Searches for Standard Model Higgs Boson Pair-Production and New Resonances in the $b\bar{b}\tau^+\tau^-$ Final State with the ATLAS Detector

A thesis submitted in accordance of the requirements of
the University of Liverpool for the degree of Doctor of Philosophy
by

Emily Graham

Department of Physics,
Oliver Lodge Laboratory
University of Liverpool

June 2019

Declaration

This thesis is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this, or any other, university. This thesis does not exceed the word limit for the respective Degree Committee.

Emily Graham

Abstract

This thesis presents a search for resonant and non-resonant Higgs boson pair-production where one Higgs decays to a pair of b-quarks and the other to a pair of τ-leptons, followed by a search for third-generation scalar leptoquark pair-production. Both searches are performed using 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV proton-proton collision data collected by the ATLAS detector in 2015 and 2016. The analysis description focuses on the final state where one τ-lepton decays leptonically and the other hadronically, though the results presented are a combination with the final state where both τ-leptons decay hadronically. No excess is observed for any search and the results are presented as 95% confidence level upper limits.

The observed (expected) upper limit on the non-resonant Higgs boson pair-production cross-section is 30.9 fb (36.0 fb), i.e. 12.7 (14.8) times the Standard Model prediction. To date, this is the world's most stringest limit on non-resonant Higgs boson pair-production in a single decay channel.

The search for resonant SM production is performed for two benchmark models: a narrow-width scalar Higgs in the hMSSM and a spin-2 Kaluza-Klein graviton in the bulk Randall-Sundrum model. For a heavy scalar Higgs boson in the hMSSM, the mass range $305 < m_X < 402$ GeV is excluded at the 95% confidence level for $\tan\beta = 2$. RS gravitons are excluded in the range $325 < m_G < 885$ GeV for $k/\bar{M}_{\mathrm{Pl}} = 1.0$; for $k/\bar{M}_{\mathrm{Pl}} = 2.0$, the entire mass range of the search ($260 < m_G < 1000$ GeV) can be excluded.

Finally, up-type (down-type) leptoquarks with $m(\mathrm{LQ}_3^u) < 1030$ GeV ($m(\mathrm{LQ}_3^d) < 930$ GeV) are excluded.

Acknowledgements

My biggest thanks goes to my supervisors, Carl Gwilliam and Andrew Mehta, for their patience and guidance over the course of my PhD. I'm extremely grateful to Carl for always making time to help me with my countless questions. I wouldn't have got here without his constant source of physics and programming knowledge and inspirational dedication to his work.

I am also grateful to the rest of the Liverpool ATLAS team and the wider HEP department for contributing to an exciting research environment. Thank you to the ATLAS $bb\tau\tau$ team, for being fantastic people to work with over the years and, most importantly, for their hard work in achieving an excellent result.

During my time at CERN, I met countless great people who contributed to a memorable 18 months. A huge thank you to Chloe, who has been there throughout this entire journey and without whom I would have given up on physics a long time ago, and Heather, for her words of encouragement when I needed them the most.

Thank you to Ben, for your unwavering love and support over my entire time at University and for your patience over the years spent apart. You have been fantastic over the past few months of thesis writing and I will make it up to you.

Finally, thank you to my family. Most importantly, my Mam and Dad, who opened up a whole world of possibilities for me. You provided me with every possible opportunity and forged the path that led me here.

Contents

Chapter 1

Introduction

In 2012, almost five decades after Peter Higgs *et al.* published their famous theoretical predictions [1–4], a new boson was observed by the ATLAS [5] and CMS [6] experiments in proton-proton collisions occurring approximately 100 m beneath the Franco-Swiss border. All measured properties of this boson are consistent with those predicted by the Standard Model of Particle Physics (SM). This is a huge success for theoretical and experimental physicists alike; the culmination of years of work performed by thousands of physicists in global collaborations has – so far – provided confirmation of a mathematically constructed theory.

Whilst the discovery of the elusive Higgs particle has closed the first chapter in probing the mechanism that generates the fundamental particle masses, the success enjoyed in 2012 has sparked a new chapter of high-energy physics (HEP) with two main objectives: the precise determination of the properties of the Higgs boson and the search for new physics beyond the Standard Model (BSM). This thesis aims to contribute to both of these quests by presenting searches for Higgs boson pair-production and third-generation scalar leptoquark pair-production.

The experimental measurement of the Higgs boson mass of $m_h = 125.09 \pm 0.24$ GeV [7] is a huge step forward in achieving a full understanding of the Higgs potential. The final piece in the puzzle could be provided by a precise measurement of Higgs pair-production. Though the SM predicts that this process occurs at a rate that is approximately three orders of magnitude lower than single Higgs production, modifications to the couplings from BSM physics could significantly alter the cross-section, making the process observable with the LHC Run II data set. The Higgs pair-production cross-section could also be boosted by the inclusion of new heavy resonances such as a spin-2 Kaluza-Klein graviton in the bulk Randall-Sundrum model or a heavy scalar Higgs boson in the hMSSM.

This thesis presents a search for pair-produced Higgs bosons decaying to a final state with two b-jets and two τ-leptons using 36.1 fb^{-1} of $\sqrt{s} = 13$ TeV proton-proton collision data collected by the ATLAS detector during Run II of the LHC. The success of the $hh \to b\bar{b}\tau^+\tau^-$ analysis motivated a search for pair-production of third-generation scalar leptoquarks in a similar final state and following a similar analysis strategy, which is also presented in this document. Leptoquarks are hypothetical particles which are predicted by a number of extensions to the SM to provide a connection between the quark and lepton sectors. Their existence is also hinted at by anomalies in measurements of B meson decays, for which they could provide an explanation.

Part I outlines the theoretical background which motivates the searches in Parts III and IV. It begins with an introduction to the SM, which describes the properties and interactions of the fundamental constituents of matter, with a particular focus on spontaneous symmetry breaking and the Higgs boson. A brief description of the motivation for BSM theories leads into an outline of the new physics which could be responsible for Higgs pair-production, followed by a description of the production mechanisms for Higgs pair-production. Theories which predict the existence of leptoquarks are outlined in Section 3.5.

Part II gives an overview of the LHC and the ATLAS detector. It outlines the MC simulations used in both analyses and, crucially, explains the particle reconstruction and identification algorithms used to derive the objects used for data analysis. In particular, Section 10 describes how the efficiency with which light-jets are tagged as b-jets is calculated to accurately describe data observations; the author contributed to this analysis in order to gain ATLAS authorship.

Part III details the search for Higgs pair-production, presenting the methods used to carefully model the large multi-jet and $t\bar{t}$ backgrounds and boost the sensitivity to the elusive signal process. Chapter 13 documents the introduction of multivariate methods to the analysis. Part IV identifies the main deviations of the leptoquark analysis from the di-Higgs strategy and presents the results of the search. The author's main contribution to both searches was to analyse events with a $\tau_{\text{lep}}\tau_{\text{had}}$ final state; for the leptoquark analysis, the author also performed the combination with the $\tau_{\text{had}}\tau_{\text{had}}$ channel.

Part I

Theory and Motivation

Chapter 2

The Standard Model of Particle Physics

Developed over the latter part of the 20^{th} century by a number of physicists throughout the world, the Standard Model of Particle Physics encapsulates our current best understanding of the fundamental particles which make up everything in the Universe and the fundamental forces through which they interact. As a relativistic Quantum Field Theory (QFT), it treats fundamental particles as excited states of quantum fields which are defined at all points in spacetime. It provides a mathematical formulation for three of the four known fundamental forces – the *strong*, *weak* and *electromagnetic* interactions – the fourth being the gravitational force.

The elementary matter particles occur in two basic types, *quarks* and *leptons*, which occur in three *generations*. All stable matter in the Universe is made from particles from the first generation; the heavier particles from the second and third generation (which can be produced in particle colliders) quickly decay to the lighter, more stable particles that belong to the first generation. The three SM forces act between particles via the exchange of corresponding force-carrying particles, which belong to a wider group of particles named *bosons*.

Critically, the entire SM is underpinned by the fundamental concept of *symmetry*. For every symmetry that exists in the SM, there is a corresponding *conservation law*. These conservation laws require that the *quantum numbers* associated with the fundamental particles involved in an interaction sum to the same value before and after the interaction occurs. This ensures that only certain particle interactions and decays are allowed.

However, these symmetries lead to the prediction that all elementary particles are massless. This is reconciled through the process of *electroweak symmetry breaking* and the

Higgs mechanism, which generates particle masses through their interaction with the Higgs field. The Higgs field is associated with a particle named the Higgs boson, the existence of which was experimentally confirmed in 2012.

This chapter gives an outline of the particle content of the SM, before providing a theoretical overview of the strong, *electroweak* and Higgs interactions and finally concentrating on Higgs pair-production, which is the focus of Part III. A full picture of the SM and the theoretical framework upon which it is built is provided by [8–10].

Throughout this thesis, natural units of $\hbar = c = 1$ are assumed, where $\hbar$ is the reduced Planck constant and c is the speed of light.

2.1 Particle Content

All particles in the SM are assumed to be point-like. They can be divided into two categories: *fermions*, which have intrinsic spin $s = 1/2$ and follow Fermi-Dirac statistics, and bosons, which have integer spin and obey Bose-Einstein statistics.

Unlike bosons, fermions obey the Pauli exclusion principle, which states that two or more identical fermions cannot occupy the same quantum state. There are 12 elementary fermions in the SM, which are further split into two categories: quarks, which interact via the strong force, and leptons, which do not. The fermions are listed in Table 2.1 alongside their electric charge, Q, weak isospin, T_3, and mass, m. There are three generations of fermions [11], with each generation containing two types of quark and two types of lepton, each of which has a different electric charge (fractional for quarks and integer for leptons) and mass. Each quark may exist with one of three different values of colour charge: *red*, *green* or *blue*. Each generation contains a lepton (electron, muon or τ-lepton) with electric charge $Q = -1e$, and a corresponding neutrino (electron-neutrino, muon-neutrino or τ-neutrino) with zero electric charge. Whilst the SM predicts that neutrinos are massless, experimental observation of neutrino oscillation[1] implies that they have non-zero mass. This mass has been experimentally constrained to $m_\nu < 2$ eV [12], with the strongest constraints provided by measurements of tritium β-decay.

The electrically charged leptons interact via the electromagnetic and weak forces, whilst the electrically neutral neutrinos interact only via the weak force. All quarks have electric charge and are the only particles to interact via all four fundamental forces. Crucially, quarks interact via the strong force according to their colour quantum number; *colour*

[1] The experimentally established phenomenon of neutrino oscillation involves the mixing of neutrino flavour states (e, μ and τ) with neutrino mass states. As a result of this process, neutrinos created with a given flavour can be later measured with a different flavour.

Generation	Name	Symbol	Electric charge, Q [e]	Weak isospin, T_3	Mass, m [GeV]
Quarks					
1	up	u	$+\frac{2}{3}$	$+\frac{1}{2}$	0.0023
	down	d	$-\frac{1}{3}$	$-\frac{1}{2}$	0.0047
2	charm	c	$+\frac{2}{3}$	$+\frac{1}{2}$	1.275
	strange	s	$-\frac{1}{3}$	$-\frac{1}{2}$	0.095
3	top	t	$+\frac{2}{3}$	$+\frac{1}{2}$	173.0
	bottom	b	$-\frac{1}{3}$	$-\frac{1}{2}$	4.18
Leptons					
1	electron	e	-1	$-\frac{1}{2}$	0.000511
	electron neutrino	ν_e	0	$+\frac{1}{2}$	$< 2 \times 10^{-9}$
2	muon	μ	-1	$-\frac{1}{2}$	0.106
	muon neutrino	ν_μ	0	$+\frac{1}{2}$	$< 2 \times 10^{-9}$
3	τ-lepton	τ	-1	$-\frac{1}{2}$	1.777
	τ neutrino	ν_τ	0	$+\frac{1}{2}$	$< 2 \times 10^{-9}$

TABLE 2.1: The elementary fermions of the SM [12].

confinement requires that quarks are not found individually in nature and are strongly bound to each other in colourless combinations.

For every quark and lepton there exists a corresponding particle with opposite-sign quantum numbers, referred to as its *antiparticle*. The existence of antimatter was first theorised in 1928 by Paul Dirac [13] to explain the appearance of a 'negative-energy' solution to his relativistic version of the Schrödinger equation,

$$(i\partial\!\!\!/ - m)\psi = 0, \tag{2.1}$$

where m is the particle mass. Experimental confirmation of antimatter occurred four years later through the discovery of the positron (the anti-electron) [14]. Though Table 2.1 does not include antimatter, an antiparticle exists for every particle in the table.

The gauge or vector ($s = 1$) bosons are the quanta of the gauge invariant fields which mediate the fundamental forces between particles. They are listed in Table 2.2, which summarises their electric charge, colour, spin and mass.

The electromagnetic force, which occurs between electrically charged particles, has infinite range and therefore it is mediated by a massless particle, the photon. Photons themselves are not electromagnetically charged and therefore cannot self-interact.

Interaction	Name	Symbol	Electric charge, Q [e]	Mass, m [GeV]	Range [m]
electromagnetic	photon	γ	0	0	∞
weak	W boson	W	± 1	80.385	$\sim 10^{-18}$
	Z boson	Z	0	91.1876	
strong	gluon	g	0	0	$< 10^{-15}$

TABLE 2.2: The fundamental forces of the SM [12] and the gauge bosons that mediate them. The charge and mass of each gauge boson and the effective range of each force are given.

The weak interaction is mediated by three massive gauge bosons, the W^+, W^- and Z bosons. The $W^\pm$ bosons are electrically charged (and can also interact via the electromagnetic force), while the Z boson is electrically neutral. The Heisenberg uncertainty principle states that

$$\Delta E \Delta t \geqslant \frac{\hbar}{2}, \tag{2.2}$$

where, in this case, ΔE is the uncertainty on the rest mass of the gauge boson and t is time. As a result, the intermediate vector bosons, and hence the weak interaction, have limited lifetime and range.

The strong interaction is mediated by the massless *gluon*, which couples to colour-charged particles (including itself). Each gluon carries both colour and anti-colour, of which there are eight different combinations, leading to eight gluon types. Although gluons are massless, the strong interaction has a shorter range than the weak and electromagnetic forces – a result of colour confinement, which requires that colour-charged particles cannot exist in isolation and must bind together to form colour-neutral hadrons. Beyond these distances (around the size of a hadron), the *residual* strong force is responsible for interactions between hadrons. This force (also referred to as the *nuclear* force) is transmitted by gluons which form virtual mesons between hadrons.

2.2 Standard Model Theory

The SM is a non-Abelian[2] gauge theory based on a product of the (special) unitary groups $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, where $\mathrm{U}(N)$ refers to a group of $N \times N$ unitary matrices and $\mathrm{SU}(N)$ is the subset of $\mathrm{U}(N)$ with unitary matrices of determinant equal to one. The group $\mathrm{SU}(3)_c$ describes *quantum chromodynamics* (QCD), which parametrises

[2]The group is *non-Abelian* due to the inclusion of the $\mathrm{SU}(3)_c$ and $\mathrm{SU}(2)_L$ groups, both of which have non-commuting gauge transformations. The $\mathrm{U}(1)_Y$ group is Abelian.

the strong interaction, while $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ constitutes the electroweak group, which combines electromagnetism and the weak force.

