly affect the result and has not been documented in this thesis, however, when quoting values or displaying plots derived from the global combined fit, the 7 TeV data is included.

Simulated signal samples were generated for Higgs masses of 100 to 150 GeV in steps of 5 GeV, but the best estimate of the Higgs mass with the ATLAS detector is 125.36 $\pm$ 0.37 (stat) $\pm$ 0.18 (syst) GeV which has been obtained from a combination of measurements [115] in the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{*}$ channels. The final results of the tau channel are therefore interpolated between the 125 and 130 GeV Higgs masses to quote values at the same m_H = 125.36 GeV.

7.1 Signal Strength Measurement and Observed Significance

An unconditional fit of the binned likelihood function to the observed data is performed, and the signal strength μ is measured at the interpolated Higgs mass m_H = 125.36 GeV. The multivariate analysis measures the value $\mu = 1.77^{+0.93}_{-0.71}$ and the cut-based analysis measures the value $\mu = 1.64^{+0.90}_{-0.74}$. The uncertainties are the combined statistical and systematic uncertainties. The probability p_0 of obtaining a result at least as signal-like as observed in the data if no signal were present is calculated using the profile likelihood ratio test statistic $q_{\mu=0} = -2\ln(\mathcal{L}(0,\hat{\hat{\vec{\theta}}})/\mathcal{L}(\hat{\mu},\hat{\vec{\theta}}))$ in the asymptotic approximation [109] as discussed in Section 6.3. The observed p_0 value is 1.8×10^{-3} for the multivariate analysis, corresponding to a deviation from the background-only hypothesis of 2.9 σ, and 9.3×10^{-3} for the cut-based analysis, corresponding to a deviation of 2.4 σ. The validity of the asymptotic approximation is tested by evaluating the profile likelihood ratio for the background-only and signal-plus-background hypotheses over 10,000 and 1000 pseudo-experiments, respectively. The observed test statistic value in data confirms a deviation of 2.9 σ.

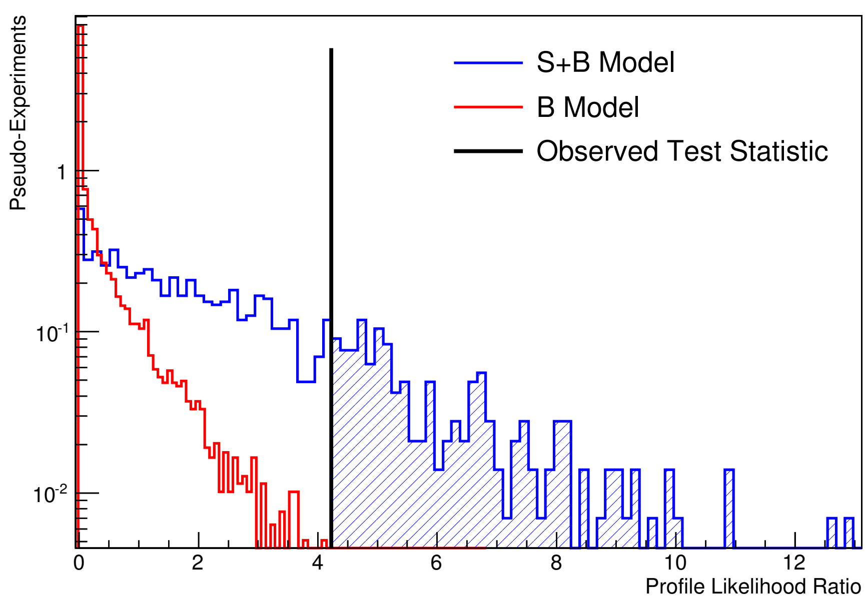


Figure 7.1: The multivariate analysis profile likelihood ratio observed in data and evaluated for the background-only and signal plus background hypotheses.

The overall combination of the multivariate analyses for all channels in the 7 TeV and 8 TeV data results in a measured signal strength of:

$$\mu = 1.42\ ^{+0.27}_{-0.26}(\text{stat.})\ ^{+0.32}_{-0.24}(\text{syst.})\ \pm 0.10(\text{theory syst.}).$$

The systematic uncertainties are split into two groups, theoretical uncertainties on the inclusive Higgs boson production cross section and $H \to \tau\tau$ branching ratio, and all other systematic uncertainties including all experimental effects as well as theoretical uncertainties on the signal region acceptance from the QCD scale and PDF choice. The observed p_0 value of the combined result is 3.0×10^{-6}, which corresponds to an excess of $4.5\,\sigma$. This can be compared to an expected significance of $3.5\,\sigma$. The 8 TeV combined signal strength measurement made by the cut-based analyses is $\mu = 1.37^{+0.57}_{-0.48}$, corresponding to an excess of $3.2\,\sigma$. These excesses observed separately by the combined multivariate and cut-based analyses exceed the threshold to claim evidence that the recently discovered Higgs boson decays into tau leptons. This therefore constitutes direct evidence that the Higgs boson couples to fermions.

In the unconditional fit of the binned likelihood function, the nuisance parameters describing the systematic uncertainties and unconstrained normalization parameters for the $Z \to \tau\tau$ and multijet backgrounds are allowed to be pulled by the observed data. The post-fit yields and uncertainties for the signal and background components over all BDT bins and for the two highest bins in the VBF and Boosted categories are shown in Table 7.1. The highest VBF bin expects roughly one background event and observes six events in data. Post-fit histograms of the BDT distributions in the VBF and Boosted categories are shown in Figure 7.3. The bins have been remapped such that they have equal width to better show the signal-sensitive bins. All BDT bins are rearranged in bins of $\log_{10}$ (S/B) in Figure 7.4 Here, S/B is the signal-to-background ratio calculated assuming $\mu = 1.4$. The expectation is shown for signal yields for both $\mu = 1$ and the best-fit value $\mu = 1.4$ for $m_H = 125$ GeV on top of the background prediction taken also from the best-fit values. The background expectation where the signal strength parameter has been fixed to $\mu = 0$ is also shown for comparison. Figure 7.5 displays the VBF event in data with the highest BDT score. This event exhibits the signature VBF topology with two forward jets and two central tau candidates. The invariant di-tau mass estimate for this event is 123 GeV.

To separately probe the relative contributions from the different production mechanisms, the signal strength parameter is split into one parameter $\mu^{\tau\tau}_{VBF+VH}$ scaling the predicted rates of the vector-boson-mediated *VBF* and *VH* processes, and another parameter $\mu^{\tau\tau}_{ggF}$ scaling the gluon-mediated ggF process. These parameters are fitted separately in the data using the full $H \to \tau\tau$ combination. The two-dimensional likelihood contours in the plane of $\mu^{\tau\tau}_{ggF}$ and $\mu^{\tau\tau}_{VBF+VH}$ [116] are shown in figure 7.2 for m_H = 125.36 GeV. The best-fit values are in agreement with the predictions from the Standard Model:

$$\mu^{\tau\tau}_{ggF} = 1.93\ ^{+0.78}_{-0.77}(\text{stat.})\ ^{+1.19}_{-0.80}(\text{syst.})\ \pm 0.29(\text{theory syst.}).$$

$$\mu^{\tau\tau}_{VBF+VH} = 1.24^{+0.48}_{-0.45}(\text{stat.})\ ^{+0.31}_{-0.28}(\text{syst.})\ \pm 0.08(\text{theory syst.}).,$$

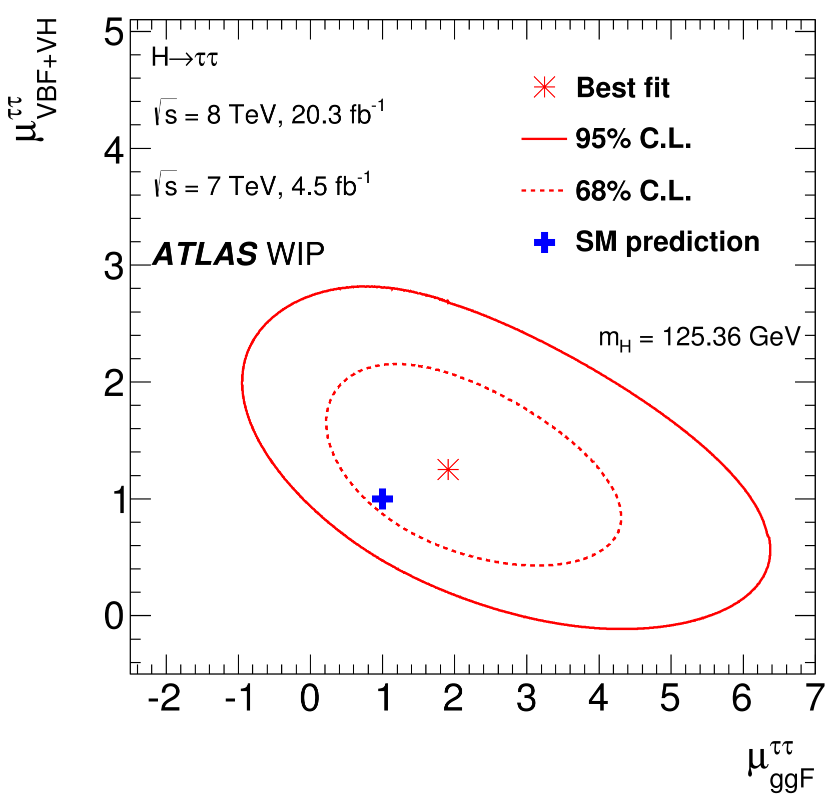


Figure 7.2: Likelihood contours for the combination of all channels in the ($\mu^{\tau\tau}_{ggF}$, $\mu^{\tau\tau}_{VBF+VH}$) plane. The 68% and 95% CL contours are shown as dashed and solid lines respectively, for m_H = 125.36 GeV. The SM expectation is shown by a filled plus symbol. The best fit to the data is shown for the case when both the $\mu^{\tau\tau}_{ggF}$ and $\mu^{\tau\tau}_{VBF+VH}$ are unconstrained.

Process/Category	VBF			Boosted		
BDT output bin	all bins	second last bin	last bin	all bins	second last bin	last bin
Multijet	370 ± 18	2.3 ± 0.9	0.6 ± 0.3	644 ± 26	35 ± 4	0.6 ± 0.3
Others	37 ± 5	0.7 ± 0.2	< 0.1	90 ± 12	16.1 ± 2.1	0.9 ± 0.2
$Z \to \tau\tau$	475 ± 16	0.6 ± 0.7	0.6 ± 0.4	2230 ± 70	92 ± 4	5.4 ± 1.6
$ggF:\ H \to \tau\tau$ (m_H = 125 GeV)	8.1 ± 2.7	0.7 ± 0.2	0.6 ± 0.2	21 ± 8	9.0 ± 3.3	1.6 ± 0.6
$VBF:\ H \to \tau\tau$	12.0 ± 3.2	1.8 ± 0.5	3.4 ± 0.9	6.3 ± 1.7	2.9 ± 0.7	0.5 ± 0.1
$WH: H \to \tau\tau$	0.3 ± 0.1	< 0.1	< 0.1	4.0 ± 1.1	1.9 ± 0.5	0.4 ± 0.1
$ZH:\ H \to \tau\tau$	0.2 ± 0.1	< 0.1	< 0.1	2.4 ± 0.6	1.1 ± 0.3	0.2 ± 0.1
Total signal	21 ± 5	2.5 ± 0.6	4.0 ± 1.0	34 ± 10	15 ± 4	2.7 ± 0.8
Total backgrounds	882 ± 18	3.6 ± 1.3	1.2 ± 1.1	2960 ± 50	143 ± 6	7.0 ± 1.8
Data	892	5	6	3020	161	10

Table 7.1: The predicted post-fit event yields for m_H = 125 GeV for the total number of events and for the two highest bins of the BDT distributions in the VBF and Boosted categories of the $\tau_{\text{had}}\tau_{\text{had}}$ channel. The background normalisations, signal normalisation, and uncertainties reflect the preferred values from the combined fit of all tau decay channels. The uncertainties on the total background and total signal reflect the full statistical and systematic uncertainty, while the uncertainties on the individual background components reflect the full systematic uncertainty only.

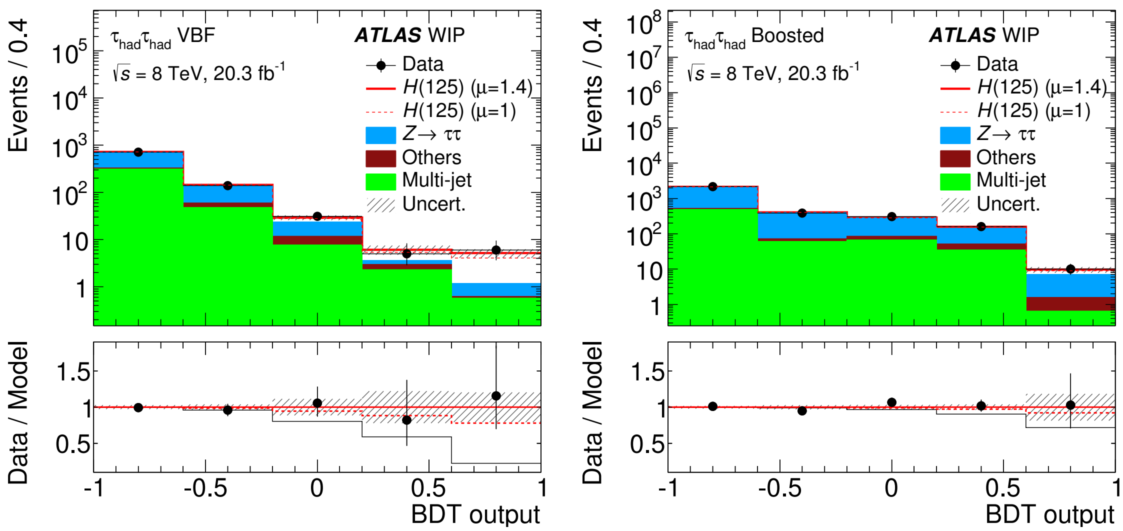


Figure 7.3: Distributions of the BDT discriminants for the data taken at $\sqrt{s}$ = 8 TeV in the signal regions of the *VBF* (left) and *Boosted* (right) categories. The Higgs boson signal (m_H = 125 GeV) is shown stacked with a signal strength of μ = 1 (dashed line) and μ = 1.4 (solid line). The background predictions are determined in the global fit (that gives μ = 1.4). The size of the statistical and systematic normalisation uncertainties is indicated by the hashed band. The ratios between the data and the model (background plus Higgs boson contributions with μ = 1.4) are shown in the lower panels. The dashed red and the solid black lines represents the changes in the model when μ = 1.0 or background only are assumed respectively.

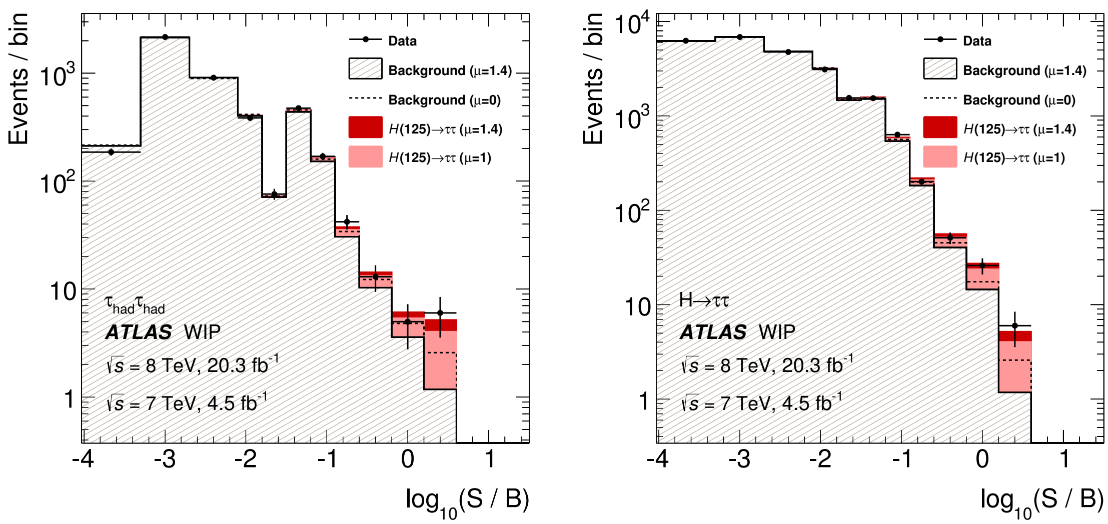


Figure 7.4: Event yields as a function of $\log_{10}(S/B)$, where S (signal yield) and B (background yield) are taken from the BDT output bin of each event, assuming $\mu = 1.4$. Events in all categories are included. The predicted background is obtained from the global fit (with $\mu = 1.4$) and signal yields are shown for $m_H = 125$ GeV, at $\mu = 1$ and $\mu = 1.4$ (the best-fit value). The background only distribution (dashed line) is obtained from the global fit, but fixing $\mu = 0$.

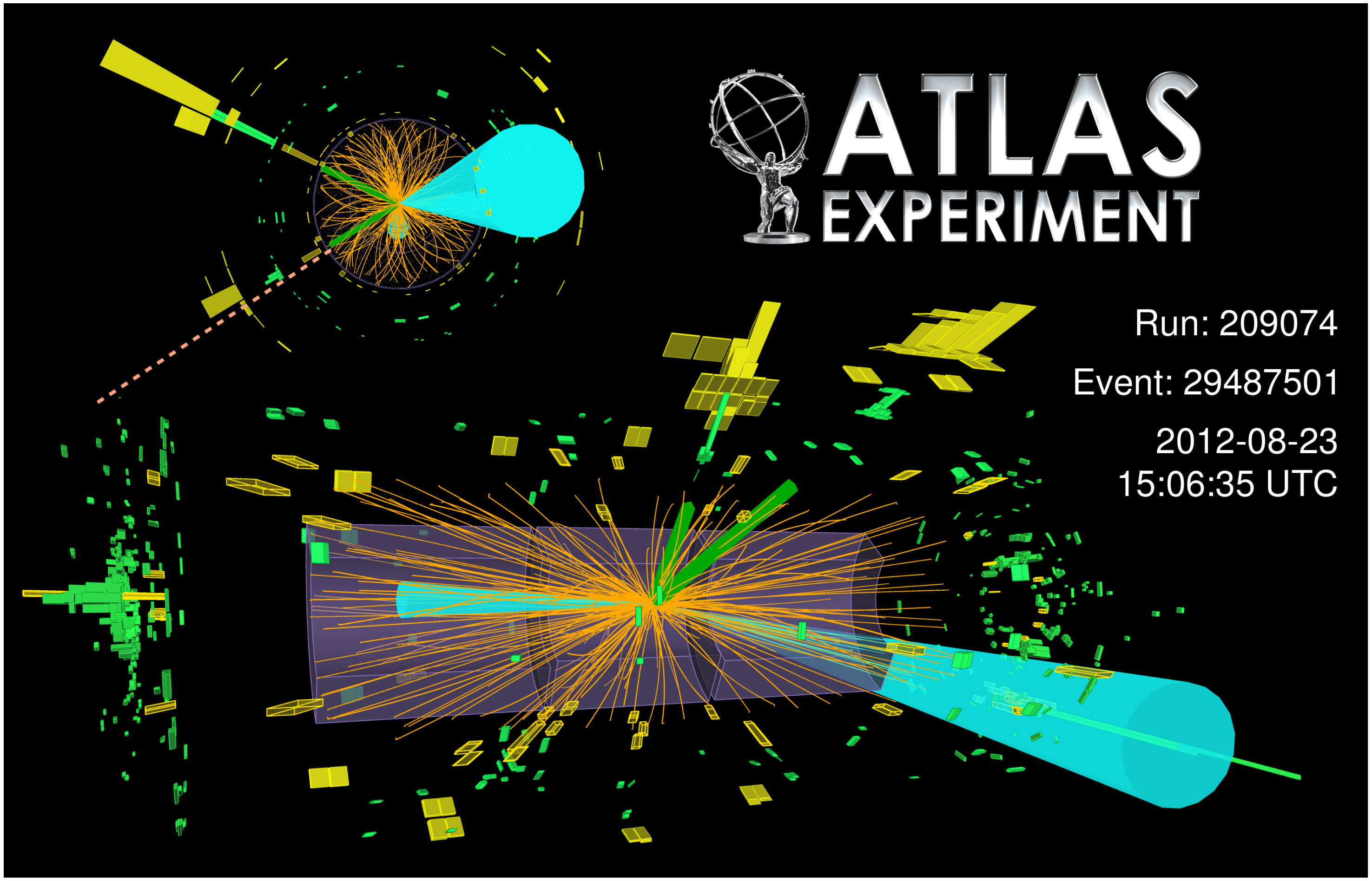


Figure 7.5: Display of the VBF event in data with the highest BDT output value of 0.91. The hadronically decaying taus are indicated by green tracks. The dashed line in the lower left quadrant of the $R - \phi$ view represents the direction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector, and there are two VBF jets marked with turquoise cones. The leading tau p_{T} is 122 GeV, the sub-leading tau p_{T} is 67 GeV, $E_{\mathrm{T}}^{\mathrm{miss}}$ = 72 GeV, m_{j_1,j_2} = 1.02 TeV and $m_{\tau\tau}^{\mathrm{MMC}}$ = 123 GeV.

7.2 Mass Compatibility

In order to visualise the compatibility of the excess of events above the background prediction with the SM Higgs boson at m_H = 125 GeV, a weighted distribution of events as a function of $m_{\tau\tau}^{\text{MMC}}$ is shown in figure 7.6. The events are weighted by a factor of $\ln(1 + S/B)$ where the signal-to-background ratio is evaluated in the bins of the BDT distribution. This effectively gives larger weights to events in the high BDT bins and suppresses the background-dominated low BDT bins. The excess of events in these mass distributions is consistent with the expectation for a Standard Model Higgs boson with m_H = 125 GeV. The distributions for the predicted excess in data over the background are also shown for alternative SM Higgs boson mass hypotheses of m_H = 110 GeV and m_H = 150 GeV. The data favour a Higgs boson mass of m_H = 125 GeV and are less consistent with the other masses considered.

A test of the mass compatibility is also performed by the cut-based analyses. Although the cut-based analyses have a lower sensitivity to the presence of a signal, they provide a more unbiased estimate of the mass. A two-dimensional likelihood fit for the signal strength μ and the mass m_H has been performed. The mass points are tested in steps of 5 GeV in the range between 100 GeV and 150 GeV. The best fit value is found at μ = 1.4 and m_H = 125 GeV. The result is shown in the (m_H, μ) plane in figure 7.7 together with the 68% and 95% CL contours. This result also indicates that the observed signal is compatible with a Higgs boson with a mass of 125 GeV.

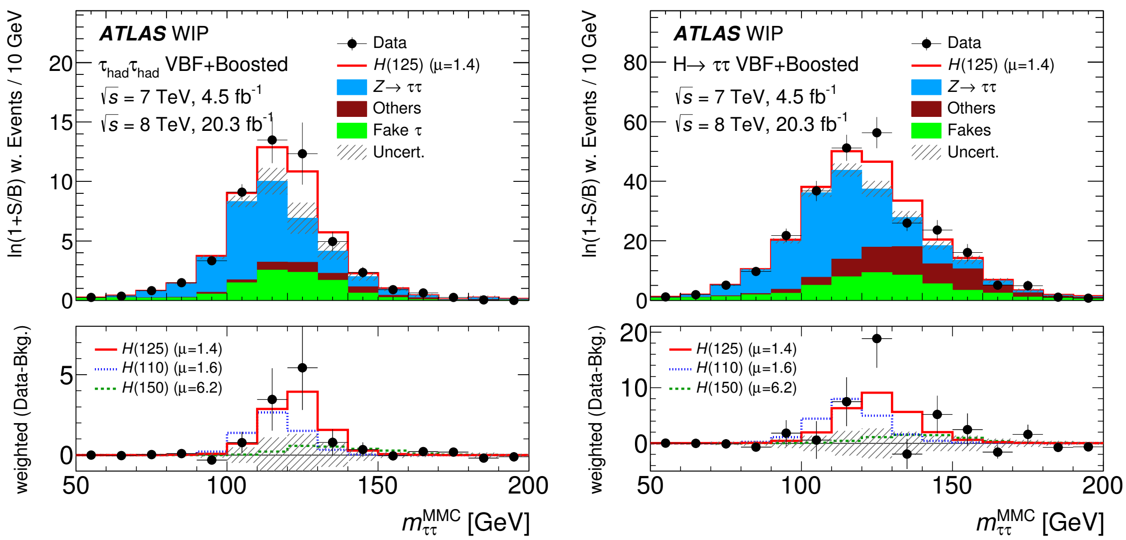


Figure 7.6: Distributions of $m_{\tau\tau}^{\mathrm{MMC}}$ where events are weighted by $\ln(1+S/B)$ for all channels. These weights are determined by the signal (S) and background (B) predictions for each BDT bin. The bottom panel in each plot shows the difference between weighted data events and weighted background events (black points), compared to the weighted signal yields. The background predictions are obtained from the global fit with the m_H = 125 GeV signal hypothesis (μ = 1.4). The m_H = 125 GeV signal is plotted with a solid red line, and, for comparison, signals for m_H = 110 GeV (blue) and m_H = 150 GeV (green) are also shown.