The SM is based on the principle of symmetry, the mathematics of which is provided by group theory. Noether's theorem [15] states that for every continuous symmetry, there is an associated conservation law. Therefore, each group in the SM introduces a conserved quantity denoted by the indices c, L and Y, which represent *colour*, *weak isospin*[3] and *weak hypercharge* respectively.

The fields in the SM are described according to the *Lagrangian* formalism. The SM Lagrangian density (more commonly referred to as simply the Lagrangian), $\mathcal{L}_{\mathrm{SM}}$, is constructed from four components:

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{QCD}} + \mathcal{L}_{\mathrm{Electroweak}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}}, \tag{2.3}$$

where $\mathcal{L}_{\mathrm{QCD}}$ describes the dynamics of the strong force, $\mathcal{L}_{\mathrm{Electroweak}}$ describes the dynamics of the electroweak force and $\mathcal{L}_{\mathrm{Higgs}}$ and $\mathcal{L}_{\mathrm{Yukawa}}$ are the terms which introduce mass to fermions and gauge bosons respectively (this concept will be introduced in Sections 2.2.3 and 2.2.4).

2.2.1 Quantum Chromodynamics

The theory of quantum chromodynamics (QCD) describes the strong interactions of coloured quarks and gluons. As well as the processes involved in high-energy collisions of elementary particles in colliders, it provides an explanation for the structure of nuclei and the behaviour of neutron stars.

In 1964-65, the colour charge was introduced as an SU(3) gauge degree of freedom [16, 17] to explain the existence of the spin $S = 3/2$ baryon, Δ^{++}, which is composed of three up-quarks with parallel spins and, therefore, would otherwise violate the Pauli exclusion principle.

QCD is a non-Abelian gauge theory based on the SU(3) symmetry group, where the quark fields are represented as a colour triplet, ψ:

$$\psi = \begin{pmatrix} \psi_r \\ \psi_b \\ \psi_g \end{pmatrix}, \tag{2.4}$$

[3] The L here indicates that the weak interaction applies only to left-handed fermions.

with the r, b and g indices corresponding to the quark colour charges: red, blue and green respectively.

The Lagrangian, $\mathcal{L}_{\text{QCD}}$, of the quarks coupled to the gluon fields, is given by

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}(i\gamma^{\mu}D_{\mu} - m)\psi - \frac{1}{4}G^{a}_{\mu\nu}G^{\mu\nu}_{a}, \tag{2.5}$$

where γ^{μ} represents the gamma matrices. The covariant derivative, D_{μ}, is given by

$$D_{\mu} = \partial_{\mu} + ig_{s}T^{a}G^{a}_{\mu}, \tag{2.6}$$

where the gauge coupling term, g_s, specifies the strength of the interaction between the quark and gluon fields. The $T^a = \frac{\lambda^a}{2}$ are the eight generators of the SU(3) gauge group (λ^a represents the Gell-Mann matrices [18]), corresponding to the eight gluons of varying colour charge. The gluon field strength tensor is given by

$$G^{a}_{\mu\nu} = \partial_{\mu}G^{a}_{\nu} - \partial_{\nu}G^{a}_{\mu} - g_{s}f^{abc}G^{b}_{\mu}G^{c}_{\mu}. \tag{2.7}$$

The final term of this equation represents the gluon self interactions, where f^{abc} are the fully antisymmetric SU(3) structure constants. These self couplings contribute to the two distinguishing characteristics of QCD: *asymptotic freedom* [19] and *confinement*.

The strong coupling constant, α_s, is related to g_s via:

$$\alpha_{s} = \frac{g_{s}^{2}}{4\pi}. \tag{2.8}$$

The strength of the strong coupling exhibits a running behaviour related to the energy scale of interactions, Q^2. In the one-loop approximation,

$$\alpha_{s}(Q^{2}) \approx \frac{1}{\beta_{0}\ln\left(\frac{Q^{2}}{\Lambda^{2}}\right)}, \tag{2.9}$$

where Λ is the scale of QCD, known as the *hadronisation scale*, and β_0 is the first order perturbative constant:

$$\beta_{0} = \frac{33 - 2N_{f}}{12\pi}. \tag{2.10}$$

Here, N_f is the number of quark flavours that can appear in the loop; for $Q^2 > m_t^2$ (m_t being the top-quark mass), $N_f = 6$. As large Q^2 corresponds to probing small distances, Equation 2.9 yields two consequences:

- At short distances (large Q^2), α_s is small and quarks behave like free particles. This is referred to as *asymptotic freedom*. Theoretical predictions of QCD processes are made using perturbation theory.
- At larger distances ($Q^2 \to \Lambda^2$), the strong coupling tends to infinity. Colour-charged quarks bind together to form colour-neutral hadrons; this is called *confinement*.

2.2.2 Electroweak Theory

The theory of the electroweak interaction was proposed by Glashow, Salam and Weinberg [20–22] to describe the electromagnetic and weak interactions between quarks and leptons. The electromagnetic and weak interactions are combined as two aspects of the same force. Electroweak theory is based on the combination of the gauge symmetry groups $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. The non-Abelian special unitary group $\mathrm{SU}(2)_L$ describes the fermion doublet fields, whilst an additional $\mathrm{U}(1)_Y$ group is required to include the electromagnetic interaction.

The fermions in Table 2.1 are comprised of components with negative and positive chirality, referred to as *left-* and *right-handed* particles, respectively. Left-handed fermions are organised in $\mathrm{SU}(2)_L$ doublets, ψ_L, with weak isospin component $T_3 = \pm 1/2$, whilst right-handed particles are represented by singlets, ψ_R, with $T_3 = 0$. The left-handed doublets interact through the weak force, while the right-handed singlets remain invariant under transformations of $\mathrm{SU}(2)_L$.

The fermion fields are written as weak isospin doublets, containing left-handed leptons and quarks:

$$\psi_L^{l=e,\mu,\tau} = \begin{pmatrix} \nu_l \\ l \end{pmatrix}_L, \quad \psi_L^{f=1,2,3} = \begin{pmatrix} u_f \\ d'_f \end{pmatrix}_L, \tag{2.11}$$

where the down-type quarks are denoted d'_f since their flavour eigenstates (d', s', b') are not equal to their mass eigenstates (d, s, b), but are related via the Cabibbo-Kobayashi-Maskawa quark-mixing matrix [23].

The Abelian unitary group $\mathrm{U}(1)_Y$ implies the conservation of the weak hypercharge, Y, which is given by the Gell-Mann-Nishijima relation [24, 25]:

$$Y = 2(Q - T_3), \tag{2.12}$$

where Q is the electric charge.

The electroweak Lagrangian, $\mathcal{L}_{\text{Electroweak}}$, is given by

$$\mathcal{L}_{\text{Electroweak}} = \bar{\psi}_L \gamma^\mu D_\mu^L \psi_L + \bar{\psi}_R \gamma^\mu D_\mu^R \psi_R - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} \vec{W}_{\mu\nu} \vec{W}^{\mu\nu}, \tag{2.13}$$

where, for simplicity, the summations over the fermion fields, ψ_L and ψ_R, are not written explicitly but are implied. The first two terms in Equation 2.13 describe the interaction of the left- and right-handed fermion fields, respectively, with the gauge boson fields. The gauge covariant derivative terms, D_μ^L and D_μ^R, are given by

$$\begin{aligned} D_\mu^L &= i\partial_\mu - \frac{g}{2} \vec{\sigma} \cdot \vec{W}_\mu - \frac{g'}{2} Y B_\mu, \\ D_\mu^R &= i\partial_\mu - \frac{g'}{2} Y B_\mu, \end{aligned} \tag{2.14}$$

where $\vec{\sigma}$ represents the Pauli matrices, which generate the $\text{SU}(2)_L$ group. These terms ensure that the Lagrangian is invariant under local gauge transformations by introducing the four gauge fields, W_μ^i (with $i = 1, 2, 3$) and B^μ. The electroweak coupling constants are labelled g and g', where g is the coupling constant for weak isospin and g' is the coupling constant for weak hypercharge; ∂_μ is the partial derivative four-vector.

The final two terms in Equation 2.13 are the Lagrangians of the vector boson fields, which describe their self-interactions. The field strength tensors, $\vec{W}_{\mu\nu}$ and $B_{\mu\nu}$, are given by

$$\begin{aligned} \vec{W}_{\mu\nu} &= \partial_\mu \vec{W}_\nu - \partial_\nu \vec{W}_\mu - g \vec{W}_\nu \times \vec{W}_\mu, \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu. \end{aligned} \tag{2.15}$$

The gauge fields in the Lagrangian mix to form the $W_\mu^\pm$, Z_μ and A_μ fields, whose quantised oscillations correspond to the physical gauge bosons (the A_μ field gives the photon):

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} (W_\mu^1 \mp i W_\mu^2), \\ \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} &= \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \end{aligned} \tag{2.16}$$

where θ_W is the weak mixing (Weinberg) angle, which is related to the SU(2) and U(1) coupling constants, g and g', as:

$$\tan\theta_W = \frac{g'}{g}. \tag{2.17}$$

2.2.3 The Higgs Mechanism

The electroweak Lagrangian in Section 2.2.2 (Equation 2.13) is invariant under local gauge transformations only if the physical $W^{\pm}$, Z and γ gauge bosons are assumed to be massless. However, as seen in Table 2.2, observations of the $W^{\pm}$ and Z bosons indicate that they are, in fact, massive. Explicitly incorporating mass terms in the electroweak Lagrangian would violate gauge invariance.

To reconcile theory with observation, the local electroweak gauge symmetry is broken spontaneously through the inclusion of a new complex scalar field [1–4]. This mechanism, referred to as *spontaneous symmetry breaking* (SSB), also produces mass terms for fermions.

Consider a complex scalar field, ϕ, which is given by

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2), \tag{2.18}$$

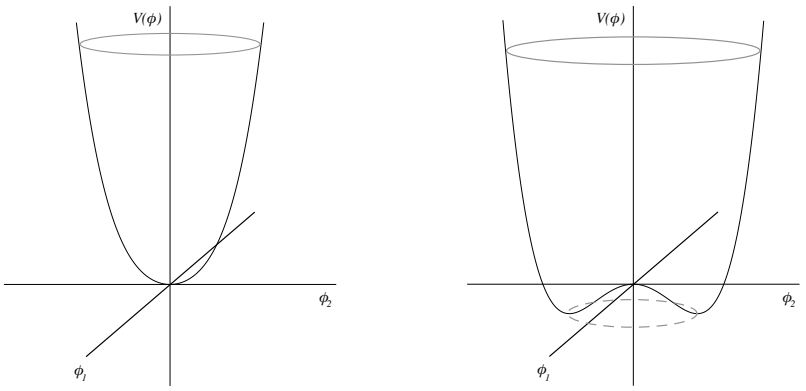
and has a Lagrangian density,

$$\begin{aligned} \mathcal{L} &= (\partial_\mu\phi)^\dagger(\partial^\mu\phi) - V(\phi), \\ V(\phi) &= \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2. \end{aligned} \tag{2.19}$$

The potential, $V(\phi)$, has a non-zero ground state when $\mu^2 < 0$, $\lambda > 0$, as shown in Figure 2.1b. The potential has degenerate minima which form a continuous complex circle around the origin in the (ϕ_1, ϕ_2) plane:

$$\phi^\dagger\phi|_{\text{min}} = \frac{1}{2}(\phi_1^2 + \phi_2^2)|_{\text{min}} = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2}, \tag{2.20}$$

where v is called the *vacuum expectation value* (VEV) of ϕ. The system also has a metastable state at $\phi = 0$ with U(1) symmetry. By choosing an arbitrary vacuum state from the complex circle of minima (by convention, this is chosen to lie along the direction of the real part of ϕ, i.e. $\phi_1 = \frac{v}{\sqrt{2}}$, $\phi_2 = 0$), the gauge invariance is *spontaneously* broken, whilst the Lagrangian itself remains gauge invariant.



(a) $\mu^2 > 0$ (b) $\mu^2 < 0$

FIGURE 2.1: Illustrations of the Higgs potential $V(\phi)$ in terms of the real (ϕ_1) and imaginary (ϕ_2) components of the scalar field, with: (a) $\lambda > 0$ and $\mu^2 > 0$; (b) $\lambda > 0$ and $\mu^2 < 0$. In case (b), the vacuum expectation value (the value of ϕ found at the minima of $V(\phi)$), v, is non-zero.

The scalar field, ϕ, can be written as an excitation around the vacuum state, in terms of two new fields, ξ and η:

$$\phi = \frac{1}{\sqrt{2}}(v + \xi + i\eta), \tag{2.21}$$

which, when substituted into the Lagrangian in Equation 2.19, gives

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) + \frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi) + \mu^2 \xi^2 - \lambda(\xi\eta^2 + \xi^3) - \frac{\lambda}{2}\xi^2\eta^2 - \frac{\lambda}{4}\xi^4 - \frac{\lambda}{4}\eta^4. \tag{2.22}$$

The term $\mu^2\xi^2$ denotes a massive scalar field, ξ, with mass $m_\xi = \sqrt{-2\mu^2}$. However, because no such term exists for the η field, η corresponds to a massless *Goldstone boson*, which results from the spontaneous breaking of a continuous global symmetry [26]. No such boson has been observed.

In fact, the particle masses enter the SM through spontaneous symmetry breaking of the *local* electroweak gauge invariance. A scalar *Higgs* field, Φ, is introduced, written as an $\mathrm{SU}(2)_L$ isospin doublet,

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{2.23}$$

comprising the charged, ϕ^+, and neutral, ϕ^0, complex scalar fields. The Lagrangian is given by

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - \mu^2 (\Phi^\dagger \Phi) - \lambda (\Phi^\dagger \Phi)^2, \tag{2.24}$$

where the covariant derivative for the SU(2) × U(1) symmetry, D_μ, is given by

$$D_\mu = \partial_\mu + \frac{ig}{2} \vec{\sigma} \cdot \vec{W}_\mu + \frac{ig'}{2} Y B_\mu. \tag{2.25}$$

The first term of the Lagrangian describes the interactions of the Higgs field with the electroweak gauge fields, whilst the final two terms, as in Equation 2.19, form a scalar potential, $V(\Phi)$, with $\mu^2 < 0$ and $\lambda > 0$. The minima of the potential $V(\Phi)$ are derived as follows:

$$\Phi^\dagger \Phi|_{\text{min}} = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2)|_{\text{min}} = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2}. \tag{2.26}$$

A vacuum state is chosen such that only the real component of the neutral scalar field has a non-zero value:

$$\begin{aligned} \phi_1 &= \phi_2 = \phi_4 = 0, \\ \phi_3 &= \sqrt{\frac{-\mu^2}{2\lambda}} \equiv v. \end{aligned} \tag{2.27}$$

Three Goldstone bosons arise from the breaking of the SU(2) symmetry. In the *unitary gauge*, the Higgs field, Φ, can be expressed as fluctuations around the minimum, v:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ h(x) + v \end{pmatrix}, \tag{2.28}$$

where $h(x)$ denotes the physical Higgs boson field. By substituting this in Equation 2.24, and using Equations 2.16 and 2.17, the Lagrangian becomes

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} = \; & \frac{1}{2}\partial_\mu h \partial^\mu h + \frac{g^2}{4}(v+h)^2 W_\mu^+ W^{-\mu} + \frac{1}{8}(g^2 + g'^2)(v+h)^2 Z_\mu Z^\mu \\ & - \mu^2 h^2 - \lambda v h^3 - \frac{\lambda}{4} h^4. \end{aligned} \tag{2.29}$$

The masses of the $W^\pm$ and Z bosons can be identified from terms of the form $m_W^2 W_\mu^+ W^{-\mu}$ and $\frac{1}{2} m_Z^2 Z_\mu Z^\mu$, respectively:

$$\begin{aligned} m_W &= \frac{1}{2}gv, \\ m_Z &= \frac{v}{2}\sqrt{g^2+g'^2} = \frac{m_W}{\cos\theta_W}, \end{aligned} \tag{2.30}$$

and the Higgs boson has mass m_h, given by

$$m_h = \sqrt{-2\mu^2} = v\sqrt{2\lambda}. \tag{2.31}$$

The expectation value of the Higgs field, given by $v = (\sqrt{2}G_{\mathrm{F}})^{-1/2} \approx 246$ GeV, is fixed by the Fermi constant, G_{F} [12]. The remaining terms in Equation 2.29 represent VVh and $VVhh$ interactions (where V represents either $W^{\pm}$ or Z) and the trilinear and quadlinear Higgs self-interactions, h^3 and h^4. The trilinear interaction is discussed in Section 2.4. Using Equation 2.30, measurement of the vector boson masses, m_W and m_Z, yields values for θ_W, g and v but not μ and, therefore, does not constrain the Higgs mass, m_h. Equation 2.29 lacks any mass terms involving the photon field A_μ; the photon remains massless because the U(1) symmetry is unbroken.

2.2.4 The Yukawa Interaction

The Higgs mechanism is also responsible for fermion masses through the addition of *Yukawa* mass terms to the SM Lagrangian:

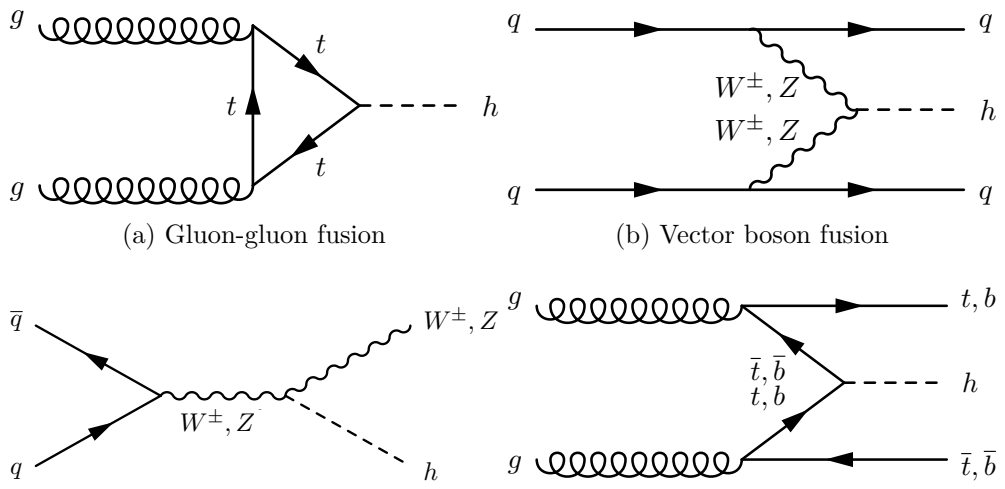
$$\mathcal{L}_{\text{Yukawa}} = -g_f(\bar{\psi}_L\Phi\psi_R + \bar{\psi}_R^{\dagger}\Phi\psi_L), \tag{2.32}$$

where g_f is the Yukawa coupling. Taking the lepton doublets as an example and substituting the Higgs field stated in Equation 2.28 into Equation 2.32 gives

$$\mathcal{L}^{l}_{\text{Yukawa}} = -\frac{g_l}{\sqrt{2}}(h+v)\bar{\psi}\psi, \tag{2.33}$$

where g_l is the lepton Yukawa coupling. The first term represents the interactions of the Higgs and lepton fields and the second is the lepton mass term; the lepton masses are given by

$$m_l = \frac{g_l v}{\sqrt{2}}, \tag{2.34}$$



(a) Gluon-gluon fusion (b) Vector boson fusion

(c) Vector boson associated production (d) $t\bar{t}/b\bar{b}$ associated production

FIGURE 2.2: Leading order Feynman diagrams for Higgs boson production in proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV: (a) gluon-gluon fusion, (b) vector boson fusion, (c) associated production with a vector boson and (d) associated production with a top or bottom quark-antiquark pair.

and the coupling of the Higgs field is proportional to the lepton mass. In the unitary gauge, neutrinos do not interact with the Higgs field or receive a mass; this is due to the charged element of Φ being chosen as zero (see Equation 2.28). The quark masses, m_q, are derived by forming a corresponding Lagrangian, $\mathcal{L}^q_{\text{Yukawa}}$, yielding

$$m_q = \frac{g_q v}{\sqrt{2}}. \tag{2.35}$$

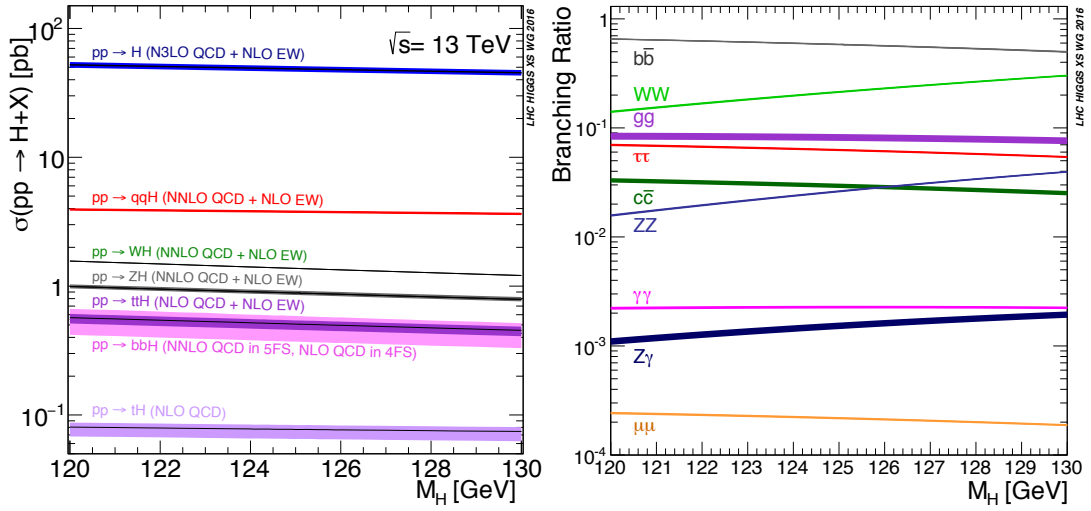
2.3 Higgs Boson Production at the LHC

Proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV produce Higgs bosons via four main mechanisms: gluon-gluon fusion, vector boson fusion, associated production with a vector boson, and associated production with a top or bottom quark-antiquark pair. Feynman diagrams for these four processes are shown in Figure 2.2.

Figure 2.3a shows the cross-sections of these production mechanisms at the LHC for $\sqrt{s} = 13$ TeV. The dominant process is gluon-gluon fusion with a rate more than one order of magnitude greater than that of next most likely process, vector boson fusion. Gluon-gluon fusion occurs via an intermediate top-quark loop.[4]

Figure 2.4 [27] shows ATLAS measurements of the cross-sections for the main Higgs production mechanisms normalised to their SM predictions. These are measured assuming

[4] In fact, all quarks are involved in the loop, but because the Higgs couples to mass, the bottom-quark contribution is highly suppressed relative to that of the top quark.



(a) Production cross-section (b) Branching ratio

FIGURE 2.3: (a) The production cross-sections of the most common Higgs production mechanisms and (b) the branching ratios of the most common Higgs decays at $\sqrt{s} = 13$ TeV as a function of the Higgs mass, m_h. Taken from [28].

SM branching ratios. The cross-sections are generally consistent with the SM expected values.

The SM Higgs boson has a short lifetime and decays rapidly into one of a number of possible final states. Its dominant decay modes are to the heaviest fermions via the Yukawa interaction and to pairs of W and/or Z bosons, one of which is off-shell. It can also decay to pairs of massless gluons or photons via loop diagrams.

Figure 2.3b [28] shows the *branching ratios* for the main Higgs decay modes, which are fixed by the Higgs mass, m_h. The branching ratio, BR, is the fraction of all decays that result in a particular final state, and is given by

$$\mathrm{BR}(h \to x) = \frac{\Gamma(h \to x)}{\Gamma}, \tag{2.36}$$

where Γ is the total decay width and $\Gamma(h \to x)$ is the partial decay width associated with final state x.

As shown by Figure 2.3b, the Higgs boson decay to a pair of b-quarks has the largest branching ratio for a Higgs boson of mass $m_h = 125$ GeV; this is because b-quarks are the heaviest particle which can be produced on-shell from a Higgs decay.

A Higgs boson candidate with mass $m_h \approx 125$ GeV was observed separately by the ATLAS and CMS experiments in 2012 (using data collected at $\sqrt{s} = 7$ TeV in 2011 and $\sqrt{s} = 8$ TeV in 2012) by combining results from searches in the $h \to ZZ^*$, $h \to \gamma\gamma$,

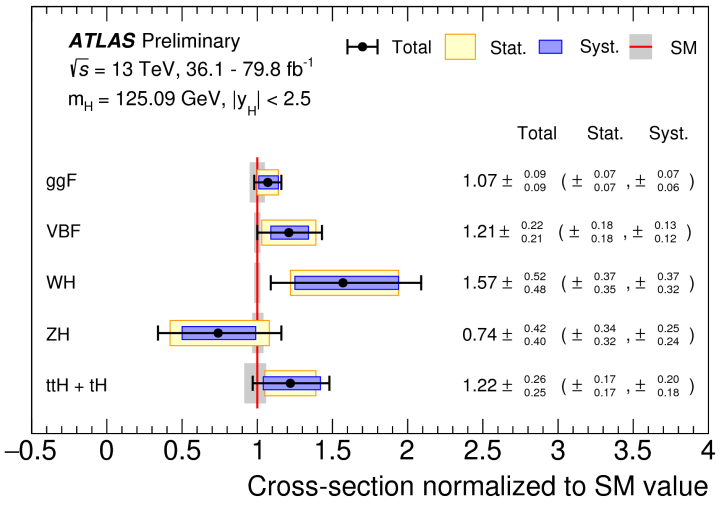


FIGURE 2.4: ATLAS measurements of cross-sections for gluon-gluon fusion ('ggF'), vector boson fusion ('VBF'), vector boson associated production ('WH' and 'ZH') and top-quark associated production ('ttH + tH') normalised to their SM predictions, assuming SM branching ratios. The black error bars, blue boxes and yellow boxes show the total, systematic, and statistical uncertainties on the measurements respectively. The grey bands indicate the theory uncertainties on the cross-section predictions. Taken from [27].

$h \to WW^*$, $h \to b\bar{b}$ and $h \to \tau\tau$ channels [5, 6]. A combination of ATLAS and CMS Run I data samples yielded a Higgs mass measurement of $m_h = 125.09 \pm 0.24$ GeV [7].

The decay of Higgs bosons to τ-leptons was observed with a combined significance of 5.5σ using Run I results from ATLAS and CMS [29–31]. Inclusion of ATLAS and CMS Run 2 data has resulted in a total significance of 6.4σ [32, 33]. Higgs boson decays to pairs of b-quarks were observed using Run II data by ATLAS and CMS in 2018 [34, 35], at a rate consistent with the SM expectation.

Figure 2.5 [27] shows ATLAS measurements of the reduced coupling strength modifiers, $\kappa_F m_F/v$ for fermions ($F = t, b, \tau, \mu$) and $\sqrt{\kappa_V} m_V/v$ for weak gauge bosons ($V = W, Z$), as a function of their masses, m_F and m_V, respectively. All measurements are consistent with the SM prediction that the Higgs boson couples proportional to mass, as shown by the dotted line.

2.4 Higgs Boson Pair-Production at the LHC

The potential of the scalar Higgs field, Φ, written in Equation 2.28, is given by

$$V(\Phi) = \frac{1}{2} m_h^2 h^2 + \lambda v h^3 + \frac{\tilde{\lambda}}{4} h^4, \tag{2.37}$$

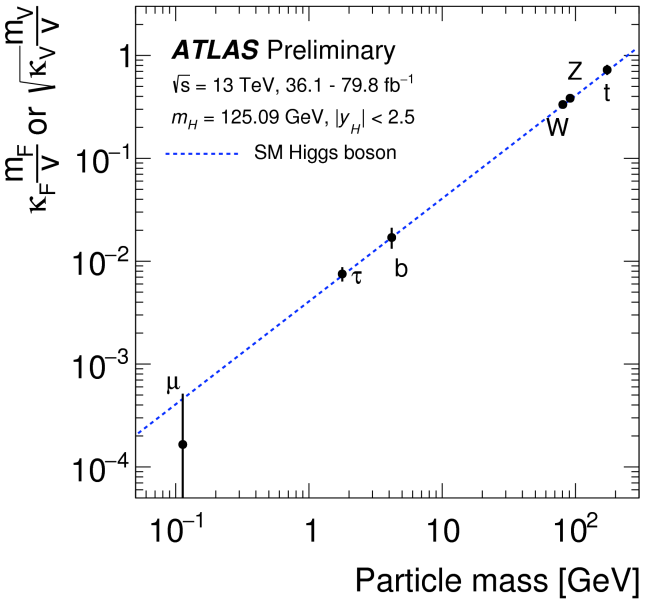


FIGURE 2.5: ATLAS measurements of the reduced coupling strength modifiers, $\kappa_F m_F/v$ for fermions ($F = t, b, \tau, \mu$) and $\sqrt{\kappa_V} m_V/v$ for weak gauge bosons ($V = W, Z$) as a function of their masses, m_F and m_V respectively. The Higgs VEV, v, is assumed to be as predicted by the SM, i.e. $v = 246$ GeV. The SM prediction is shown by the dotted line. Taken from [27].