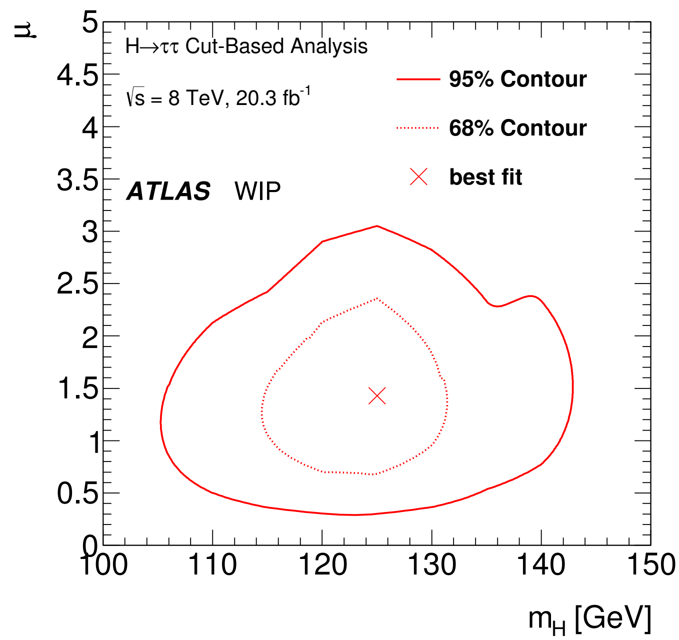


Figure 7.7: The results of the two-dimensional likelihood fit in the (m_H, μ) plane for the cut-based analysis for the data taken at $\sqrt{s}$ = 8 TeV. The 68% and 95% CL contours are shown as dashed and solid lines respectively. The best-fit value is indicated as a red cross.

7.3 Compatibility of the Multivariate and Cut-based Analyses

The signal strengths measured by the multivariate and cut-based analyses are clearly compatible within uncertainty, but with the possibility of there being a much larger difference, a method was developed to properly quote an uncertainty on $\Delta\mu = \mu_{\text{MVA}} - \mu_{\text{CBA}}$. The analyses are correlated since the signal regions of the cut-based analysis is selecting a subset of the events in the multivariate analysis. The correlation coefficient ρ must be determined to compute the uncertainty on the difference:

$$\delta(\Delta\mu)^2 = (\delta\mu_{\text{MVA}})^2 + (\delta\mu_{\text{CBA}})^2 - 2\rho \times \delta\mu_{\text{MVA}}\delta\mu_{\text{CBA}}$$

The bootstrap [117] method is a very powerful technique of estimating statistical parameters by only using the existing data. A new observed data set of the same size is generated by resampling the original data set *with replacement*. The bootstrapped data set is not necessarily the same as the original since events can be duplicated or omitted. Unconditional fits of the multivariate and cut-based likelihood functions then produce new bootstrapped estimates of μ. This has been repeated 10,000 times to obtain a robust estimate of the correlation between μ_{MVA} and μ_{CBA}. The correlation coefficient is determined from the covariance matrix of μ_{MVA} and μ_{CBA}:

$$\rho = \frac{\text{cov}(\mu_{\text{MVA}}, \mu_{\text{CBA}})}{\sqrt{\text{var}(\mu_{\text{MVA}}) \times \text{var}(\mu_{\text{CBA}})}}$$

Histograms of the bootstrapped signal strengths are shown in Figure 7.8 where the bottom left plot displays the correlation.

For simplicity, the bootstrapping has been performed at the Higgs mass of 125 GeV and the uncertainties on the signal strengths are the symmetric `MINUIT` [114] errors:

$$\mu_{\text{MVA}} = 1.75 \pm 0.79 \qquad \mu_{\text{CBA}} = 1.61 \pm 0.79$$

Taking the bootstrapped correlation coefficient of $\rho = 0.53$, the difference with uncertainty is consistent with zero:

$$\Delta\mu = 0.14 \pm 0.59$$

The bootstrap method can automatically perform full error propagation, and the bottom right plot in Figure 7.8 shows a Gaussian fit of $\mu_{\text{MVA}} - \mu_{\text{CBA}}$ histogrammed over all bootstrap iterations. The width of this Gaussian is similar to the uncertainty on $\Delta\mu$ quoted above.

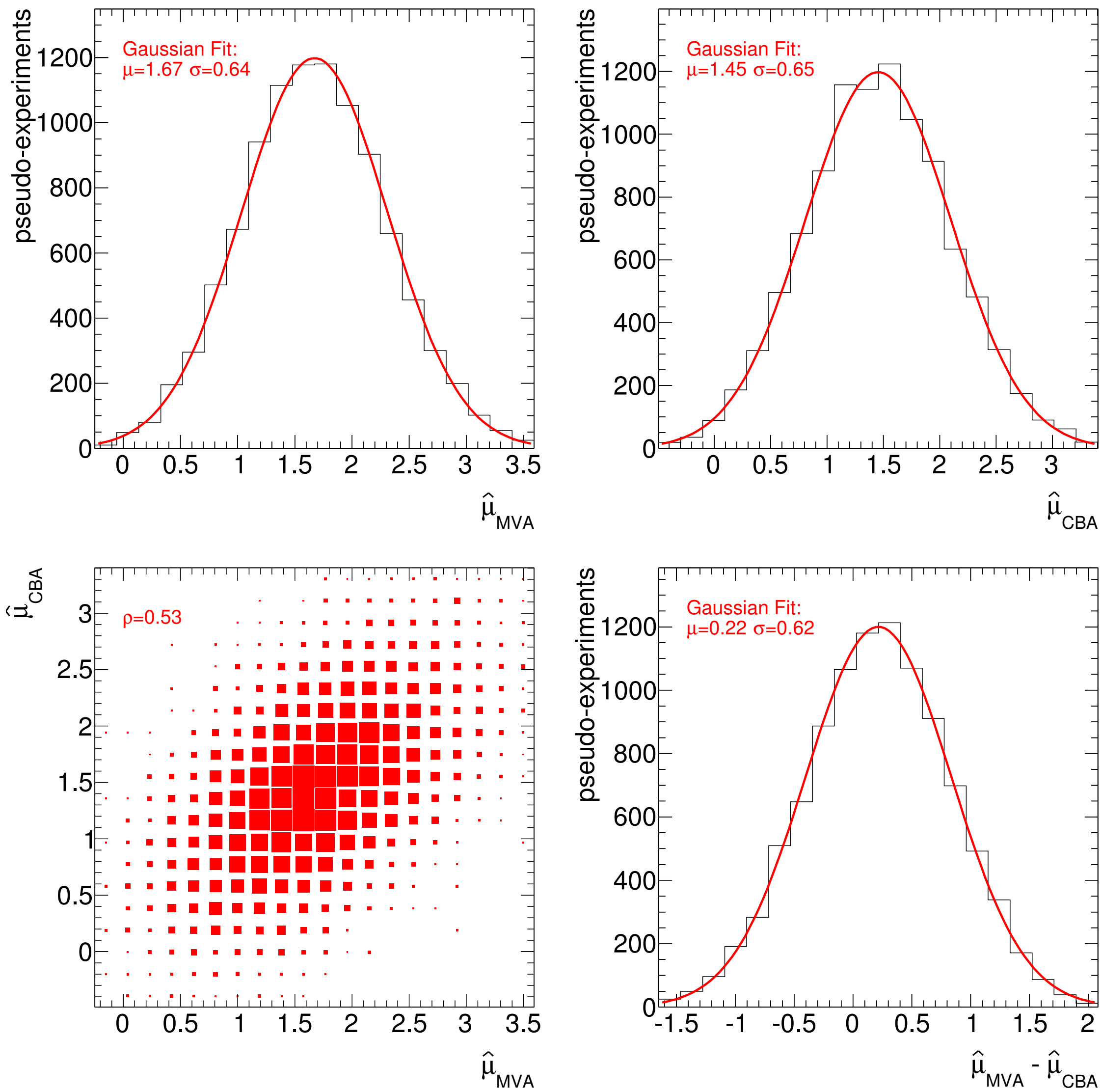


Figure 7.8: The best-fit $\hat{\mu}$ in the cut-based analysis (CBA) and the multivariate analysis (MVA) for 10,000 bootstrapped pseudo-experiments. The bottom left plot displays the correlation between the analyses.

Chapter 8

Conclusions

This thesis has presented a search for $H \to \tau_{\text{had}}\tau_{\text{had}}$ as part of a combined search in all tau decay channels. The tau channel is especially promising at low Higgs masses but is made challenging by overwhelming backgrounds. In comparison with the fully leptonic and semi-leptonic final states, the fully hadronic final state benefits from both a high branching ratio and a better resolution on the di-tau mass estimate because of the presence of only two neutrinos. The performance of the tau trigger and tau identification have been critical in making a search in this channel possible. Prior to undertaking this search, significant efforts were made in pioneering a multivariate tau identification to reach the required levels of fake tau rejection.

Following the discovery in 2012 of a Higgs boson with a mass of about 125 GeV in the $\gamma\gamma$, ZZ^*, and WW^* decay channels, the goal of this analysis was to determine if this Higgs boson also decays to taus, and thus provide *direct* confirmation of a coupling to fermions. The sensitivity of this search has therefore been optimized for m_H = 125 GeV. Data-driven background estimation techniques have been used to model the two main backgrounds from $Z \to \tau\tau$ and QCD multijet production, and a multivariate analysis using BDTs was optimized in two categories designed to separately probe the VBF and ggF Higgs production mechanisms. A cut-based analysis fitting the di-tau invariant mass has also been developed to support the multivariate result and to probe a range of Higgs boson masses, although with a more limited sensitivity.

An excess of events over the expected background is found, with an observed (expected) significance of 2.9 (1.8) standard deviations and a measured signal strength of μ = 1.7 $\pm$ 0.8. The compatibility of the multivariate and cut-based results at m_H = 125 GeV is demonstrated with the bootstrap method. In combination with the fully leptonic and semi-leptonic final states, the observed (expected) significance reaches 4.5 (3.2) standard deviations with a measured signal strength of μ = 1.5 $\pm$ 0.5. This excess constitutes direct evidence that the recently discovered Higgs boson couples to fermions.

In the coming years the LHC and detectors will undergo a series of upgrades to deliver and cope with ever higher collision energies and luminosities. Run II of the LHC constitutes "Phase 0" of the LHC upgrade plan and is scheduled to begin with first beams circulating in the spring of 2015. The center-of-mass energy will be increased to 13-14 TeV. The proton bunch spacing will also be decreased from 50ns to 25ns. In addition to the potential gain in sensitivity from the increase in the Higgs production cross sections at the higher collision energy, the LHC will operate at the higher luminosity of 10^{34}cm^{-2}s^{-1} and is estimated to deliver up to 100 fb^{-1} by the year 2018. Following the "Phase 1" upgrades, the LHC will deliver over 300 fb^{-1} by the year 2022. Although the dates are merely estimates and plans could certainly change, the "Phase 2" upgrade will launch the High Luminosity LHC (HL-LHC), aiming to increase the luminosity to 10^{35}cm^{-2}s^{-1} and deliver 3000 fb^{-1} over several years.

The forthcoming LHC runs present many exciting possibilities as well as new challenges in the future analysis of $H \to \tau\tau$ with the ATLAS detector. Significant improvements on the Run I $H \to \tau\tau$ results will of course be possible along with the ability to claim discovery in the tau channel alone. Beyond the scope of the Run I analysis, ATLAS will have the potential to perform measurements of the Higgs mass, CP state, and spin using the di-tau final state. This will not come easily, however, as the higher luminosity, and increased pileup and underlying event, will present serious challenges throughout the ATLAS trigger and data acquisition systems, and event reconstruction.

Increased activity in the very forward regions of the ATLAS detector, especially in the HL-LHC regime, will adversely affect the jet selection in the VBF category, degrading the sensitivity. These effects could be suppressed with an increase in the pseudorapidity coverage of the ATLAS tracking system, extending the reach of cuts on the jet-vertex fraction to reject jets from pileup and the underlying event. Work is well underway to optimize the tau triggers for Run II and the next phase of tau reconstruction and identification will take advantage of new algorithms analyzing the substructure of tau decays. An improvement in the tau energy resolution and decay classification should correspondingly improve the resolution on the estimate of the di-tau mass. With a larger data sample it might even be possible to exploit the fully hadronic di-tau final state where both decays are 3-prong. With the reconstruction of a secondary vertex for each 3-prong tau decay, the directions of the original taus are known with respect to the primary vertex, and the di-tau invariant mass can be determined *analytically* up to a two-fold ambiguity [118, 119]. The subset of events in which analytical solutions exist is, however, severely limited by the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution.

More protons circulating the LHC means fresh minds exchanging new ideas, and there is no telling what unforeseen achievements will become reality along the path toward the next revolutionary discoveries.

"When you see someone putting on his Big Boots,
you can be pretty sure that an Adventure is going to happen."

— A.A. Milne
WINNIE-THE-POOH

Bibliography

[1] The ATLAS Collaboration. *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*. *Phys. Lett. B* 716.CERN-PH-EP-2012-218 (2012), p. 1. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex] [see p. 1].

[2] The CMS Collaboration. *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*. *Phys. Lett. B* 716.CMS-HIG-12-028, CERN-PH-EP-2012-220 (2012), p. 30. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex] [see p. 1].

[3] Wikimedia Commons. *Standard Model of Elementary Particles*. 2006. URL: http://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg [see p. 4].

[4] S Heinemeyer, C Mariotti, G Passarino, and R (Eds.) Tanaka. *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*. arXiv:1307.1347. CERN-2013-004. Geneva, 2013 [see pp. 5, 9, 10, 12, 98].

[5] S Dittmaier, C Mariotti, G Passarino, and R (Eds.) Tanaka. *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*. Geneva: CERN, 2011 [see pp. 6, 7, 12].

[6] J. Beringer et al. *Review of Particle Physics*. *Phys. Rev. D* 86 (2012), p. 010001. DOI: 10.1103/PhysRevD.86.010001 [see pp. 8, 44, 45].

[7] Lyndon R Evans and Philip Bryant. *LHC Machine*. *J. Instrum.* 3 (2008). This report is an abridged version of the LHC Design Report (CERN-2004-003), S08001. 164 p [see pp. 16, 18].

[8] LEP Higgs Working Group. *Search for the Standard Model Higgs boson at LEP*. *Phys. Lett.* B565 (2003), p. 61 [see p. 16].

[9] Christiane Lefèvre. *The CERN accelerator complex* (Dec. 2008) [see p. 16].

[10] Thomas Taylor. *Incident at the CERN LHC*. *TEION KOGAKU (Journal of Cryogenics and Superconductivity Society of Japan)* 45.7 (2010), pp. 344–351. DOI: `10.2221/jcsj.45.344` [see p. 18].

[11] The ATLAS Collaboration. *ATLAS Experiment - Public Results*. 2014 [see pp. 19, 33].

[12] The ATLAS Collaboration. *The ATLAS Experiment at the CERN Large Hadron Collider*. *JINST* 3 (2008), S08003. DOI: `10.1088/1748-0221/3/08/S08003` [see pp. 21, 23, 25, 27, 29, 31].

[13] The ATLAS Collaboration. *Expected Performance of the ATLAS Experiment – Detector, Trigger and Physics*. CERN-OPEN-2008-020 (2008). arXiv: `0901.0512 [hep-ex]` [see p. 35].

[14] The ATLAS Collaboration. *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*. ATLAS-CONF-2012-042 (Mar. 2012) [see p. 35].

[15] I Gavrilenko. *Description of Global Pattern Recognition Program (XKalman)*. ATL-INDET-97-165. ATL-I-PN-165 (Apr. 1997) [see p. 35].

[16] The ATLAS Collaboration. *Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment*. ATLAS-CONF-2010-069 (July 2010) [see p. 35].

[17] Wolfgang Waltenberger, Rudolf Frühwirth, and Pascal Vanlaer. *Adaptive vertex fitting*. *Journal of Physics G: Nuclear and Particle Physics* 34.12 (2007), N343 [see p. 35].

[18] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al. *Event generation with SHERPA 1.1*. *JHEP* 0902 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph] [see p. 38].

[19] W. Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. ATL-LARG-PUB-2008-002. 2008 [see p. 37].

[20] The ATLAS Collaboration. *Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC*. *Eur. Phys. J. C* 73 (3 2013), p. 2305. arXiv: 1203.1302 [hep-ex] [see pp. 37, 45].

[21] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. *The anti-k_t jet clustering algorithm*. *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph] [see p. 40].

[22] Matteo Cacciari and Gavin P. Salam. *Dispelling the N^3 myth for the k_t jet-finder*. *Phys. Lett.* B 641 (2006), p. 57. DOI: 10.1016/j.physletb.2006.08.037. arXiv: hep-ph/0512210 [see p. 40].

[23] The ATLAS Collaboration. *Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s}$ = 7 TeV with ATLAS 2011 data*. ATLAS-CONF-2013-004 (2013) [see pp. 40, 101].

[24] The ATLAS Collaboration. *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*. *Eur. Phys. J. C* arXiv:1404.2240. CERN-PH-EP-2014-040 (Apr. 2014), 74. 38 p [see p. 41].

[25] The ATLAS Collaboration. *ATLAS detector and physics performance: Technical Design Report, 1*. Geneva: CERN, 1999 [see p. 41].

[26] The ATLAS Collaboration. *Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data* (2013) [see p. 41].

[27] The ATLAS Collaboration. *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at* 8 *TeV* (2013) [see pp. 42, 43, 102].

[28] The ATLAS Collaboration. *Identification of Hadronic Decays of Tau Leptons in 2012 Data with the ATLAS Detector* (2013) [see pp. 45, 51–57, 100].

[29] The ATLAS Collaboration. *Performance of the Reconstruction and Identification of Hadronic Tau Decays in ATLAS with 2011 Data*. ATLAS-CONF-2012-142 (Oct. 2012) [see pp. 46, 47, 55].

[30] EN Dawe, DC O'Neil, and S Protopopescu. *Using Boosted Decision Trees for Hadronic Tau Identification*. ATL-PHYS-INT-2011-004 (2012) [see pp. 47, 69].

[31] The ATLAS Collaboration. *Performance of the Reconstruction and Identification of Hadronic Tau Decays with ATLAS*. ATLAS-CONF-2011-152 (2011) [see pp. 50, 183].

[32] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. *TMVA: Toolkit for Multivariate Data Analysis*. *PoS* ACAT (2007), p. 040. arXiv: `physics/0703039` [see pp. 50, 120].

[33] The 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) [see pp. 59–61, 100].

[34] The ATLAS Collaboration. *Performance of the ATLAS tau trigger in 2011*. ATLAS-CONF-2013-006 (Jan. 2013) [see p. 62].

[35] A. Elagin et al. *A new mass reconstruction technique for resonances decaying to di-tau*. *Nucl. Instrum. Meth.* A 654.FERMILAB-PUB-11-918-PPD (2011), p. 481. DOI: `10.1016/j.nima.2011.07.009`. arXiv: `1012.4686 [hep-ex]` [see p. 63].

[36] *Electron Quality*. `http://twiki.cern.ch/twiki/bin/view/AtlasProtected/LArCleaningAndObjectQuality` [see p. 68].

[37] *Muon Quality*. `http://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesRel17MC11a` [see p. 68].

[38] The ATLAS Collaboration. *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons*. ATLAS-CONF-2011-077 (May 2011) [see pp. 69, 183].

[39] *Jet Cleaning*. http://twiki.cern.ch/twiki/bin/view/AtlasProtected/HowToCleanJets [see p. 72].

[40] S. Jadach et al. *The tau decay library Tauola: Version 2.4*. *Comput. Phys. Commun.* 76.CERN-TH-6793-93 (1993), p. 361. DOI: 10.1016/0010-4655(93)90061-G [see pp. 82, 97].

[41] The ATLAS Collaboration. *Search for the Standard Model Higgs boson in H* $\rightarrow \tau^{+}\tau^{-}$ *decays in proton-proton collisions with the ATLAS detector*. ATLAS-CONF-2012-160 (2012) [see p. 82].

[42] The ATLAS Collaboration. *Search for the Standard Model Higgs boson in the H* $\rightarrow \tau\tau$ *decay mode with 4.7 fb*$^{-1}$ *of ATLAS data at 7 TeV*. ATLAS-CONF-2012-014 (Mar. 2012) [see p. 82].

[43] N. Davidson, T. Przedzinski, and Z. Was. *Photos Interface in C++: Technical and Physics Documentation*. (2010), [arXiv:1011.0937] [see pp. 83, 97].

[44] Z. Czyczula, T. Przedzinski, and Z. Was. *TauSpinner Program for Studies on Spin Effect in tau Production at the LHC*. *Eur. Phys. J.* C 72.CERN-PH-TH-20011-307 (2012), p. 1988. DOI: 10.1140/epjc/s10052-012-1988-z. arXiv: 1201.0117 [hep-ph] [see p. 83].

[45] Paolo Nason. *A new method for combining NLO QCD with shower Monte Carlo algorithms*. *JHEP* 11.BICOCCA-FT-04-11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 [see pp. 96, 98].

[46] Stefano Frixione, Paolo Nason, and Carlo Oleari. *Matching NLO QCD computations with parton shower simulations: the* POWHEG *method*. *JHEP* 0711.BICOCCA-FT-07-9, GEF-TH-21-2007 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph] [see pp. 96, 98].

[47] Simone Alioli et al. *A general framework for implementing NLO calculations in shower Monte Carlo programs: the* POWHEG *BOX*. *JHEP* 1006.DESY-10-018, SFB-CPP-10-22, IPPP-10-11, DCPT-10-22 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph] [see pp. 96, 98].

[48] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini. *Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM*. *JHEP* 1202.RM3-TH-11-5, IFUM-990-FT (2012), p. 088. DOI: 10.1007/JHEP02(2012)088. arXiv: 1111.2854 [hep-ph] [see pp. 96, 98].

[49] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. *A brief introduction to* PYTHIA *8.1*. *Comput. Phys. Commun.* 178.CERN-LCGAPP-2007-04, LU-TP-07-28, FERMILAB-PUB-07-512-CD-T (2008), p. 852. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph] [see pp. 96–98].

[50] H.-L. Lai et al. *New generation of parton distributions with uncertainties from global QCD analysis*. *Phys. Rev.* D 82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [see p. 96].

[51] A. Djouadi, M. Spira, and P.M. Zerwas. *Production of Higgs bosons in proton colliders: QCD corrections*. *Phys. Lett.* B 264.PITHA-91-6 (1991), p. 440. DOI: 10.1016/0370-2693(91)90375-Z [see pp. 96, 98].

[52] S. Dawson. *Radiative corrections to Higgs boson production*. *Nucl. Phys.* B 359 (1991), p. 283. DOI: 10.1016/0550-3213(91)90061-2 [see pp. 96, 98].

[53] M. Spira et al. *Higgs boson production at the LHC*. *Nucl. Phys.* B 453.DESY-94-123, UDEM-GPP-TH-95-16, CERN-TH-95-30, CERN-TH-95-030, DESY-94–123, GPP–UDEM–TH–95–16-AND-CERN–TH–95–30 (1995), p. 17. DOI: 10.1016/0550-3213(95)00379-7. arXiv: hep-ph/9504378 [hep-ph] [see pp. 96, 98].