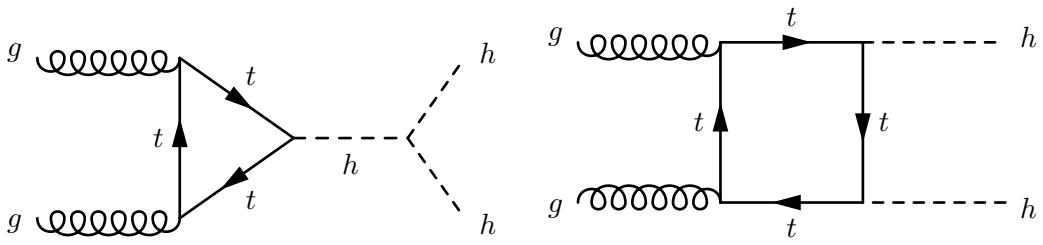
where λ is known as the Higgs trilinear self-coupling and $\tilde{\lambda}$ is the quartic Higgs self-coupling. In the SM, λ is given by

$$\lambda = \tilde{\lambda} = \frac{m_h^2}{2v^2}, \tag{2.38}$$

and, therefore, is predicted when the Higgs mass is known. Experimental measurement of the true value of λ is an important task in determining the shape of the Higgs potential; any deviation from the predicted value implies the existence of BSM phenomena. By experimentally measuring m_h and λ, the value of the VEV, v, can be calculated and compared with the predicted value - a direct test of electroweak symmetry breaking and, hence, the SM.

Terms of the form $\lambda v h^3$ and $\frac{\lambda}{4} h^4$ (assuming $\lambda = \tilde{\lambda}$) in the SM Lagrangian, Equation 2.29, generate di-Higgs and triple-Higgs production, respectively. Higgs pair-production is the only LHC process in which the coupling between three Higgs bosons can be tested at tree level.

At the LHC, di-Higgs production proceeds mainly via gluon-gluon fusion (as for single-Higgs production) via two categories of processes:



(a) 'Triangle' (b) 'Box'

FIGURE 2.6: Feynman diagrams for the two processes by which Higgs pair-production proceeds in the SM: (a) an s-channel virtual Higgs boson propagator decays to a Higgs pair via the Higgs self-coupling, λ; (b) two Higgs bosons are directly radiated from a heavy quark. These diagrams are shown at leading order.

- a virtual Higgs boson, produced by the same mechanism as in single Higgs production, decays to a Higgs pair, involving the self-coupling, λ;
- the two Higgs bosons are both directly radiated from a heavy (top-)quark in a process involving only the Yukawa couplings.

At leading order (LO), the two mechanisms can be attributed to 'triangle' and 'box' Feynman diagrams, as shown in Figure 2.6a and 2.6b respectively. At next-to-leading-order (NLO), the diagram topologies become more complicated and the triangle approximation is no longer valid. The cross-sections of both processes suffer from a reduced phase space due to the inclusion of two heavy final state particles. Destructive interference [36] between the two diagrams in Figure 2.6 further reduces the cross-section for Higgs pair-production, such that an observation is even more challenging to obtain. The cross-sections for the production of Higgs pairs at the LHC are around three orders of magnitude smaller than for single-Higgs production. The SM cross-section for Higgs pair-production by gluon-gluon fusion, calculated using the full top-quark mass dependence [37], is given by

$$\sigma_{\text{SM}} = 33.49^{+1.44}_{-2.00}(\text{scale}) \pm 1.67(\text{theory}) \pm 0.70(\text{PDF}) \pm 0.77(\alpha_s) \text{ fb [28]}. \tag{2.39}$$

2.4.1 The $b\bar{b}\tau^+\tau^-$ Decay Channel

For two Higgs bosons originating from the same interaction, the final state can be one of many possible combinations of single-Higgs decays. The branching ratios associated with some of these are shown in Figure 2.7. The $hh \to b\bar{b}b\bar{b}$ decay is dominant, accounting for 33.7% of all decays, followed by $hh \to b\bar{b}WW^*$, which occurs 25.0% of the time. Whilst these decays have the largest branching ratios, the analyses are subject to large multi-jet and $t\bar{t}$ backgrounds which are difficult to separate from the signal. On the other hand,

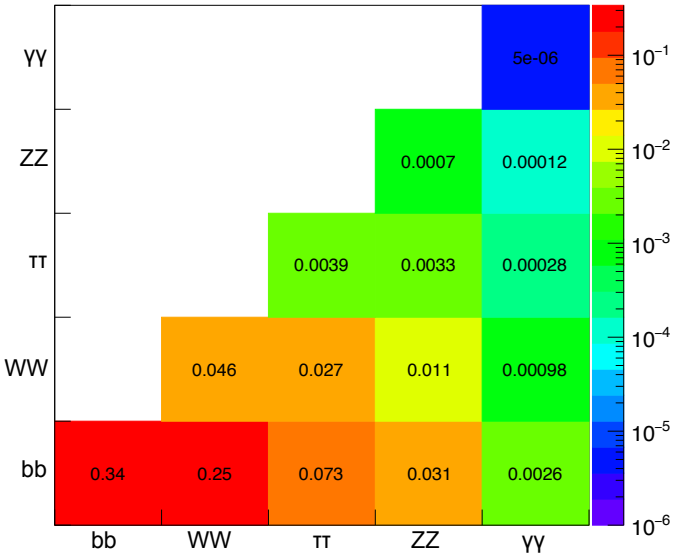


FIGURE 2.7: The most common di-Higgs final states and their branching ratios.

the $hh \to b\bar{b}\gamma\gamma$ decay boasts an exceptionally clean signature but suffers from a very low branching ratio.

Processes where one Higgs decays to b-quarks and the other to τ-leptons, i.e. $hh \to b\bar{b}\tau^+\tau^-$, offer a compromise: a relatively clean signature with better discrimination against multi-jet and $t\bar{t}$ processes than the $hh \to b\bar{b}b\bar{b}$ and $hh \to b\bar{b}WW^*$ final states, but a higher rate than $hh \to b\bar{b}\gamma\gamma$. With a branching ratio of 7.27%, the $hh \to b\bar{b}\tau^+\tau^-$ decay presents a promising channel to search for Higgs boson pair-production.

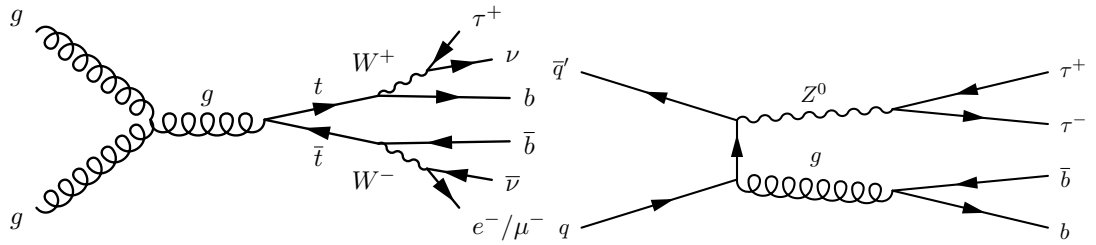
2.4.1.1 τ-Lepton Decays

As the heaviest lepton in the SM, the τ-lepton decays before reaching the detector volume and can only be observed through its decay products. The τ-lepton decays either leptonically (τ_{lep}) or hadronically (τ_{had}) via the weak interaction, with branching fractions BR(τ_{lep}) $\approx$ 35.2% and BR(τ_{had}) $\approx$ 64.8% [12].

Leptonic decays are characterised by two processes: $\tau \to e\nu_e\nu_\tau$ and $\tau \to \mu\nu_\mu\nu_\tau$. For the purposes of the searches presented in this thesis, these decays are simply tagged by the identification of an electron or muon.

Hadronic decays of τ-leptons usually result in either one or three charged pions, some neutral pions and a τ-neutrino (ν_τ or $\bar{\nu}_\tau$). This results in an experimental signature of a collimated calorimeter shower with either one or three associated tracks.

A decay channel involving a pair of τ-leptons, such as the $hh \to b\bar{b}\tau^+\tau^-$ decay (and the leptoquark decay, $\text{LQ}_3^{u,d}\overline{\text{LQ}_3^{u,d}} \to b\tau^+\bar{b}\tau^-$, described in Section 3.5), can be further



(a) $t\bar{t}$ (b) $Z \rightarrow \tau\tau \ + b\bar{b}$

FIGURE 2.8: Leading order Feynman diagrams depicting two important backgrounds for the analyses in this thesis: (a) the pair-production of top-quarks and (b) Z boson production in association with b-jets.

divided into channels involving different combinations of τ-lepton decays: $\tau_{\text{had}}\tau_{\text{had}}$, which makes up 42.0% of di-τ decays; $\tau_{\text{lep}}\tau_{\text{had}}$, which makes up 45.6%; and $\tau_{\text{lep}}\tau_{\text{lep}}$, which occurs the remaining 12.4% of the time. The searches presented in Parts III and IV were performed in the $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels, which were combined to produce the final results. The work of the author is centred on the $\tau_{\text{lep}}\tau_{\text{had}}$ channel.

2.5 Background Processes

The dominant background process in both analyses in this thesis is the pair-production of top-quarks, referred to as $t\bar{t}$ production. Figure 2.8a, the Feynman diagram for this process, illustrates how the decay of the two top-quarks mimics the final state of the signal process. The top (or anti-top) quark decays to produce a b-jet and a W boson. The W bosons then decay, one to a hadronically decaying τ-lepton and neutrinos, and the other to either a light lepton (which is misidentified as having decayed from a τ-lepton) and neutrinos or a leptonically-decaying τ-lepton and neutrinos. The W boson may also decay to jets, one of which is misidentified as a hadronic τ-lepton. In this document, a jet that is misidentified as hadronic τ-lepton may be referred to as a fake-τ_{had}.

Another important background process is that where a Z boson, which is produced in association with any combination of b-, c- and light-flavour jets, decays to a pair of τ-leptons. The Feynman diagram for this process is shown in in Figure 2.8b.

Chapter 3

Physics Beyond the Standard Model

The current formulation of the SM was finalised in the 1970s as a result of the experimental confirmation of the existence of quarks. Since then, the power of the SM has been proven through the experimental discoveries of the top quark, the W and Z bosons and the Higgs boson - all of which were theoretically predicted by the SM. The majority of the measurements made by experimental physicists are in strong agreement with the values predicted by the SM.

However, the SM does have a number of shortcomings which indicate that it is not a complete theory of physics, some of which are outlined in Section 3.1. As stated in Chapter 2, the gravitational force is a glaring omission. Though gravity is successfully encompassed by general relativity, this is not compatible with QFT. Further, experimental evidence shows that the matter particles in the SM constitute only a fraction of the mass in the observable Universe, alluding to the existence of a hypothetical form of matter named *dark matter*. Other issues with the SM include a lack of explanation for the observed matter-antimatter asymmetry, and the hierarchy problem; some of these are summarised in Section 3.1.

A number of BSM theories aim to address one or several of these obstacles. Though a number of other models exist, of particular relevance to this thesis are supersymmetry (SUSY) and the Randall-Sundrum (RS) model (Sections 3.2 and 3.3, respectively), both of which introduce new particles (some of which may be produced in high-energy hadron-hadron collisions) with a view to solving the hierarchy problem. SUSY proposes a symmetry between fermions and bosons which, for each particle of one group, introduces a *superpartner* which belongs to the other group. The RS model is based on the existence of a fifth dimension beyond the usual four (the three space dimensions and time). It

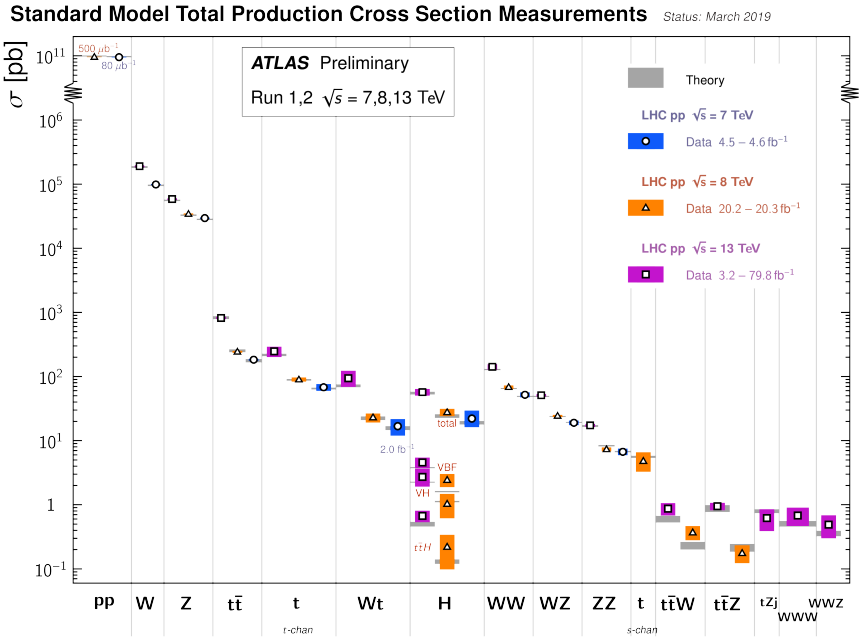


FIGURE 3.1: Summary of several Standard Model total production cross-section measurements, corrected for leptonic branching fractions, compared to the corresponding theoretical expectations. Taken from [38].

introduces *gravitons* and, crucially, predicts their masses to be on a scale that is accessible at the LHC. As will be shown in Section 3.4, BSM theories such as these could modify the rate at which Higgs pair-production occurs.

Finally, a number of extensions to the SM predict the existence of *leptoquarks*, which allow interactions between quarks and leptons. The theoretical and experimental motivation for leptoquarks is summarised in Section 3.5.

3.1 Shortcomings of the Standard Model

The SM is a hugely successful theoretical framework, predicting the existence of particles before their experimental discovery and providing accurate and precise predictions of physical parameters. This is shown in Figure 3.1, which compares several cross-section measurements with the corresponding theoretical predictions.

However, the SM cannot be described as a complete theory of physics. A number of open questions, some of which are outlined here, reveal the need for modifications or extensions to the current framework.

Gravity

Of the four fundamental forces, only three - the strong, weak and electromagnetic forces

- are included in the SM. Attempts to formulate a quantum theory of gravity result in a theory which is not renormalisable (i.e. it predicts infinite values for observables such as particle masses). However, because gravity is many orders of magnitude weaker than the strong, weak and electromagnetic forces (at the energies associated with typical particle interactions), SM predictions of particle interactions remain robust.

Dark matter and dark energy

Measurements of the cosmic microwave background [39] indicate that the matter described by the SM only constitutes $\sim 5\%$ of the energy content in the Universe. Measurements of galaxy rotation curves [40] and gravitational lensing [41] provide evidence for dark matter, which contributes a further $\sim 27\%$. Dark matter does not interact through the electromagnetic or strong[1] interaction and the only SM candidate, the neutrino, whilst being dark and massive, has been shown to only constitute a tiny fraction of the necessary mass. The final $\sim 68\%$ of the total energy in the observable Universe is known as dark energy, which is required to explain the indications that the Universe is expanding at an accelerating rate [43].

Matter-antimatter asymmetry

We observe a clear imbalance in baryonic matter and antimatter in the observable Universe, for which the SM provides no explanation. In the early stages of the Universe, matter, antimatter and energy were in an equilibrium state, with a temperature high enough that pair-production and annihilation could occur at similar rates. However, as the Universe cooled, the reaction could no longer proceed in both directions; the matter and antimatter could only annihilate. If the properties of particles and antiparticles are symmetric, such that matter and antimatter are produced at the same rate, this process should leave only energy remaining. In fact, in addition to a large amount of energy, a small amount of matter survived, with a ratio of matter to energy of 10^{-19}. For an initially symmetrical Universe to give rise to an asymmetrical Universe requires an interaction named *baryogenesis* which, according to the three Sakharov conditions [44]: violates baryon number, B, violates C- and CP-symmetry, and occurs out of thermal equilibrium. The CP violation in the SM is too small to be responsible for the matter-antimatter asymmetry.

Neutrino masses

In the SM, neutrinos are included as left-handed particles represented by Weyl spinors. As their chirality is a permanent property, they are required to be massless. Further, only left-handed neutrinos have been observed in nature so far. However, observations of neutrino oscillations [45, 46] prove that neutrinos are, in fact, massive. All other fermions

[1] In most theories, dark matter candidates are not strongly-interacting, though some do propose Strongly Interacting Massive Particles (SIMPs) [42].

in the SM are massive and are comprised of components with negative and positive chirality, forming Dirac spinors. The inclusion of neutrino mass would suggest that neutrinos also form Dirac spinors, requiring the introduction of right-handed neutrinos. An alternative theory is that neutrinos are Majorana fermions, i.e. that neutrinos and antineutrinos are the same particle. If neutrinos are Majorana particles, then lepton number-violating processes such as neutrinoless double β-decay would be allowed; this process provides a direct probe for new physics.

Free parameters

The most general SM Lagrangian depends on 19 numerical constants, such as particle masses and coupling strengths, for which the SM provides no predictions. These values are constrained only by experiment. When extending the SM to accommodate massive neutrinos, a further seven parameters must be incorporated. Further, the generational structure of the fermions and the differences in the particle masses between generations is not explained.

Hierarchy problem

The hierarchy problem in particle physics has two facets relating to the scale of the gravitational force and the Higgs mass. The first arises from the disparity in strength between the weak force and gravity; the ratio of the Fermi constant (which sets the weak scale), G_F, to the Newton constant (which determines the strength of the gravitational force), G_N, is

$$\frac{G_F}{G_N} \sim 10^{33}. \tag{3.1}$$

There is no theoretical principle to explain why the two forces occur at such different energy scales.

Another aspect of this problem appears when we adopt the QFT framework. The bare mass of the Higgs boson receives quantum corrections from one-loop diagrams including virtual heavy fermions according to

$$\Delta m_h^2(f) = -\frac{(y_f^2 \Lambda^2)}{16\pi^2}, \tag{3.2}$$

and from bosonic loops according to

$$\Delta m_h^2(s) = \frac{(\lambda_S^2 \Lambda^2)}{16\pi^2}, \tag{3.3}$$

where y_f and λ_S are the coupling strengths and Λ is the highest scale of the theory. The highest energy at which a QFT is valid is the Planck mass, M_{Pl}. With quadratic divergences up to $\Lambda = M_{\text{Pl}}$, to restore the Higgs mass to its observed value without introducing new physics would require a cancellation of one part in $\sim 10^{28}$.

3.2 Supersymmetry

A popular solution to the hierarchy problem discussed in Section 3.1 is a class of theories known as *supersymmetry* (SUSY). SUSY cancels the quadratic divergences of the Higgs mass by creating a new symmetry between fermions and bosons. This symmetry introduces superpartners for each SM particle; each fermion has a corresponding boson partner and vice versa. The loop corrections to the Higgs boson mass cancel between particles and their superpartners. For a complete cancellation to occur, particles and their superpartners would be required to have the same mass; however, the prospect of new physics at such low energy scales has been experimentally excluded. Therefore, supersymmetry, should it exist, is a spontaneously broken symmetry and the loop cancellations are not exact. Some fine-tuning is required to achieve the natural Higgs boson mass and the hierarchy problem becomes the *little hierarchy problem* [47].

Whilst addressing the hierarchy problem, supersymmetric models with R-parity conservation introduce a potential dark matter candidate in the form of the lightest supersymmetric particle (LSP), usually a stable neutralino [48]. The LSP is unable to decay into any SM particle without violating R-parity.

One supersymmetric extension to the Standard Model is the Minimal Supersymmetric Standard Model (MSSM), named as such because it considers only the minimum number of new particle states and interactions consistent with phenomenology [49].

The MSSM is an example of a two-Higgs doublet model (2HDM) of type II [50], which means that it requires two Higgs doublets, Φ_u and Φ_d, to break the electroweak symmetry by generating mass for both up- and down-type quarks and charged leptons:

$$\Phi_u = \begin{pmatrix} \phi_u^+ \\ \phi_u^0 \end{pmatrix}, \quad \Phi_d = \begin{pmatrix} (\phi_d^0)^* \\ -\phi_d^- \end{pmatrix}. \tag{3.4}$$

The doublets Φ_u and Φ_d have hypercharge $Y = 1$ and $Y = -1$ respectively, and are expressed in terms of the charged and neutral components, $\phi_i^\pm$ and ϕ_i^0 $(i = u, d)$ [51].

This leads to a total of five Higgs particles: two CP-even states, h and H, a CP-odd state, A, and two charged $H^\pm$ states. At tree level, the MSSM Higgs sector is described

in terms of one Higgs mass, usually m_A, and $\tan\beta$, the ratio of vacuum expectation values of the two Higgs fields, given by

$$\tan\beta = \frac{v_d}{v_u}. \tag{3.5}$$

In the *decoupling regime*, where the mass of the CP-even state, m_A, is large ($m_A \gg m_Z$), the neutral CP-even state, h, is light and displays almost exactly the same properties as the SM Higgs boson. The H and $H^{\pm}$ states are heavy and decouple from the massive gauge bosons, such that the MSSM Higgs sector maintains consistency with the unique Higgs in the SM.

However, the mass of the lightest CP-even Higgs in the MSSM is subject to an upper bound of approximately 130 GeV, which ensures that the SUSY breaking scale, m_S, is not too high (to avoid excessive fine-tuning in the model). For the lightest CP-even state to coincide with the one discovered at the LHC would imply a high value for m_S and place very strong constraints on the MSSM parameters.

The hMSSM (the *habemus* MSSM or the MSSM we currently have) [52] is a benchmark scenario based on the MSSM with the assumption of $m_h \approx 125$ GeV. Fixing m_h removes the explicit dependence of the Higgs sector on other SUSY parameters through radiative corrections, such that the model can be described in terms of m_A and $\tan\beta$ as at tree-level.

3.3 The Randall-Sundrum Kaluza-Klein Model

The existence of higher dimensions was first theorised in the Kaluza-Klein (KK) model, a field theory which unifies the electromagnetic forces and four-dimensional gravity as components of a single five-dimensional gravitational field [53, 54]. The very small fifth dimension is compact and periodic, taking the form of a circle with radius R. The space obtained is the product of the four-dimensional Minkowski space, $\mathbb{R}^{1,3}$, with the circle, denoted S^1: $\mathcal{M} := \mathbb{R}^{1,3} \times S^1$.

Reducing a five-dimensional massless scalar field, ϕ, to four-dimensional Minkowski space, yields an infinite tower of increasingly massive scalars, known as the *Kaluza-Klein tower*. For smaller R, i.e. a smaller extra dimension, the lowest-mass state of the tower becomes more massive and, therefore, more difficult to detect. Furthermore, the five-dimensional theory also yields an infinitesimal U(1) gauge symmetry in four-dimensional space, corresponding to electromagnetism.

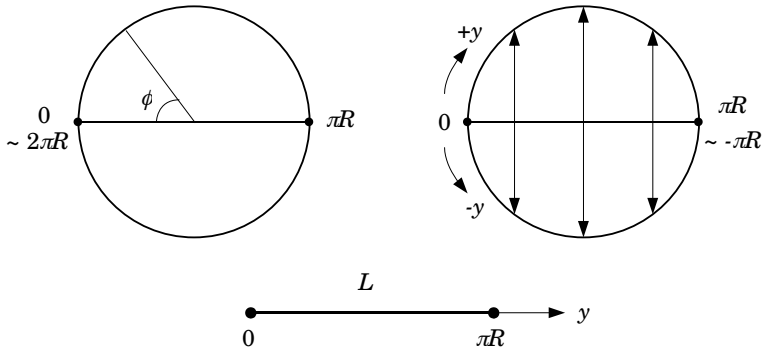


FIGURE 3.2: The geometry of the warped extra dimension, known as an orbifold. The extra dimension is compactified on a circle whose upper and lower halves are identified. Two 3-branes exist: one at $y = 0$ and one at $y = \pi R \equiv L$. Taken from [57].

The Randall-Sundrum (RS) [55, 56] model builds upon KK theory by proposing a warped fifth dimension compactified on a circle whose upper and lower halves are identified. This construction entails two points, at opposite sides of the circle (at $y = 0$ and $y = \pi R \equiv L$), upon each of which a four-dimensional world, or *3-brane*, exists; the five-dimensional *bulk* resides in the anti-de Sitter space between the two branes. The geometry of this model is shown in Figure 3.2 [57].

In the RS1 model (which is the model proposed in [55, 56]), the matter fields are confined to the second brane at $y = L$. The Higgs scalar acts as defined in the SM, but the vacuum expectation value (VEV), v, is exponentially suppressed:

$$v_{eff} = e^{-kL} v, \tag{3.6}$$

where k is the warp factor of the fifth dimension. Because all mass parameters in the SM are based on the Higgs VEV, this means that all mass parameters on the second brane are submitted to an exponential suppression. If the value of the bare Higgs mass is of the order of the Planck mass, M_{Pl}, the physical Higgs mass could be warped down to its observed value via this mechanism. It is for this reason that the brane at $y = 0$ is often referred to as the *Planck brane*, whilst the brane at $y = L$ is the *TeV brane*. Figure 3.3 shows the two 3-branes and the exponential suppression of the weak scale along the fifth dimension.

The hierarchy problem is solved because whilst the weak scale is exponentially suppressed along the extra dimension, the gravity scale is largely independent of it, naturally generating a hierarchy between the weak and gravitational forces.

The presence of two branes forces the masses of the KK states to be quantised according to

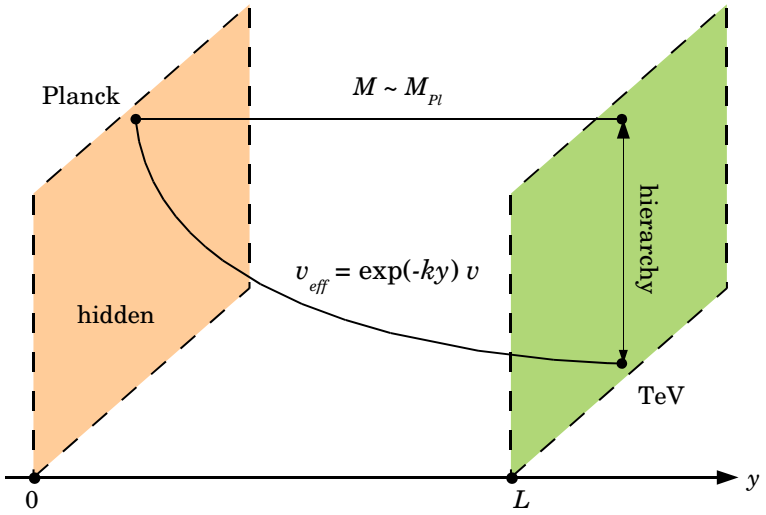


FIGURE 3.3: The Randall-Sundrum setup with two 3-branes - one at $y = 0$ and one at $y = \pi R \equiv L$. An exponential hierarchy is generated along the warped extra dimension. Taken from [57].

$$m_n = ke^{-kL} j_n, \tag{3.7}$$

where j_n are the zeroes of the Bessel function, $J_1(j_n) = 0$. Fixing the factor $\exp(-kL)$ to solve the hierarchy problem yields masses m_n of order TeV with order TeV splitting [58]. Therefore, the first KK gravitons in the tower could be accessible at the LHC.

The decay widths of the KK gravitons are proportional to

$$c = \frac{k}{\bar{M}_{\mathrm{Pl}}}, \tag{3.8}$$

where $\bar{M}_{\mathrm{Pl}}$ is the reduced four-dimensional Planck mass, related to the Planck mass by:

$$\bar{M}_{\mathrm{Pl}} = \frac{M_{\mathrm{Pl}}}{\sqrt{8\pi}} = \sqrt{\frac{\hbar c}{8\pi G}}, \tag{3.9}$$

where G is the gravitational constant.

A key feature of the RS1 model is that the couplings of the KK excitations to SM particles are of order TeV, implying that they would be produced at the LHC with observable rates for masses up to the kinematic limit [58]. This would produce clear signals in di-lepton and di-photon final states which have not been observed experimentally.

A well-motivated extension of the original RS1 model is known as the bulk RS model [59]. For the model to provide a solution to the hierarchy problem, only the Higgs

particle is required to be confined to the TeV brane. The other matter fields are allowed to propagate in the fifth dimension. The SM particles are described as the zero-modes of the five-dimensional fields. The first- and second- generation fermions, which have small Yukawa couplings to the Higgs, are localised near the Planck brane, whilst the top-quark is localised near the TeV brane, accounting for its large Yukawa coupling to the Higgs. As such, the bulk RS model provides a natural hierarchy for the fermion masses.

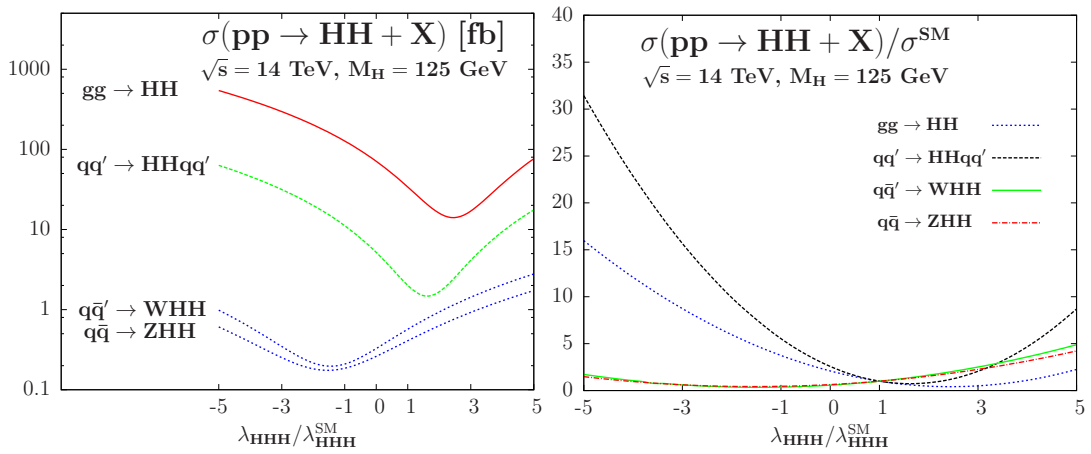
The KK gravitons themselves are localised near the TeV brane and, therefore, their couplings to light fermions are highly suppressed. The KK gravitons have a profile which is similar to that of the Higgs, such that their production via quark-antiquark annihilation is almost negligible, whilst the gluon-gluon fusion production mechanism, though suppressed relative to the RS1 model, remains non-negligible. The dominant decay of the KK gravitons is to top-quarks and Higgs bosons, due to their profile being near the TeV brane, where the Higgs and top-quark also reside.