[54] Robert V. Harlander and William B. Kilgore. *Next-to-next-to-leading order Higgs production at hadron colliders*. *Phys. Rev. Lett.* 88.BNL-HET-02-3, CERN-TH-2002-006 (2002), p. 201801. DOI: 10.1103/PhysRevLett.88.201801. arXiv: hep-ph/0201206 [hep-ph] [see pp. 96, 98].

[55] Charalampos Anastasiou and Kirill Melnikov. *Higgs boson production at hadron colliders in NNLO QCD*. *Nucl. Phys.* B 646.SLAC-PUB-9273 (2002), p. 220. DOI: 10.1016/S0550-3213(02)00837-4. arXiv: hep-ph/0207004 [hep-ph] [see pp. 96, 98].

[56] V. Ravindran, J. Smith, and W. L. van Neerven. *NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions*. *Nucl. Phys.*

B 665.YITP-SB-03-02, INLO-PUB-01-03 (2003), p. 325. DOI: 10.1016/S0550-3213(03)00457-7. arXiv: hep-ph/0302135 [hep-ph] [see pp. 96, 98].

[57] S Catani et al. *Soft gluon resummation for Higgs boson production at hadron colliders*. *JHEP* 0307 (2003), p. 028. DOI: 10.1088/1126-6708/2003/07/028. arXiv: hep-ph/0306211 [see p. 96].

[58] U. Aglietti et al. *Two loop light fermion contribution to Higgs production and decays*. *Phys. Lett.* B 595.ROME1-1373-04, FREIBURG-THEP-04-05, RM3-TH-04-06, IFUM-788-FT (2004), p. 432. DOI: 10.1016/j.physletb.2004.06.063. arXiv: hep-ph/0404071 [hep-ph] [see p. 96].

[59] Stefano Actis et al. *NLO electroweak corrections to Higgs boson production at hadron colliders*. *Phys. Lett.* B 670.PITHA-08-20, SFB-CPP-08-62, TTP08-38 (2008), p. 12. DOI: 10.1016/j.physletb.2008.10.018. arXiv: 0809.1301 [hep-ph] [see p. 96].

[60] M. Ciccolini, Ansgar Denner, and S. Dittmaier. *Strong and electroweak corrections to the production of Higgs + 2-jets via weak interactions at the LHC*. *Phys. Rev. Lett.* 99.MPP-2007-82, PSI-PR-07-03 (2007), p. 161803. DOI: 10.1103/PhysRevLett.99.161803. arXiv: 0707.0381 [hep-ph] [see pp. 96–98, 104].

[61] Mariano Ciccolini, Ansgar Denner, and Stefan Dittmaier. *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC*. *Phys. Rev.* D 77.MPP-2007-152, PSI-PR-07-06, UWTHPH-2007-26 (2008), p. 013002. DOI: 10.1103/PhysRevD.77.013002. arXiv: 0710.4749 [hep-ph] [see pp. 96–98, 104].

[62] K. Arnold et al. *VBFNLO: A parton level Monte Carlo for processes with electroweak bosons*. *Comput. Phys. Commun.* 180.KA-TP-31-2008, SFB-CPP-08-95 (2009), p. 1661. DOI: 10.1016/j.cpc.2009.03.006. arXiv: 0811.4559 [hep-ph] [see pp. 96, 98].

[63] Paolo Bolzoni et al. *Higgs production via vector-boson fusion at NNLO in QCD*. *Phys. Rev. Lett.* 105.DESY-10-042, SFB-CPP-10-23, CP3-10-12 (2010), p. 011801. DOI: 10.1103/PhysRevLett.105.011801. arXiv: 1003.4451 [hep-ph] [see p. 96].

[64] J. Pumplin et al. *New generation of parton distributions with uncertainties from global QCD analysis*. *JHEP* 0207.MSU-HEP-011101 (2002), p. 012. DOI: 10.1088/1126-6708/2002/07/012. arXiv: hep-ph/0201195 [see p. 96].

[65] Oliver Brein, Abdelhak Djouadi, and Robert Harlander. *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*. *Phys. Lett.* B 579.MPP-2003-35, CERN-TH-2003-161, PM-03-16 (2004), p. 149. DOI: 10.1016/j.physletb.2003.10.112. arXiv: hep-ph/0307206 [hep-ph] [see pp. 96, 98].

[66] M.L. Ciccolini, S. Dittmaier, and M. Krämer. *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*. *Phys. Rev.* D 68 (2003), p. 073003. DOI: 10.1103/PhysRevD.68.073003. arXiv: hep-ph/0306234 [hep-ph] [see p. 96].

[67] Massimiliano Grazzini and Hayk Sargsyan. *Heavy-quark mass effects in Higgs boson production at the LHC*. *JHEP* 1309.ZU-TH-10-13 (2013), p. 129. DOI: 10.1007/JHEP09(2013)129. arXiv: 1306.4581 [hep-ph] [see p. 96].

[68] D. de Florian et al. *Higgs boson production at the LHC: transverse momentum resummation effects in the $H \to 2\gamma, H \to WW \to \ell\nu\ell\nu$ and $H \to ZZ \to 4\ell$ decay modes*. *JHEP* 1206.ZU-TH-04-12, IFUM-993-FT (2012), p. 132. DOI: 10.1007/JHEP06(2012)132. arXiv: 1203.6321 [hep-ph] [see p. 96].

[69] John M. Campbell, R. Keith Ellis, and Giulia Zanderighi. *Next-to-Leading order Higgs + 2 jet production via gluon fusion*. *JHEP* 0610 (2006), p. 028. DOI: 10.1088/1126-6708/2006/10/028. arXiv: hep-ph/0608194 [hep-ph] [see p. 96].

[70] Andrea Banfi et al. *Higgs and Z-boson production with a jet veto*. *Phys. Rev. Lett.* 109 (2012), p. 202001. DOI: 10.1103/PhysRevLett.109.202001. arXiv: 1206.4998 [hep-ph] [see p. 96].

[71] Andrea Banfi, Gavin P. Salam, and Giulia Zanderighi. *NLL+NNLO predictions for jet-veto efficiencies in Higgs-boson and Drell-Yan production*. *JHEP* 1206.CERN-PH-TH-2012-065, OUTP-12-04P (2012), p. 159. DOI: 10.1007/JHEP06(2012)159. arXiv: 1203.5773 [hep-ph] [see p. 96].

[72] Andrea Banfi, Pier Francesco Monni, and Giulia Zanderighi. *Quark masses in Higgs production with a jet veto*. *JHEP* 1401 (2014), p. 097. DOI: 10.1007/JHEP01(2014)097. arXiv: 1308.4634 [hep-ph] [see p. 96].

[73] LHC Higgs Cross Section Working Group: S. Dittmaier C. Mariotti G. Passarino R. Tanaka (Eds.) *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions* (2012). arXiv: 1201.3084 [see pp. 97, 105].

[74] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. PYTHIA *6.4 physics and manual*. *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [see pp. 97, 98].

[75] G. Corcella et al. HERWIG *6.5 release note*. (2002), [hep-ph/0210213] [see pp. 97, 98].

[76] M. Mangano et al. ALPGEN*, a generator for hard multiparton processes in hadronic collisions*. *JHEP* 07 (2003), p. 001. DOI: 10.1088/1126-6708/2003/07/001. arXiv: hep-ph/0206293 [see pp. 97, 98].

[77] M. Mangano et al. *Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b}$ + n jets as a case study*. *Nucl. Phys.* B 632 (2002), p. 343. DOI: 10.1016/S0550-3213(02)00249-3. arXiv: hep-ph/0108069 [see p. 97].

[78] T. Binoth et al. *Gluon-induced W-boson pair production at the LHC*. *JHEP* 0612 (2006), p. 046. DOI: 10.1088/1126-6708/2006/12/046. arXiv: hep-ph/0611170 [see pp. 97, 98].

[79] Borut Paul Kersevan and Elzbieta Richter-Was. *The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*. *Comput. Phys. Commun.* 184.TPJU-6-2004 (2013), p. 919. DOI: 10.1016/j.cpc.2012.10.032. arXiv: hep-ph/0405247 [hep-ph] [see pp. 97, 98].

[80] The ATLAS Collaboration. *The ATLAS Simulation Infrastructure*. *Eur. Phys. J.* C 70 (2010), p. 823. DOI: 10.1140/epjc/s10052-010-1429-9. arXiv: 1005.4568 [hep-ex] [see p. 97].

[81] S. Agostinelli et al. *Geant4: A Simulation toolkit*. *Nucl. Instrum. Meth.* A 506.SLAC-PUB-9350, FERMILAB-PUB-03-339 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8 [see p. 97].

[82] *Summary of ATLAS* PYTHIA *8 tunes*. ATL-PHYS-PUB-2012-003. 2012 [see p. 97].

[83] Stefano Catani et al. *Vector boson production at hadron colliders: A fully exclusive QCD calculation at next-to-next-to-leading order*. *Phys. Rev. Lett.* 103 (8 2009), p. 082001. DOI: 10.1103/PhysRevLett.103.082001 [see p. 98].

[84] 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*. *Phys. Rev. Lett.* 98 (22 2007), p. 222002. DOI: 10.1103/PhysRevLett.98.222002 [see p. 98].

[85] T. Gleisberg et al. *Event generation with SHERPA 1.1*. *JHEP* 0902 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph] [see p. 98].

[86] Matteo Cacciari et al. *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*. *Phys. Lett.* B 710 (2012), p. 612. DOI: http://dx.doi.org/10.1016/j.physletb.2012.03.013 [see p. 98].

[87] P. Bärnreuther, M. Czakon and A. Mitov. *Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$*. *Phys. Rev. Lett.* 109.CERN-PH-TH-2012-092 (2012), p. 132001. DOI: 10.1103/PhysRevLett.109.132001. arXiv: 1204.5201 [hep-ph] [see p. 98].

[88] M. Czakon and A. Mitov. *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*. *JHEP* 1212 (2012), p. 054. DOI: 10.1007/JHEP12(2012)054. arXiv: 1207.0236 [hep-ph] [see p. 98].

[89] M. Czakon and A. Mitov. *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*. *JHEP* 1301 (2013), p. 080. DOI: 10.1007/JHEP01(2013)080. arXiv: 1210.6832 [hep-ph] [see p. 98].

[90] M. Czakon, P. Fiedler, A. Mitov. *The total top quark pair production cross-section at hadron colliders through $\mathcal{O}(\alpha_S^4)$*. *Phys. Rev. Lett.* 110 (2013), p. 252004. DOI: 10.1103/PhysRevLett.110.252004. arXiv: 1303.6254 [hep-ph] [see p. 98].

[91] M. Czakon and A. Mitov. *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*. *Comput. Phys. Commun.* 185.CERN-PH-TH-2011-303, TTK-11-58 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph] [see p. 98].

[92] Nikolaos Kidonakis. *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-*. *Phys. Rev.* D 82 (5 2010), p. 054018. DOI: 10.1103/PhysRevD.82.054018. arXiv: 1005.4451 [hep-ph] [see p. 98].

[93] Nikolaos Kidonakis. *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*. *Phys. Rev.* D 83 (9 2011), p. 091503. DOI: 10.1103/PhysRevD.83.091503. arXiv: 1103.2792 [hep-ph] [see p. 98].

[94] Nikolaos Kidonakis. *Next-to-next-to-leading logarithm resummation for s-channel single top quark production*. *Phys. Rev.* D 81 (5 2010), p. 054028. DOI: 10.1103/PhysRevD.81.054028. arXiv: 1001.5034 [hep-ph] [see p. 98].

[95] J. M. Campbell, K. R. Ellis, and C. Williams. *Vector boson pair production at the LHC*. *JHEP* 1107 (2011), p. 018. DOI: 10.1007/JHEP07(2011)018. arXiv: 1105.0020 [hep-ph] [see pp. 98, 104].

[96] The ATLAS Collaboration. *Improved luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC*. *Eur. Phys. J.* C 73.CERN-PH-EP-2013-026 (2013), p. 2518. DOI: 10.1140/epjc/s10052-013-2518-3. arXiv: 1302.4393 [hep-ex] [see p. 99].