3.4 Higgs Boson Pair-Production Beyond the Standard Model

3.4.1 Non-Resonant Higgs Boson Pair-Production

Di-Higgs production occurs non-resonantly in the SM, as outlined in Section 2.4, at a very low cross-section due to destructive interference between the two gluon-gluon fusion Feynman diagrams. Modifications to any of the SM couplings involved in either process, i.e. the trilinear Higgs self-coupling, λ, or the top Yukawa coupling, would vary (either up or down) the interference between the processes and, therefore, the cross-section. Figure 3.4 shows how the di-Higgs production cross-sections vary with λ for a collision energy of $\sqrt{s} = 14$ TeV [60]. For example, reversing the sign of λ with respect to its SM value would quadruple the gluon-gluon fusion pair-production cross-section.

Further variations of the $gg \rightarrow hh$ process could occur due to new couplings, such as a direct coupling of two fermions to two Higgs bosons, or new particles, such as heavy quark partners in the loop or additional virtual Higgs bosons. In *composite Higgs models* [61, 62], the Higgs boson is a bound state of new strong interactions which is not a point-like particle but is of finite size. In such models, the Higgs couplings to SM particles and the Higgs self couplings are modified with respect to the SM. Finally, Higgs pair-production is also sensitive to anomalous couplings such as the two-Higgs-two-fermion coupling which emerges in composite Higgs models.



(a) Total cross-section

(b) Ratio of given cross-section to SM cross-section

FIGURE 3.4: The sensitivity of the various Higgs pair-production processes to the trilinear SM Higgs self–coupling at a centre-of-mass energy of $\sqrt{s} = 14$ TeV: (a) the total cross-sections and (b) the ratio between the cross-sections at a given λ/λ_{SM} and the cross-sections at $\lambda/\lambda_{SM} = 1$. The gluon-gluon fusion cross-sections are labelled 'gg → HH'. Taken from [60].

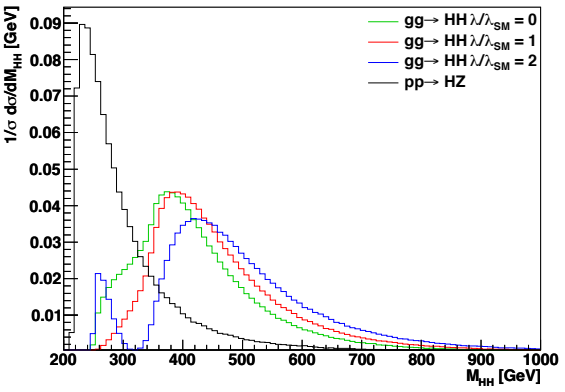


FIGURE 3.5: Normalised distributions of m_{hh} for $\lambda/\lambda_{SM} = 0, 1, 2$, with a typical background distribution provided by $q\bar{q} \to Zh$ (where the Z boson fakes a Higgs). Taken from [60].

Figure 3.5 [60] shows how the distribution of the mass of the Higgs pair, m_{hh}, depends on the value of the trilinear coupling, λ. Distributions of m_{hh} are shown for gluon-gluon fusion with $\lambda/\lambda_{SM} = 0, 1, 2$.

3.4.2 Resonant Higgs Boson Pair-Production

A heavy resonance with mass greater than twice the Higgs mass could decay to a pair of SM Higgs bosons, as shown in Figure 3.6. Whilst the SM does not contain such a particle, a number of BSM theories such as the ones covered earlier in this chapter contain particles which could decay to a Higgs pair. Those studied in Part III are:

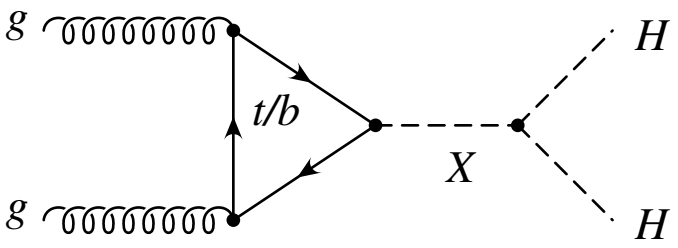


FIGURE 3.6: Feynman diagram for a heavy resonance, X, produced by gluon-gluon fusion, decaying to a pair of SM-like Higgs bosons. The SM-like Higgs bosons, here labelled H, are referred to as h in the text.

- The Randall-Sundrum (RS) model, outlined in Section 3.3, contains a tower of spin-2 KK graviton states, G, with TeV scale masses. Their decays are dominated by top-quarks and Higgs bosons.

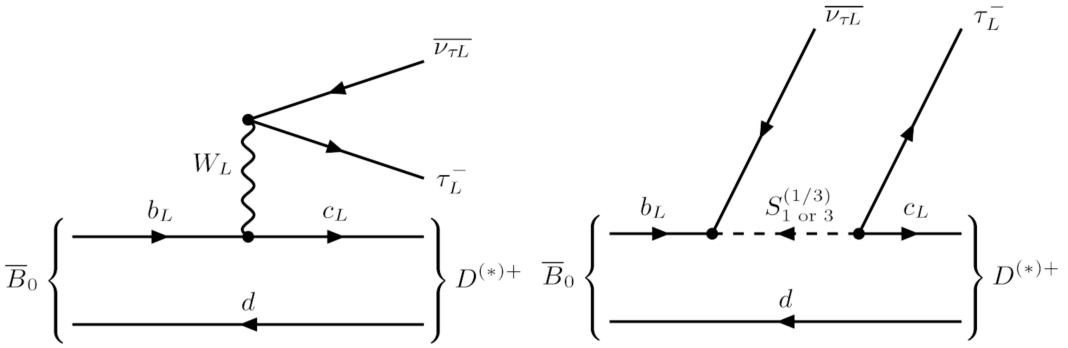
- Two-Higgs-doublet models, such as the hMSSM (described in Section 3.2) contain five Higgs particles, including two CP-even neutral states, usually labelled h and H. The former is lighter than the latter and is associated with the SM-like Higgs boson observed at the LHC. With sufficiently high mass, a heavy Higgs, H, could decay to a pair of the lighter h.

3.5 Leptoquarks

The structural symmetry between the quark and lepton sectors in the SM hints toward a more fundamental relationship between them. A number of extensions of the SM permit interactions between quarks and leptons, mediated by new hypothetical particles named *leptoquarks*. Examples include technicolour models [63, 64], superstring theories [65], composite models [66] and SU(5) grand unification [67].

Leptoquarks are either scalar (spin-0) or vector (spin-1) colour-triplet bosons which carry both lepton and baryon number and have fractional electric charge. Scalar leptoquarks couple to quark-lepton pairs via Yukawa interactions, the couplings of which are determined by two parameters: a model parameter, β, and the coupling parameter, λ. On the other hand, vector leptoquarks have additional magnetic moment and electric quadrupole interactions with strengths proportional to two couplings, κ_G and λ_G.

A general theory of leptoquark production and decay is presented by the model-independent Buchmüller-Rückl-Wyler (BRW) approach [68], in which leptoquark interactions are renormalisable and invariant under the $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)Y$ SM gauge group. In the BRW model, leptoquarks couple only to SM particles and their interactions conserve both lepton and baryon number. Further, in the *minimal* BRW model, leptoquarks are



(a) SM (b) Leptoquark

FIGURE 3.7: Feynman diagrams of contributions to the $B \to D^*\tau\nu$ process from (a) the SM and (b) a hypothetical third-generation scalar leptoquark. The leptoquark in (b) is labelled $S_{1 \text{ or } 3}^{(1/3)}$. Taken from [76].

assumed to interact only with leptons and quarks of the same family, giving three generations of leptoquarks: LQ_1, LQ_2 and LQ_3.

Independent measurements of B meson decays at Belle [69–71], BaBar [72, 73] and LHCb [74] have yielded measurements which persistently deviate from the SM predictions. These include discrepancies in the charged current process, $B \to D^*\ell\nu$, which is mediated by the W boson at tree level in the SM, specifically in the ratio between the branching ratios for $\ell = \mu$ and $\ell = \tau$,

$$R(D^*) = \frac{\text{BR}(B \to D^*\tau\nu)}{\text{BR}(B \to D^*\mu\nu)}. \tag{3.10}$$

An excess of $B \to D^*\tau\nu$ events is observed at a level of 4σ [75]. These measurements challenge lepton flavour universality and may suggest new physics beyond the SM. The existence of a third-generation scalar leptoquark would introduce a new diagram for the $B \to D^*\tau\nu$ process (shown in Figure 3.7b) which would interfere constructively with the SM process (shown in Figure 3.7a).

In proton-proton collisions, leptoquarks can be produced either singly or in pairs. The pair-production mechanism is independent of the coupling parameter, λ. The main production modes of leptoquark pairs, shown in Figure 3.8 in the form of Feynman diagrams, are gluon-gluon fusion (which is dominant) and quark-antiquark annihilation. The cross-section is dependent on the leptoquark mass, $m(\text{LQ})$.

Third generation leptoquarks can be either up- (LQ_3^u) or down-type (LQ_3^d), the decays of which are shown in Figure 3.9a and 3.9b, respectively. Up-type (down-type) third-generation leptoquarks can decay to either a top-quark (bottom-quark) and a neutrino, or a bottom-quark (top-quark) and a τ-lepton. The coupling to the τ-lepton is given by

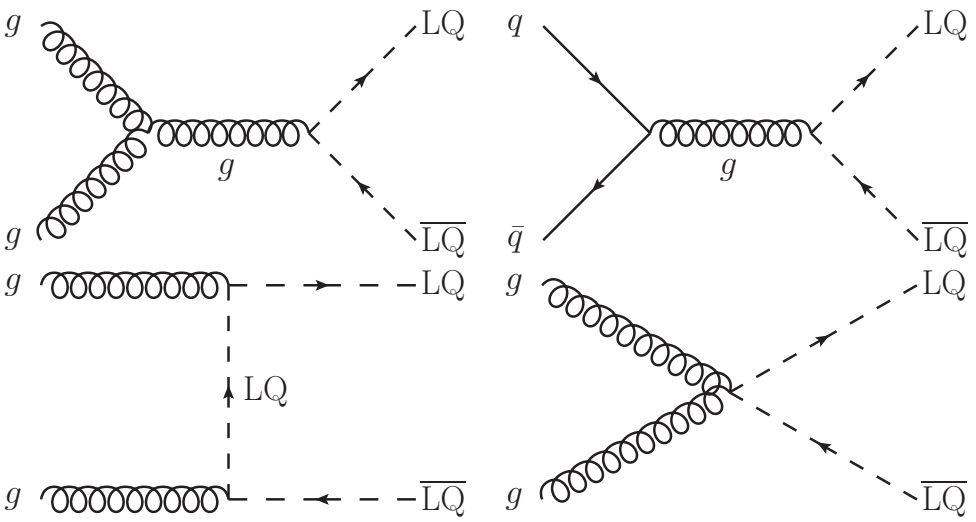
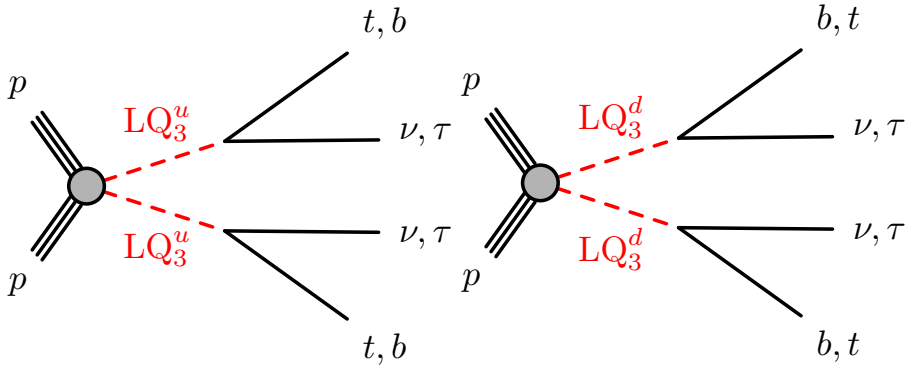


FIGURE 3.8: Feynman diagrams of leptoquark pair-production.



(a) LQ_3^u (b) LQ_3^d

FIGURE 3.9: Feynman diagrams of up- and down-type leptoquark pair-production and subsequent decays.

$\sqrt{\beta}\lambda$ and the coupling to the τ-neutrino is given by $\sqrt{1-\beta}\lambda$. The branching ratio into charged leptons (i.e. $\mathrm{BR}(LQ_3^u \to b\tau)$ and $\mathrm{BR}(LQ_3^d \to t\tau)$) is labelled B,[2] and is shown as a function of the leptoquark mass, $m(LQ_3^{u,d})$, in Figure 3.10.

[2] β and B are not equal for third-generation leptoquarks due to the sizeable top-quark mass.

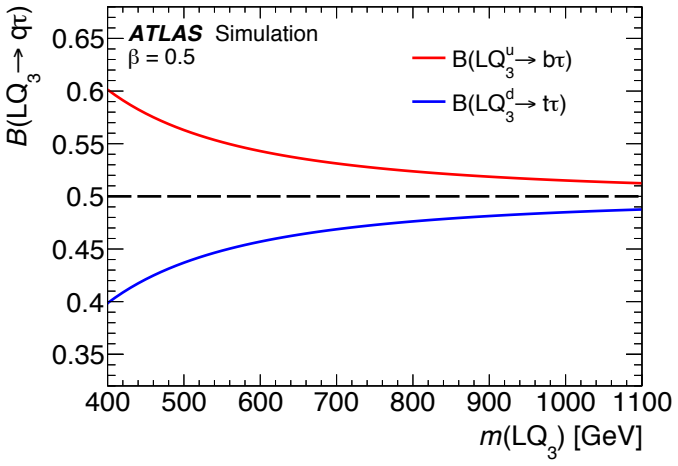


FIGURE 3.10: Branching ratio into charged leptons for third-generation leptoquarks with $\beta = 0.5$ as a function of leptoquark mass, $m(\mathrm{LQ}_3^{u,d})$.

Chapter 4

Machine Learning Theory

Machine learning refers to the process of building a statistical model based on a sample of *training data* which is capable of making decisions and predictions – without being explicitly programmed to do so. Machine learning is used in a number of areas of high-energy physics (HEP): from *regression* problems (where a continuous function is learned), such as estimating the energy of a particle, to *classification* problems, such as identifying the particle type. The most popular machine learning algorithms in HEP are Boosted Decision Trees (BDTs) [77, 78] and Neural Networks (NNs) – though improvements in the training algorithms and computing power have led to a rise in popularity of deep learning techniques [79].

Machine learning techniques are commonly used to search for rare signal processes among large backgrounds. Specifically, in the analyses in this thesis, BDTs are used for the classification of signal and background events. A BDT combines a number of kinematic variables into a single discriminant. This is more successful than placing individual requirements on variables, as per traditional cut-based analysis techniques, because it can keep events that would be incorrectly rejected by a given cut based on whether they pass other selection criteria. Whereas a cut-based analysis selects a single hypercube of phase space, a decision tree splits the phase space into a large number of hypercubes (multi-dimensional 'cubes') and is therefore able to consider the correlations between variables.

The BDT is trained using MC-simulated events, which are labelled as 'signal' or 'background'. To produce an effective, unbiased model, a large MC sample is required and the simulation must provide a good description of the data. The BDT learns the signatures typically associated with signal events and is then applied to the unlabelled, real data, which it categorises as signal or background.

There are two ways in which the presence of a signal can be determined using the BDT discriminant. By requiring events to have a minimum BDT score, the more signal-like events are extracted and the statistical analysis can be performed in the same way as a traditional cut-based analysis. Alternatively, hypothesis testing is performed using the BDT score distribution in the statistical fit – this is the method chosen for the analyses documented here.

The BDTs in both analyses in this thesis were implemented using TMVA, which is a machine learning toolkit designed specifically for HEP applications [80].

4.1 Decision Trees

A single decision tree [77] is a series of binary splits to the input data. The algorithm begins with all events, both signal and background, on the first node (known as the *root* node). It iterates over each input variable to find the variable and corresponding cut value that achieve the best separation between signal and background. From this cut, two new child nodes are produced.

There are several separation criteria that may be used as the figure of merit for selecting the optimum cut at each stage of the algorithm. The BDTs used for the analyses in this thesis use the *gini index*, g:

$$g = p(1-p)\sum_{i=1}^{N} W_i, \tag{4.1}$$

where the i^{th} of N events has weight W_i and p is the purity of a node, given by

$$p = \frac{s}{s+b}, \tag{4.2}$$

where s and b represent the number of signal and background events on the node respectively.

The gini index is symmetric with respect to the event classes; for purity $p = 0$ or $p = 1$ (i.e. a data sample that is 100% signal or 100% background), $g = 0$. Therefore, in order to choose a splitting variable and value, $g_{\text{left node}} + g_{\text{right node}}$ is minimised.

The algorithm is repeated recursively on each node; the next node chosen to split is the one that will give the maximum change in the gini index. A variable may be used

more than once. The decision tree is finalised when the *stopping criteria* are reached. Depending on the configuration, the algorithm may be terminated when:

- the minimum leaf size is reached,
- insufficient improvement is made from further splitting,
- the maximum tree depth is reached, or
- the maximum number of final nodes are produced.

The route to each final node (the final nodes are known as *leaves*) represents a series of cuts which gives a specific region of hyperspace; all events in that hypercube are classified as either signal or background according to which makes up the majority.

4.2 Boosting

Decision trees are known to be unstable with respect to statistical fluctuations in the training sample. For example, if two variables give similar separation, a relatively small fluctuation in the training sample could lead to a different variable being chosen and, therefore, a significantly different tree being grown from that node. This is improved by the use of Random Forests (RFs), where an ensemble of individual decision trees are grown and combined to form a single classifier. Whilst a number of methods exist for growing and combining RFs, the most popular in HEP is *boosting* [78].

Events in the training sample that were misclassified in the original tree have their weights increased (boosted) and a new tree is grown. This procedure is repeated recursively for each new tree such that a *forest* of trees is constructed, with each tree using a different version of the training data. A weighted average of the trees in the forest is taken to produce the final classification.

The analyses in this thesis use the *adaptive boost*, or AdaBoost, method [81, 82]. Consider a training set of N examples, $A = \{(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_N, y_N)\}$, where $\boldsymbol{x}_i$ is the set of input variables for the i^{th} event and y_i is its label (i.e. signal or background); $y_i \in \{-1, 1\}$. The weight for each point is initialised as $W(\boldsymbol{x}_i, y_i) = 1/N$. The result yielded for the i^{th} event by the j^{th} tree in the ensemble is given by $h_j(\boldsymbol{x}_i) = \pm 1$ depending on whether the leaf containing the event is labelled as signal or background.

In order to grow the $(j+1)^{\text{th}}$ tree, the weights of the misclassified events in the previous tree, j, are multiplied by a boost weight:

$$W_i^{(j+1)} = e^{\alpha_j} W_i^{(j)}, \tag{4.3}$$

where α_j is given by

$$\alpha_j = \beta \ln\left(\frac{1-\epsilon_j}{\epsilon_j}\right). \tag{4.4}$$

Here β represents the *learning rate*, which can be optimised for the analysis, and the weighted sum of misclassified events, ϵ_j, is given by

$$\epsilon_j = \frac{\sum\limits_{y_i \neq h_j(\boldsymbol{x}_i)} W_i^{(j)}}{\sum\limits_{i=1}^{N} W_i^{(j)}}. \tag{4.5}$$

where the numerator is the weight of the misclassified events and the denominator is the total weight of the tree.

The boosted classification of the i^{th} event, $h_{\text{boost}}(\boldsymbol{x}_i)$, is then given by

$$h_{\text{boost}}(\boldsymbol{x}_i) = \frac{1}{N_{\text{trees}}} \sum_{j=1}^{N_{\text{trees}}} \alpha_j h_j(\boldsymbol{x}_i), \tag{4.6}$$

where N_{trees} is the number of trees in a forest.

4.3 Training and Validation

Machine learning methods can result in a model that is too closely fit to the training data but does not generalize well to unseen data. This is referred to as overtraining; it occurs when there is insufficient training data to justify the number of parameters in the model.

It is important to validate the model by evaluating the results on a test data set for which the labels (signal or background) are known. A common method is the k-fold validation method in which the training data is split into k equally sized, statistically independent, random sets (known as folds). The classifier is trained using the data from $k-1$ of these folds and tested on the remaining fold, with the process repeated for all possible permutations (of which there are k). The BDTs in Parts III and IV are validated using a two-fold procedure; the MC samples are split into two sets according

BDT Training Parameter	Description	Value
BoostType	Type of boost	AdaBoost
AdaBoostBeta	AdaBoost β value	0.15
NTrees	Number of trees	200
MaxDepth	Maximum depth of tree	4
MinNodeSize	Minimum node size	5%
NCuts	Number of cuts	100

TABLE 4.1: The configuration used for the BDT training.

to their event number (one set for even event numbers, one for odd). Therefore, two classifiers (odd and even) are applied to data; the odd classifier is applied to even event numbers and the even classifier is applied to odd.

When employing a two-fold training strategy, the BDT is overtrained if there is a significant deviation between the training and test distributions for either signal or background, or both, for either of the two classifiers. A Kolmogorov-Smirnov goodness-of-fit test is used to quantify the degree of overtraining by comparing the output distributions from applying the BDT to the training and test MC sets. [1]

The BDT training is performed using a specific set of parameters which are optimised to minimise the expected upper limits whilst monitoring for overtraining. The parameters used for both analyses in this thesis are given in Table 4.1.

Variables in a BDT that provide little improvement to the signal and background separation achieved or have high correlation with another variable should be removed. TMVA derives a ranking of the variables used by the decision trees (calculated for the full decision tree forest) according to how often each variable is used to split decision tree nodes, where each split occurrence is weighted by the *separation gain-squared* it achieves and the number of events in the node [77]. The separation gain achieved when a single parent node is split into two child nodes is given by

$$\text{separation gain} = g_{\text{parent node}} - g_{\text{left node}} - g_{\text{right node}}, \tag{4.7}$$

where the gini index, g, is given in Equation 4.1.

[1] The Kolmogorov-Smirnov test provided by TMVA is susceptible to bias due to being applied to binned data.

Chapter 5

Statistical Interpretation

When searching for physics processes that have been predicted but not observed, statistical methods are employed to determine whether observations are consistent with expectations and, if not, quantify the significance of the deviation.

This is known as *hypothesis testing*; the search for a new signal process can be formalised by defining the *null hypothesis*, H_0, as the known (Standard Model) background processes, while the *alternative hypothesis*, H_1, is defined as the background processes plus the signal for which the search is performed. When setting exclusion limits, these definitions are reversed; H_0 is the signal-plus-background hypothesis and H_1 is the background-only scenario.

If a deviation from the null hypothesis is observed, the statistical significance of the observed signal is quantified by a p-value (or the equivalent Gaussian significance, Z), which is discussed in Section 5.1. This is used to determine whether the signal strength is sufficient to claim a discovery. The sensitivity of an experiment can be characterised by the *expected* significance associated with a given signal hypothesis.

In the absence of a signal, 95% confidence level limits are set on model parameters, such as the cross-section of a particular process or the mass of a particle. An upper (lower) limit excludes values of the given parameter below (above) the quoted limit at the 95% confidence level.

5.1 Likelihood Fit

The level of agreement of the data with a given hypothesis, H, can be quantified in terms of a p-value, i.e. the probability, under assumption of H, that the data will be

of equal or greater incompatibility with the predictions of H. The hypothesis can then be excluded if the observed p-value is below a specified threshold. The p-value is often converted to a significance, Z, defined such that a Gaussian distributed variable found Z standard deviations above its mean has an upper-tail probability equal to p, i.e.

$$Z = \Phi^{-1}(1-p), \tag{5.1}$$

where Φ^{-1} is the quantile (inverse of the cumulative distribution) of the standard Gaussian. In high-energy physics, rejection of the background-only hypothesis typically requires a significance of $Z \geqslant 5$ to constitute a discovery, corresponding to $p = 2.87 \times 10^{-7}$. Exclusion of the signal hypothesis occurs at a threshold p-value of $p = 0.05$ (hence the term 95% *confidence level*), which is equivalent to $Z = 1.64$.

The sensitivity of an experiment can be quantified in terms of the *expected* significance that is obtained for a given signal hypothesis. For example, the sensitivity to discovery of a given signal process, H_1, can be quantified by the expectation value of Z obtained by testing the background-only model, H_0, under the assumption of H_1.

In high-energy physics, a discovery (or exclusion) can be determined from a frequentist significance test using a *likelihood function* [83]. This is performed using a measured histogram with N bins, $\boldsymbol{n} = (n_1, n_2, ..., n_N)$. By writing the expected number of data events in the i^{th} bin as $E[n_i] = \mu s_i + b_i$, where s_i and b_i are the expected signal and background yields in the i^{th} bin, respectively, the *signal strength*, μ is defined. A value of $\mu = 0$ corresponds to the background-only hypothesis and $\mu = 1$ corresponds to the nominal signal-plus-background hypothesis.

A Poisson distribution is assumed for the number of signal and background events in each bin, n_i. The likelihood function, L, is the product of the Poisson probabilities for all bins in the measured histogram, given by

$$L(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))^{n_i}}{n_i!} \exp^{-(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))} \text{ [83].} \tag{5.2}$$

In addition to the parameter of interest, the signal strength, μ, the signal and background models contain a number of *nuisance parameters*, $\boldsymbol{\theta}$. These are used to parametrise the effect of each systematic uncertainty on the expected number of signal and background events in each bin. These are usually modelled by Gaussian functions with width representing the size of the uncertainty. Log-normal functions are used for the background

normalisation parameters, which are required to remain positive. The nuisance parameters are fitted from the data, introducing additional flexibility to the model which reflects the loss of information about μ which is imposed by the existence of uncertainty.

In addition to the measured histogram, $\boldsymbol{n}$, further measurements can be made which constrain the nuisance parameters, $\boldsymbol{\theta}$. A control region, dominated by background events, can be chosen in order to constrain nuisance parameters that are related to background modelling. A histogram in this region can be constructed, where the number of entries in M bins (labelled j) form the set $\boldsymbol{m} = (m_1, m_2, ..., m_M)$. The expectation value of the number of events in the j^{th} bin is written as $E[m_j] = u_j(\boldsymbol{\theta})$. Equation 5.2 becomes:

$$L(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))^{n_i}}{n_i!} \exp^{-(\mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))} \prod_{j=1}^{M} \frac{u_j(\boldsymbol{\theta})^{m_j}}{m_j!} \exp^{-u_j(\boldsymbol{\theta})} \text{ [83]}. \tag{5.3}$$

Maximising the likelihood function yields the most probable values of μ and $\boldsymbol{\theta}$. To test the fitted values, the *profile likelihood ratio* is given by

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}, \tag{5.4}$$

where $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the parameters that maximise the likelihood, referred to as the *unconditional maximum likelihood estimation*, and $\hat{\hat{\boldsymbol{\theta}}}$ are the values of the nuisance parameters that maximise L for the specified μ, referred to as the *conditional* maximum likelihood estimation. When searching for a new physics signal, the discovery test statistic used is

$$q_0 = \begin{cases} -2 \ln \lambda(0), & \text{if } \hat{\mu} \geqslant 0. \\ 0, & \text{if } \hat{\mu} < 0. \end{cases} \tag{5.5}$$

This tests the hypothesis that $\mu = 0$ (the background-only hypothesis); rejecting $\mu = 0$ corresponds to the observation of the signal-plus-background hypothesis. The $\hat{\mu} \geqslant 0$ case is that in which there is an excess of events; $\hat{\mu} < 0$ corresponds to a deficit. By using this test statistic, a discrepancy with the $\mu = 0$ hypothesis is only reported in the case of an excess of events.

In the absence of a signal, a new test statistic is defined in order to establish upper limits on μ:

$$q_\mu = \begin{cases} -2\ln\lambda(\mu), & \text{if } \hat{\mu} \leqslant \mu. \\ 0, & \text{if } \hat{\mu} > \mu. \end{cases} \tag{5.6}$$

For $\hat{\mu} \leqslant \mu$, this statistic is used to test the compatibility of the hypothesised value of μ and the data. For $\hat{\mu} > \mu$, q_μ vanishes because $\hat{\mu}$ cannot be excluded by an upper limit of μ [83]. The level of agreement between the data and the hypothesised value of μ is expressed by the p-value. For an observed value of $q_{\mu,\text{obs}}$, the p-value is given by

$$p_\mu = \int_{q_{\mu,\text{obs}}}^{\infty} f(q_\mu|\mu)\mathrm{d}q_\mu, \tag{5.7}$$

where $f(q_\mu|\mu)$ is the probability density function of q_μ assuming the hypothesis μ.

5.2 The CL$_s$ Method

The analyses in this thesis use the CL$_s$ method [84] to set upper limits at the 95% confidence level.

The p-value calculated using the test statistic in Equation 5.6 is defined as the probability, under assumption of the signal-plus-background hypothesis, to find a value of q_μ with equal or less compatibility with the signal-plus-background hypothesis than is observed from data. Using this prescription, signal hypotheses to which the analysis has little or no sensitivity are excluded with probability slightly greater than the threshold p-value, typically 5%. This exclusion occurs when there is a sufficiently large downward fluctuation in the number of background events.