[97] The ATLAS Collaboration. *Jet energy resolution in proton-proton collisions at $\sqrt{s}$ = 7 TeV recorded in 2010 with the ATLAS detector*. *Eur. Phys. J.* C 73 (2013), p. 2306. DOI: 10.1140/epjc/s10052-013-2306-0. arXiv: 1210.6210 [hep-ex] [see p. 102].

[98] S. Dittmaier et al. (LHC Higgs Cross Section Working Group). *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables* (2011). arXiv: 1101.0593 [hep-ph] [see p. 104].

[99] I.W. Stewart and F.J. Tackmann. *Theory uncertainties for Higgs and other searches using jet bins*. *Phys. Rev.* D 85 (2012), p. 034011. DOI: 10.1103/PhysRevD.85.034011. arXiv: 1107.2117 [hep-ph] [see p. 104].

[100] S. Frixione and B. R. Webber. *Matching NLO QCD computations and parton shower simulations*. *JHEP* 06 (2002), p. 029. arXiv: hep-ph/0204244 [see p. 105].

[101] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. *Parton distributions for the LHC*. *Eur. Phys. J.* C 63 (2009), p. 189. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph] [see p. 105].

[102] NNPDF Collaboration, Richard D. Ball et al. *Impact of heavy quark masses on parton distributions and LHC phenomenology*. *Nucl. Phys.* B 849 (2011), p. 296. DOI: 10.1016/j.nuclphysb.2011.03.021. arXiv: 1101.1300 [hep-ph] [see p. 105].

[103] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al. *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*. *JHEP* 1407.CERN-PH-TH-2014-064, CP3-14-18, LPN14-066, MCNET-14-09, ZU-TH-14-14 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph] [see p. 105].

[104] Yoav Freund and Robert E. Schapire. *A decision-theoretic generalization of on-line learning and an application to boosting*. *J. Comput. Syst. Sci.* 55 (1997), p. 119. DOI: 10.1006/jcss.1997.1504 [see p. 118].

[105] R. Brun and F. Rademakers. *ROOT: An object oriented data analysis framework*. *Nucl.Instrum.Meth.* A389 (1997), pp. 81–86. DOI: 10.1016/S0168-9002(97)00048-X [see p. 120].

[106] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. *Scikit-learn: Machine Learning in Python*. *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830 [see p. 121].

[107] L. Breiman, J. Friedman, R. Olshen, and C. Stone. *Classification and Regression Trees*. Monterey, CA: Wadsworth and Brooks, 1984 [see p. 121].

[108] Trevor Hastie, Saharon Rosset, Ji Zhu, and Hui Zou. *Multi-class AdaBoost*. *Statistics and Its Interface* 2.3 (2009), pp. 349–360. DOI: 10.4310/sii.2009.v2.n3.a8 [see p. 121].

[109] Glen Cowan et al. *Asymptotic formulae for likelihood-based tests of new physics*. *Eur. Phys. J.* C 1 (2011), p. 1554. DOI: `10.1140/epjc/s10052-011-1554-0`. arXiv: `1007.1727 [physics.data-an]` [see pp. 128, 131, 132, 145].

[110] The ATLAS Collaboration. *Procedure for the LHC Higgs boson search combination in summer 2011*. ATL-PHYS-PUB-2011-011 (Aug. 2011) [see p. 128].

[111] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. CERN-OPEN-2012-016 (Jan. 2012) [see p. 133].

[112] *RooStats*. http://twiki.cern.ch/twiki/bin/view/RooStats/WebHome [see p. 133].

[113] Wouter Verkerke and David P. Kirkby. *The RooFit toolkit for data modeling*. *eConf* C0303241 (2003), MOLT007. arXiv: `physics/0306116 [physics]` [see p. 133].

[114] F. James and M. Roos. *Minuit - a system for function minimization and analysis of the parameter errors and correlations*. *Computer Physics Communications* 10.6 (1975), pp. 343–367. DOI: `http://dx.doi.org/10.1016/0010-4655(75)90039-9` [see pp. 133, 155].

[115] The ATLAS Collaboration. *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data*. *Phys.Rev.* D 90.CERN-PH-EP-2014-122 (2014), p. 052004. DOI: `10.1103/PhysRevD.90.052004`. arXiv: `1406.3827 [hep-ex]` [see p. 144].

[116] A. David et al. (LHC Higgs Cross Section Working Group). *LHC HXSWG interim recommendations to explore the coupling structure of a Higgs-like particle* (2012). arXiv: `1209.0040 [hep-ph]` [see p. 147].

[117] B. Efron and R. Tibshirani. *Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy*. *Statistical Science* 1.1 (1986), pp. 54–75 [see p. 155].

[118] Lars Perchalla and Achim Stahl. “Kinematic Tau Reconstruction and Search For The Higgs Boson in Hadronic Tau Pair Decays with the CMS Experiment”. presented 11 May 2011. PhD thesis. Aachen, Tech. Hochsch., 2011 [see p. 159].

[119] Ben Gripaios, Keiko Nagao, Mihoko Nojiri, Kazuki Sakurai, and Bryan Webber. *Reconstruction of Higgs bosons in the di-tau channel via 3-prong decay*. *JHEP* 1303 (2013), p. 106. DOI: 10.1007/JHEP03(2013)106. arXiv: 1210.1938 [hep-ph] [see p. 159].

[120] The ATLAS Collaboration. *ATLAS Computing: technical design report*. Geneva: CERN, 2005 [see p. 182].

[121] The ATLAS Collaboration. *Measurement of the W $\rightarrow \tau\nu_\tau$ Cross Section in pp Collisions at $\sqrt{s}$ = 7 TeV with the ATLAS experiment*. *Phys. Lett. B* 706.arXiv:1108.4101 CERN-PH-EP-2011-122 (Sept. 2011), pp. 276–294. [See p. 183].

Appendix A

Additional Plots

A.1 Distribution Contour Plots

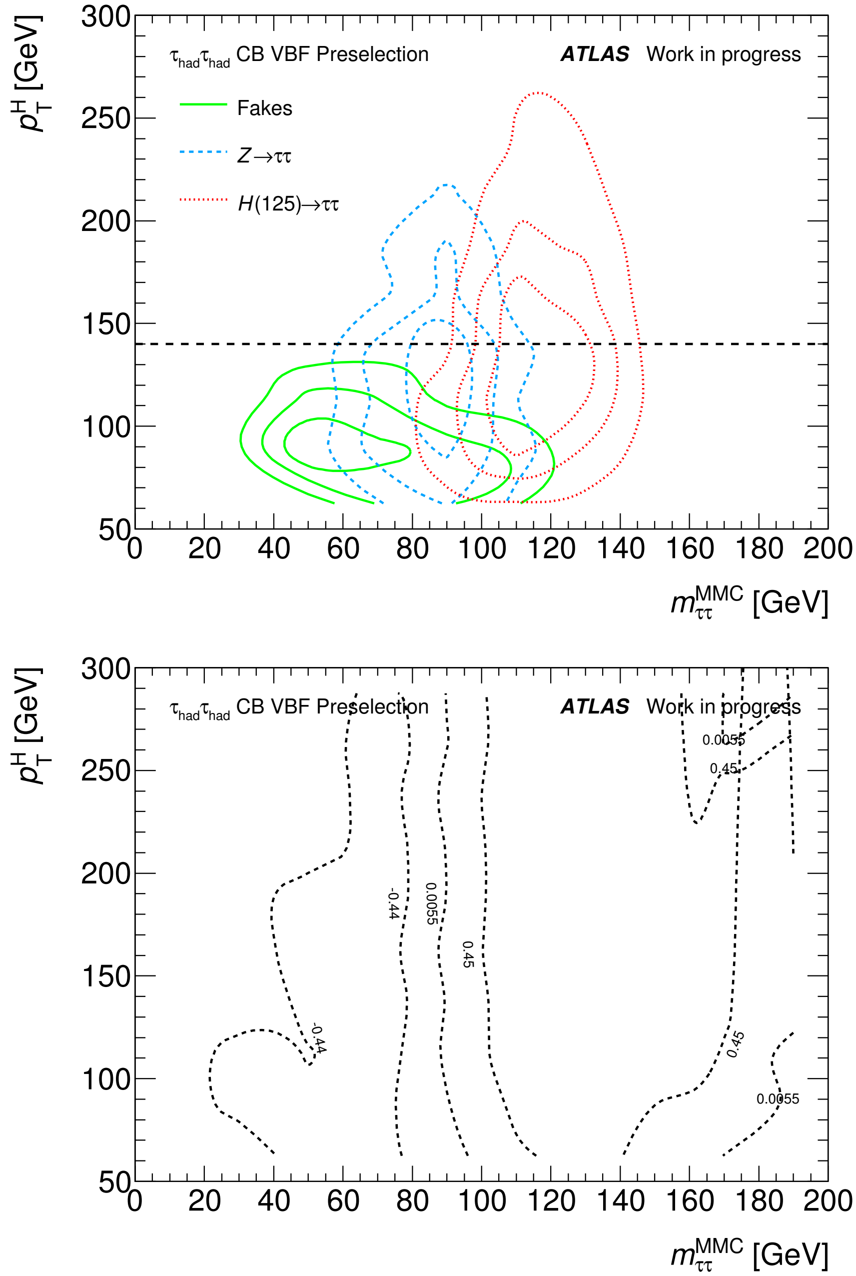
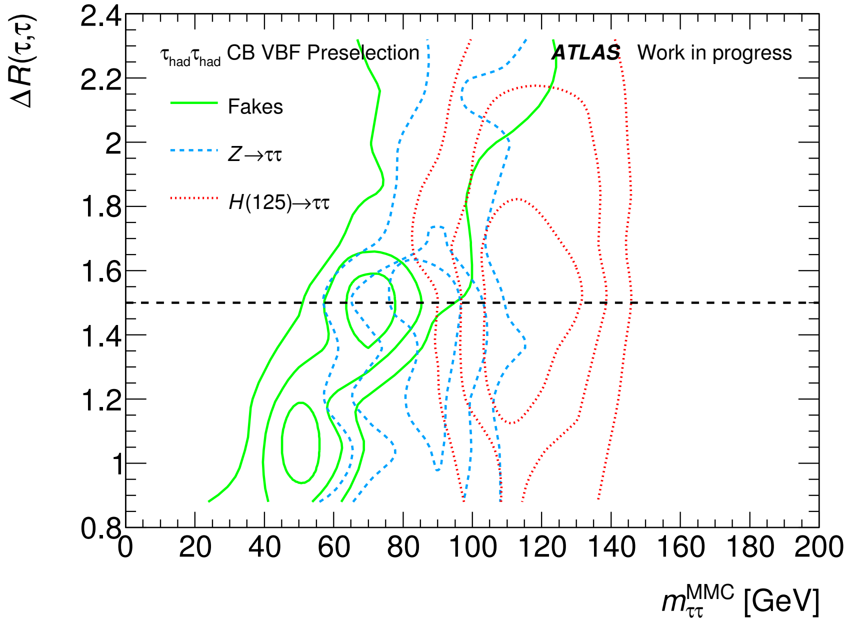


Figure A.1: Top: two-dimensional contours of the sample distributions at the cut-based VBF preselection stage in the $(m^{\mathrm{MMC}}_{\tau\tau}, p^{H}_{\mathrm{T}})$ plane. The black dashed line illustrates the categories as listed in Table 4.3. The samples are simulated Higgs boson events with $m_H = 125$ GeV, embedded $Z \to \tau\tau$, and fake tau multijet events estimated from data. Bottom: contours of the profiled BDT scores in the same two-dimensional plane. Regions of higher average BDT score overlap with the Higgs boson event distribution in the top figure.



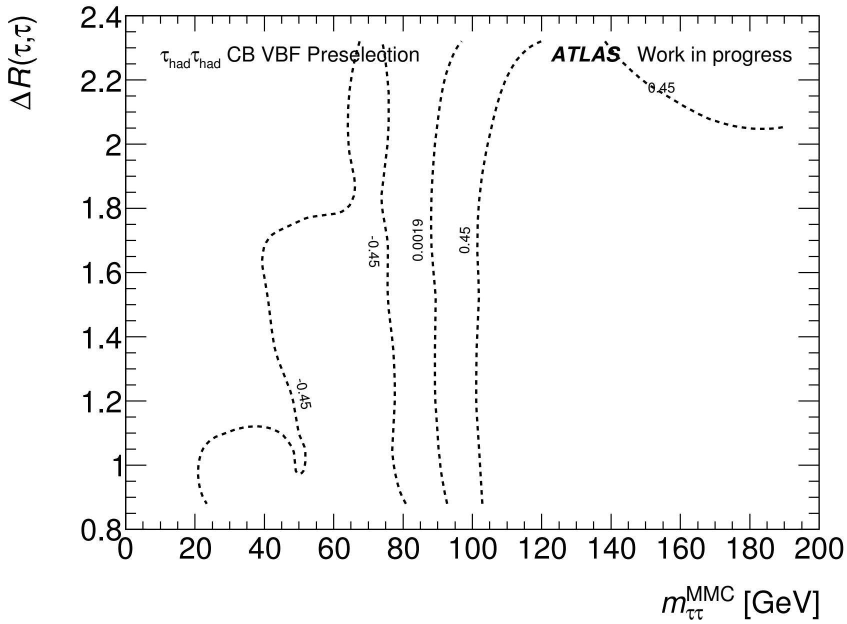


Figure A.2: Top: two-dimensional contours of the sample distributions at the cut-based VBF preselection stage in the ($m_{\tau\tau}^{\text{MMC}}$, $\Delta R_{\tau\tau}$) plane. The black dashed line illustrates the categories as listed in Table 4.3. The samples are simulated Higgs boson events with m_H = 125 GeV, embedded $Z \rightarrow \tau\tau$, and fake tau multijet events estimated from data. Bottom: contours of the profiled BDT scores in the same two-dimensional plane. Regions of higher average BDT score overlap with the Higgs boson event distribution in the top figure.

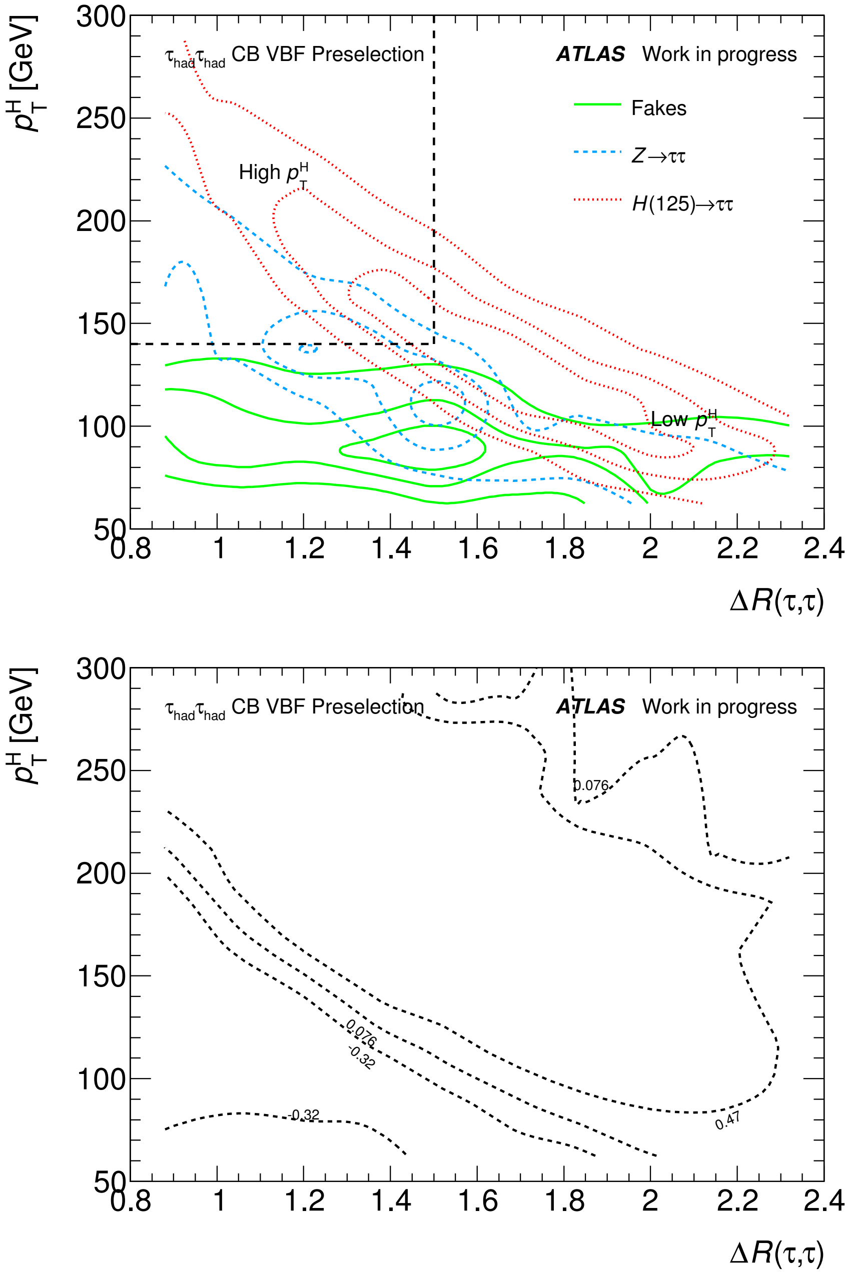


Figure A.3: Top: two-dimensional contours of the sample distributions at the cut-based VBF preselection stage in the ($\Delta R_{\tau\tau}$, p_T^H) plane. The black dashed line illustrates the categories as listed in Table 4.3. The samples are simulated Higgs boson events with m_H = 125 GeV, embedded $Z \rightarrow \tau\tau$, and fake tau multijet events estimated from data. Bottom: contours of the profiled BDT scores in the same two-dimensional plane. Regions of higher average BDT score overlap with the Higgs boson event distribution in the top figure.

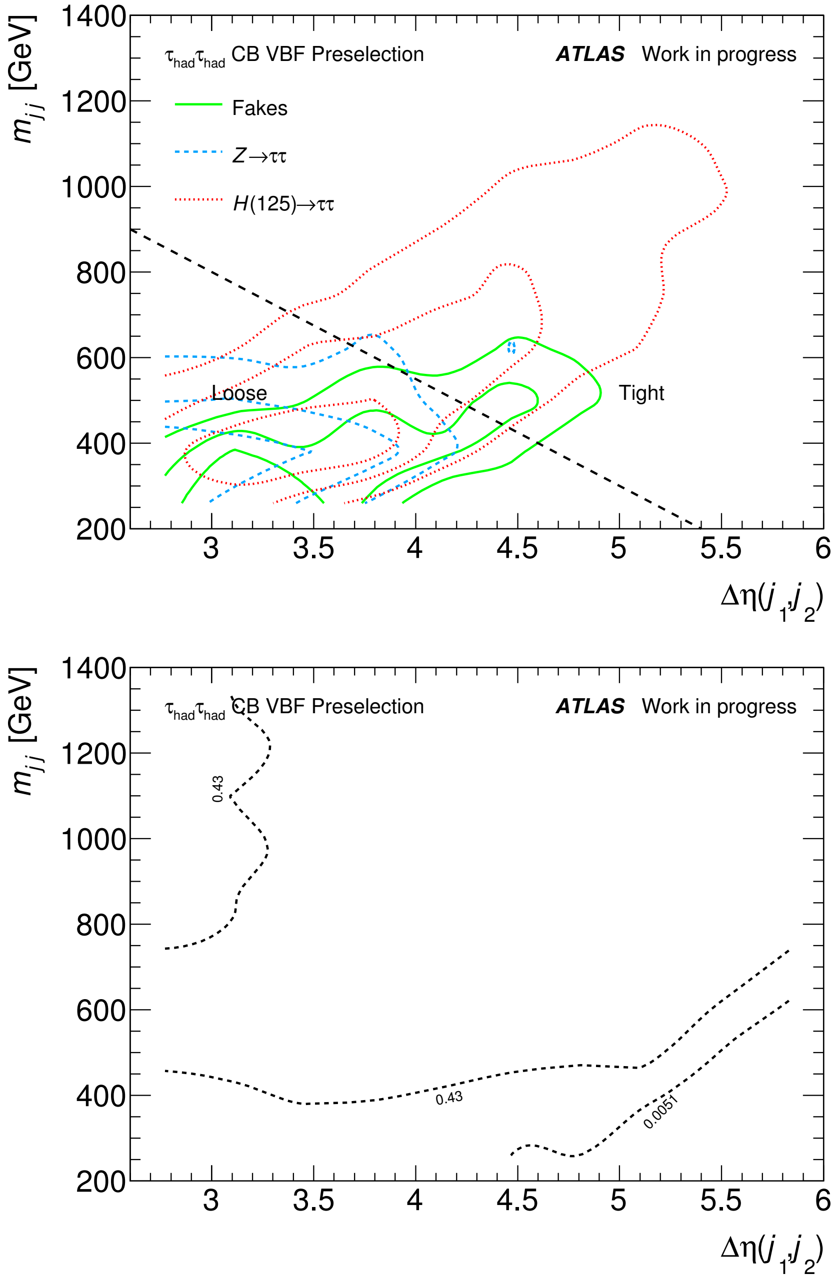


Figure A.4: Top: two-dimensional contours of the sample distributions at the cut-based VBF preselection stage in the ($|\Delta\eta_{jj}|$, m_{jj}) plane. The black dashed line illustrates the categories as listed in Table 4.3. The samples are simulated Higgs boson events with $m_H = 125$ GeV, embedded $Z \rightarrow \tau\tau$, and fake tau multijet events estimated from data. Bottom: contours of the profiled BDT scores in the same two-dimensional plane. Regions of higher average BDT score overlap with the Higgs boson event distribution in the top figure.

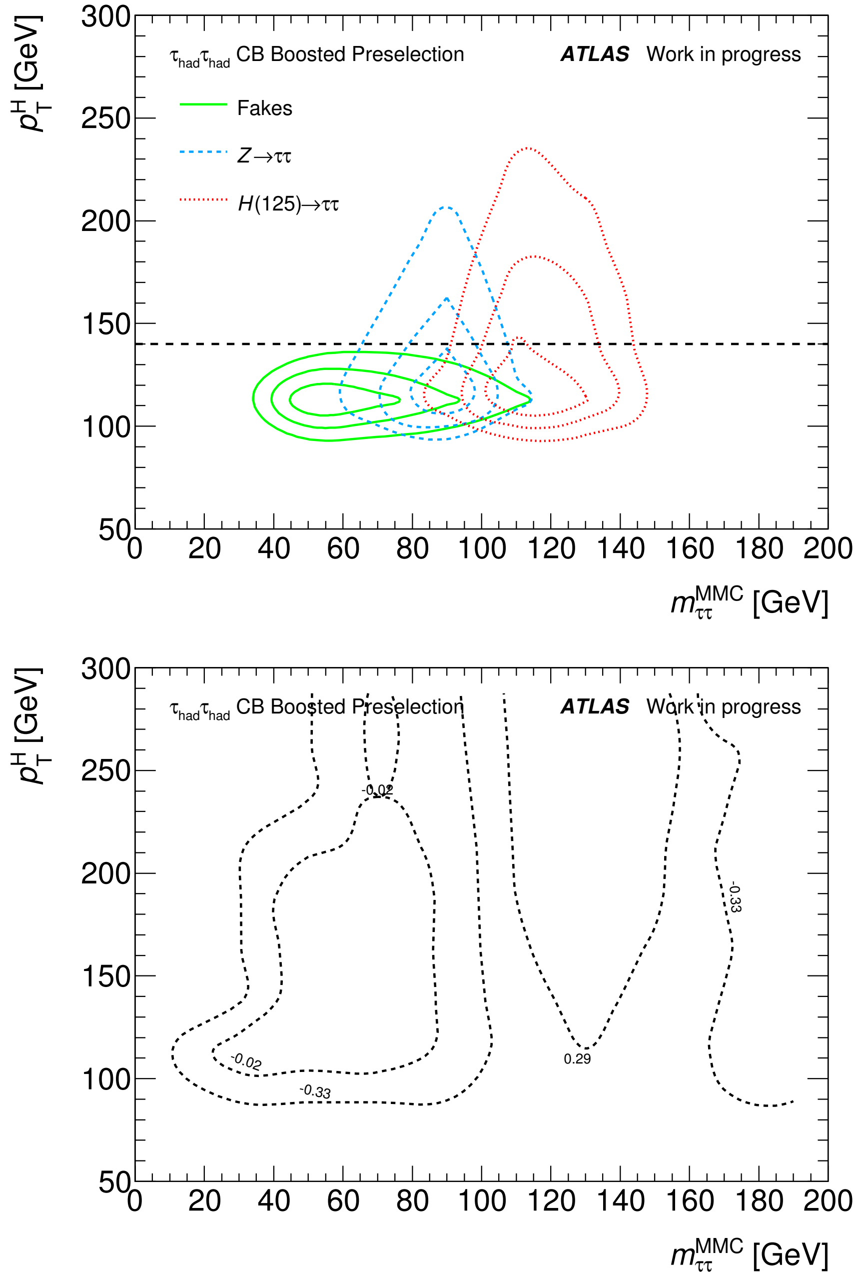


Figure A.5: Top: two-dimensional contours of the sample distributions at the cut-based boosted preselection stage in the ($m_{\tau\tau}^{\text{MMC}}$, p_{T}^{H}) plane. The black dashed line illustrates the categories as listed in Table 4.3. The samples are simulated Higgs boson events with m_H = 125 GeV, embedded $Z \to \tau\tau$, and fake tau multijet events estimated from data. Bottom: contours of the profiled BDT scores in the same two-dimensional plane. Regions of higher average BDT score overlap with the Higgs boson event distribution in the top figure.

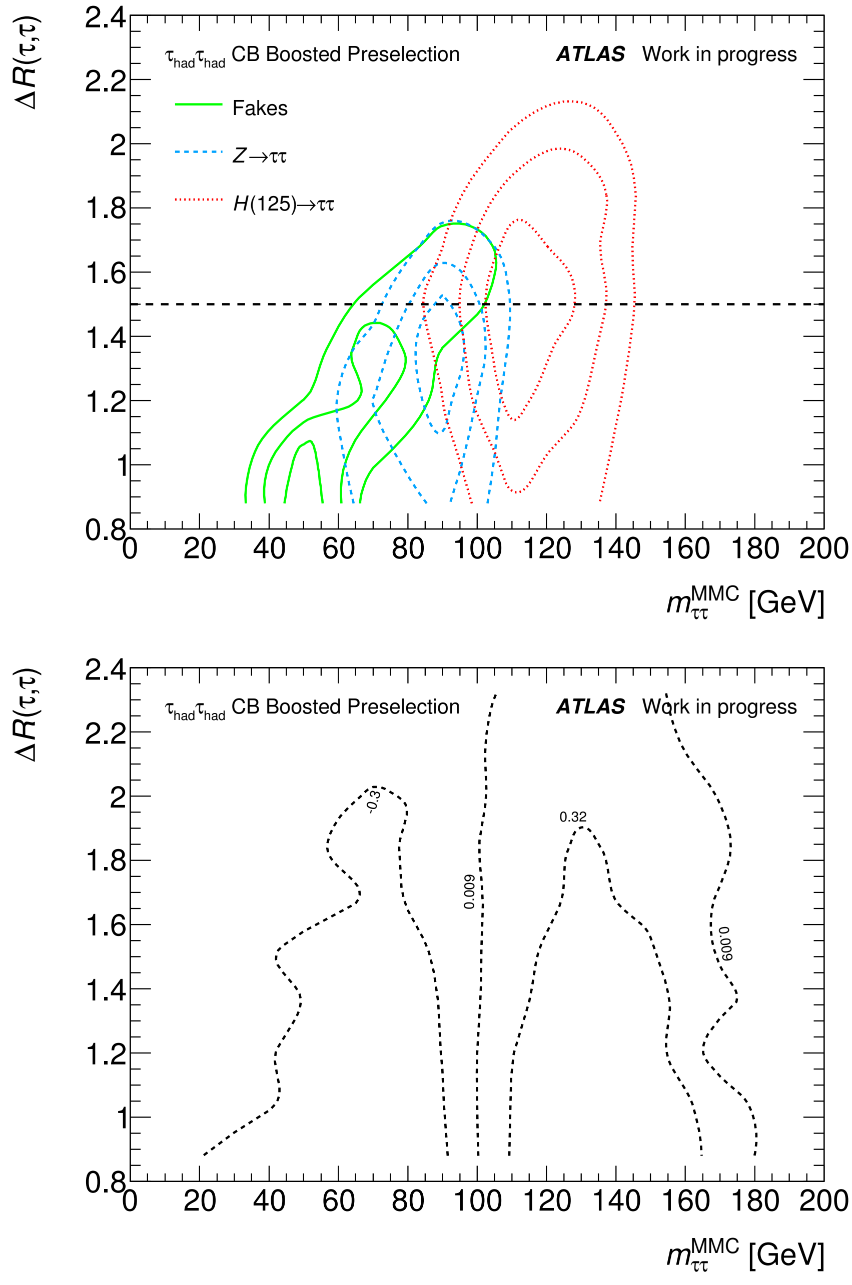


Figure A.6: Top: two-dimensional contours of the sample distributions at the cut-based boosted preselection stage in the ($m_{\tau\tau}^{MMC}$, $\Delta R_{\tau\tau}$) plane. The black dashed line illustrates the categories as listed in Table 4.3. The samples are simulated Higgs boson events with m_H = 125 GeV, embedded $Z \to \tau\tau$, and fake tau multijet events estimated from data. Bottom: contours of the profiled BDT scores in the same two-dimensional plane. Regions of higher average BDT score overlap with the Higgs boson event distribution in the top figure.

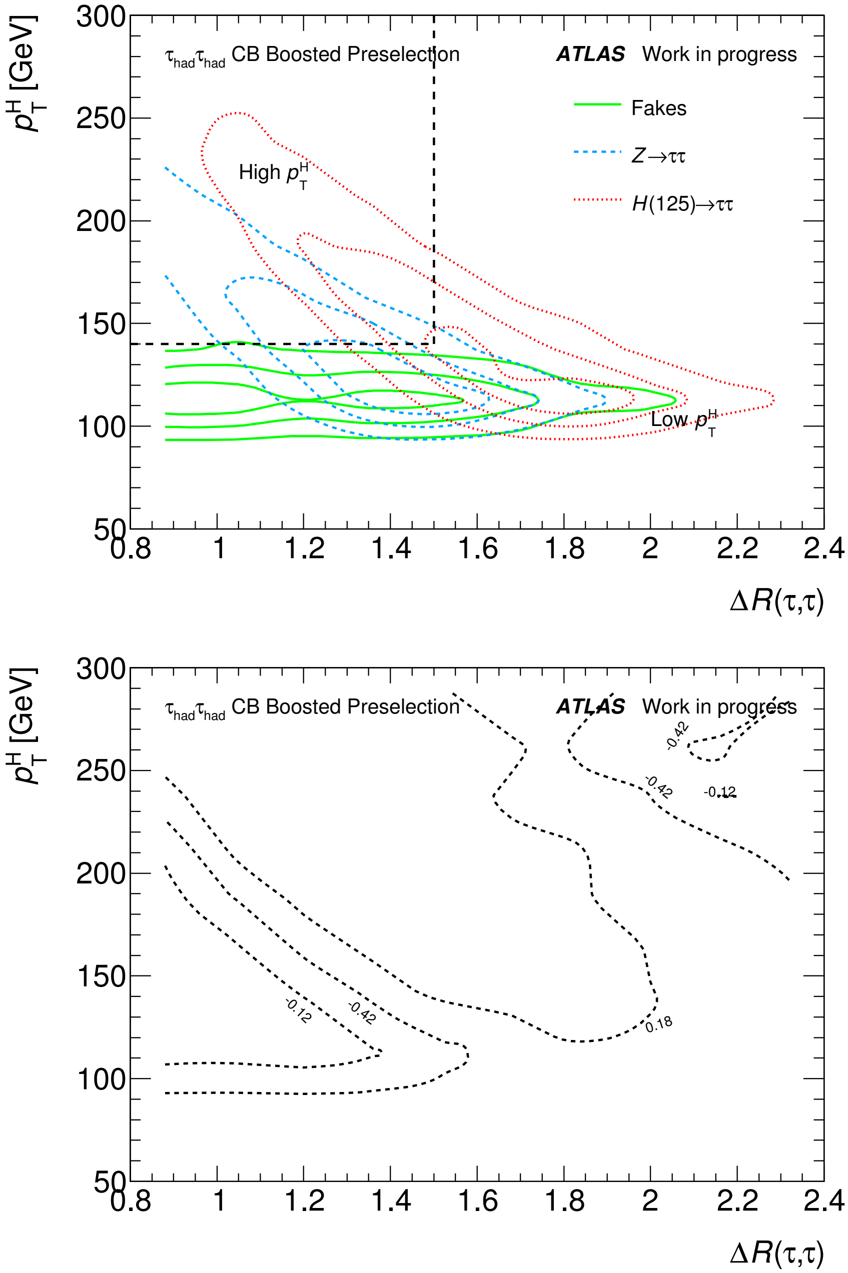


Figure A.7: Top: two-dimensional contours of the sample distributions at the cut-based boosted preselection stage in the ($\Delta R_{\tau\tau}$, p_T^H) plane. The black dashed line illustrates the categories as listed in Table 4.3. The samples are simulated Higgs boson events with m_H = 125 GeV, embedded $Z \rightarrow \tau\tau$, and fake tau multijet events estimated from data. Bottom: contours of the profiled BDT scores in the same two-dimensional plane. Regions of higher average BDT score overlap with the Higgs boson event distribution in the top figure.

Appendix B

Personal Contributions

The ATLAS experiment is driven by a large collaboration of approximately 3,000 members from many institutions around the world. All studies performed with ATLAS data depend on the work of numerous others. This appendix highlights the work I have personally contributed as part of supporting the analysis presented in this thesis and other areas of ATLAS research.

B.1 As a member of the Tau Working Group

- Developed the BDT-based tau identification used in first ATLAS data that was later updated by other members of the Tau Working Group for the 2012 data-taking period. My work included the selection of tau features to be used as inputs to the BDT training, the BDT training itself, performance evaluation, and comparison with other identification methods. This led to the adoption of the BDT technique as the default ATLAS tau identification.
- Developed parametrized cut thresholds on the BDT output in a way that yield approximately constant tau efficiency over p_{T}. These are the “loose”, “medium“, and “tight” working points used by all ATLAS tau-related analyses and reconstruction software.
- Developing and maintaining the tau identification software and maintenance of the Tau Event Data Model software in the Athena framework [120].

- Contributed to the documentation of the ATLAS tau performance in Refs. [31, 38].
- Contributed tau identification expertise in the measurement of the $W \to \tau\nu$ cross section published in Ref. [121].

B.2 As a member of the Higgs Working Group

- Implemented the data reduction software used to select the relevant events for the $H \to \tau_{\text{had}}\tau_{\text{had}}$ analysis.
- Developed the multivariate analysis in search of $H \to \tau_{\text{had}}\tau_{\text{had}}$ using BDTs.
- Implemented the $H \to \tau_{\text{had}}\tau_{\text{had}}$ cut-based analysis using updated ATLAS recommendations as a cross-check of the multivariate analysis.
- Used the bootstrap technique to assess the compatibility of the multivariate and cut-based results.
- This is the only ATLAS thesis presenting the final Run 1 $H \to \tau_{\text{had}}\tau_{\text{had}}$ multivariate and cut-based analyses. I was involved in many aspects of developing the $H \to \tau_{\text{had}}\tau_{\text{had}}$ analysis in its final form and implemented the majority of the $\tau_{\text{had}}\tau_{\text{had}}$-specific analysis software and statistical models entering the final plots, tables, and combined fit of all thee tau channels.

B.3 Other Contributions

- Developer on pyAMI, the CLI and Python API for the ATLAS Metadata Interface
- Liquid Argon Shifter in the ATLAS control room
- Creator and lead developer of the rootpy project: `http://rootpy.org`. This pythonic interface for the ROOT libraries formed the core dependency of the $H \to \tau_{\text{had}}\tau_{\text{had}}$ analysis framework and is used by other analyses in ATLAS and other experiments.
- Contributed an implementation of AdaBoost and support for weighted events in the `scikit-learn` software library for machine learning: `http://scikit-learn.org`. `scikit-learn` was then used by the $H \to \tau_{\text{had}}\tau_{\text{had}}$ analysis framework.

List of Figures

List of Tables