To avoid excluding models to which there is no sensitivity, the CL$_s$ method requires for exclusion:

$$\text{CL}_s \equiv \frac{p_{s+b}}{1-p_b} < \alpha \tag{5.8}$$

where p_{s+b} is the p-value for the signal-plus-background hypothesis, p_b is the p-value for the background only hypothesis, and α is the threshold for exclusion– typically $\alpha = 0.05$. In the case of low signal sensitivity, $1-p_b$ becomes small and p_{s+b} is penalised, preventing the signal model from being excluded.

Part II

Experimental Setup

Chapter 6

The Large Hadron Collider

The Large Hadron Collider (LHC) [85], built by the European Organisation for Nuclear Research (CERN) between 1998 and 2008, is the world's largest and highest energy particle accelerator and collider. It is contained in a circular tunnel with a circumference of 26.7 km at a depth of 50-175 m below the Franco-Swiss border.

Protons (and sometimes heavy ions, such as lead nuclei) are accelerated in counter-rotating beams to an energy of 6.5 TeV per beam, at which they travel at close to the speed of light. Bunches of up to 1.15×10^{11} protons are collided 40 million times per second at four crossing points, around which are positioned four main detectors.

6.1 The LHC Accelerator Complex

Figure 6.1 shows the succession of machines used to accelerate protons to increasingly high energies before injecting them into the LHC. The protons begin as atoms of hydrogen gas, which are stripped of their electrons in an electric field. Linac 2, the first accelerator in the chain, is a linear accelerator which accelerates the protons to an energy of 50 MeV. They then enter the Proton Synchrotron Booster, which accelerates the protons to 1.4 GeV, followed by the Proton Synchrotron (PS), where they reach 25 GeV. The series of radio frequency cavities in the PS splits the beam into discrete *bunches* of protons with 25 ns spacing. These bunches are then accelerated to 450 GeV in the Super Proton Synchrotron, from which they are finally injected into the two parallel beam pipes of the LHC, which are kept at ultra-high vacuum.

The beams travel in opposite directions around the ring, guided around their circular path by 1232 superconducting dipole magnets, with an additional 392 quadrupole magnets used to keep the beams focused. The beams are accelerated to their maximum

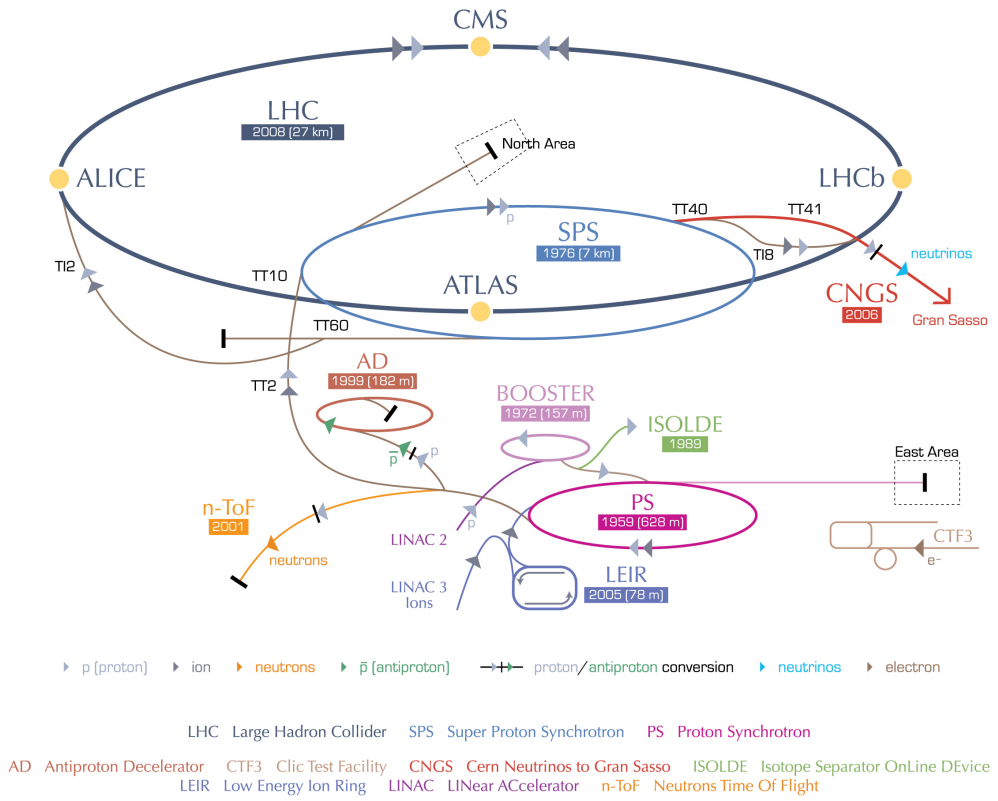


FIGURE 6.1: Schematic of the CERN accelerator complex showing the succession of machines used to prepare protons for injection into the main ring, as well as the locations of the four main detectors. Image taken from [86].

energy of 6.5 TeV (giving a centre-of-mass collision energy of $\sqrt{s} = 13$ TeV), at which they travel at close to the speed of light. The beams cross at four interaction points around the ring, where stronger quadrupole magnets focus the beam further to maximise the chance of interaction. The four main LHC detectors are built around these interaction points: ATLAS [87], CMS [88], LHCb [89] and ALICE [90].

6.1.1 Luminosity

The rate of collisions at the LHC is known as the *instantaneous luminosity*, $\mathcal{L}$, calculated according to

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} \mathcal{F}, \tag{6.1}$$

where N_b is the number of protons per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency of the radio frequency cavities, γ_r is the relativistic gamma factor, ϵ_n is the normalised transverse beam emission at the interaction point, β^* is the

beta function describing the beam envelope at the interaction point and $\mathcal{F}$ is a reduction factor which accounts for the crossing angle of the beams at the interaction point.

The *integrated luminosity*, L, is the integral over time of the instantaneous luminosity and is a measure of the total number of collisions occurring in a given time. It has units of barns, b, where 1 b = 10^{-24} cm^{-2}.

The expected number of events, N, for a given process is the product of the cross-section, σ, of the process and the integrated luminosity, L:

$$N = \sigma L = \sigma \int \mathcal{L} \mathrm{d}t. \tag{6.2}$$

6.2 Pile-Up

In a single bunch crossing, multiple proton-proton interactions give rise to final state particles which traverse the detector. Proton-proton collisions surplus to the collision of interest are collectively referred to as *pile-up*. Pile-up events pollute the reconstruction of the final state of the collision of interest, forming a background which is difficult to model and must be considered in every physics analysis.

The pile-up background originates from five sources [91]:

- *in-time* pile-up, which refers to the additional proton-proton collisions occurring in the same bunch-crossing as the collision of interest, of which up to 80 are expected in Run II;
- *out-of-time* pile-up, which refers to the additional proton-proton collisions occurring in bunch-crossings just before and after the collision of interest, which are a challenge for the detector subsystems that have sensitivity windows longer than 25 ns (the interval between bunch crossings);
- the cavern background, i.e. the gas of neutrons and protons inundating the detector during an LHC run, which give rise to random hits in the muon detector;
- beam halo events, which are sprays of sprays of muons induced by an outlier of the proton beam hitting beam collimators;
- beam gas events, which occur when protons collide with residual gas inside the beam pipe.

The two main sources of pile-up are in-time pile-up and out-of-time pile-up. These present a growing concern as the instantaneous luminosity increases.

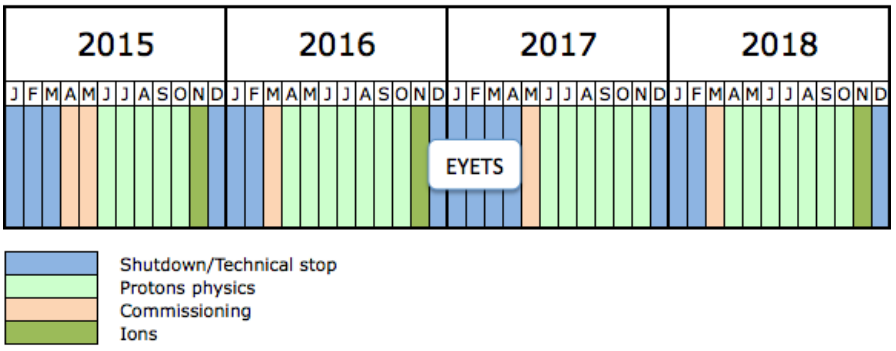


FIGURE 6.2: The running schedule of the LHC for Run II. The majority of the schedule is dedicated to proton-proton collisions, with short heavy ion runs included in 2015, 2016 and 2018. A short shutdown occurs each year, with an Extended Year-End Technical Stop (EYETS) between 2016 and 2017.

6.3 The LHC Run Schedule

The first operational run of the LHC, referred to as Run I, occurred between 2010 and 2013 with the beams colliding at a centre-of-mass energy of $\sqrt{s} = 8$ TeV. During the two-year Long Shutdown between 2013 and 2015, the LHC was upgraded to enable collisions with a centre-of-mass energy of $\sqrt{s} = 13$ TeV for Run II.

Run II of the LHC occurred between 2015 and 2018, following the schedule outlined in Figure 6.2, with the first physics data being delivered in June 2015. In June 2016, the design luminosity for proton-proton collisions was reached and later exceeded, with the peak collision rate reaching up to 140% of the design value [92].

The analyses in Parts III and IV use data collected by the ATLAS experiment in 2015 and 2016, as explained in Section 8.1.

Chapter 7

The ATLAS Detector

At 46 m long and 25 m wide, the ATLAS detector (A Large Toroidal LHC ApparatuS) [87] is the largest of the four experiments on the LHC ring and is situated 93 m underground. It is a general purpose detector designed to provide high-quality measurements of the particles resulting from proton-proton collisions to study a wide range of physics predicted by the SM and beyond.

The detector comprises numerous sub-detector components arranged as a series of concentric cylinders surrounding the beam pipe, each designed with a specific purpose such as tracking particle trajectories, measuring their energies, or detecting a particular type of particle. The high interaction rates, particle energies and radiation dose pose a series of challenges for the detector design, making ATLAS one of the largest and most complex scientific instruments ever constructed.

7.1 The ATLAS Co-Ordinate System

ATLAS uses a right-handed co-ordinate system to describe the detector and the trajectories of particles emerging from proton-proton collisions. The origin is at the interaction point and the z-axis lies along the beam line. The x-y plane is transverse to the beam axis, with the positive x-axis pointing toward the centre of the ring and the positive y-axis pointing upward. The azimuthal angle, ϕ, is measured around the beam axis and the radial co-ordinate, R, measures the distance from the beam line. The polar angle, θ, is the angle from the positive direction of the beam axis. The *pseudorapidity*, η, is a function of θ:

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right), \tag{7.1}$$

hence η is zero when θ is perpendicular to the beam axis. The distance in η and ϕ space is given by

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{7.2}$$

Kinematic variables which take the form of a vector are often expressed in terms of their transverse component, perpendicular to the beam axis. For example, the transverse momentum, p_{T}, is given by

$$p_{\mathrm{T}} = |\boldsymbol{p}| \sin\theta. \tag{7.3}$$

where $\boldsymbol{p}$ is the three-momentum of the particle.

7.2 Detector Overview

Figure 7.1 depicts the numerous sub-detectors which make up the four major sub-systems of the ATLAS detector: the Inner Detector (ID), the Electromagnetic Calorimeter (ECAL), the Hadronic Calorimeters and the Muon Spectrometer (MS).

The detector components are arranged as a series of concentric cylinders surrounding the beam pipe with forward-backward symmetry about the interaction point, full coverage in azimuthal angle, ϕ, and large acceptance in pseudorapidity, η. The detector can be divided into three regions: a central barrel region and two endcap regions.

The ID is used for tracking charged particles and is described in Section 7.4. The ECAL and the Hadronic Calorimeters, described in Sections 7.5.1 and 7.5.2 respectively, are responsible for measuring the energies of electromagnetic and hadronic particles respectively. The MS, described in Section 7.6, forms the outermost part of the detector and is designed to precisely measure the momentum of muons leaving the calorimeters. A unique magnet system provides the magnetic field required for momentum measurements and is described in Section 7.3.

Table 7.1 lists the ATLAS subdetectors and their intended resolutions and coverage [87].

7.3 Magnet System

The ATLAS detector uses a unique hybrid system of four large superconducting magnets [94] to bend charged particles to facilitate the measurement of their momenta. The

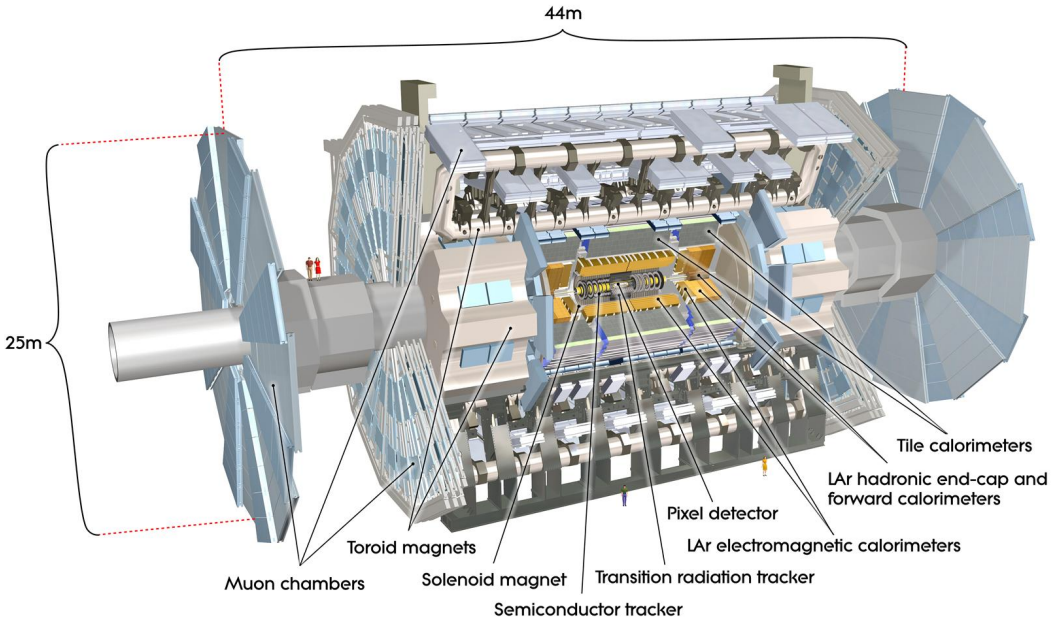


FIGURE 7.1: Schematic depicting the numerous sub-detectors which make up the ATLAS detector. Figure taken from [93].

Subdetector	Resolution	Coverage in η
ID	$\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 0.05\%\ p_\mathrm{T} \oplus 1\%$	± 2.5
ECAL	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2
HCAL (barrel and endcap)	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2
HCAL (forward)	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$
MS	$\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 10\%$ at $p_\mathrm{T} = 1$ TeV	± 2.7

TABLE 7.1: The subdetectors of the ATLAS experiment and their design resolutions and coverage. 'HCAL' refers to the Hadronic Calorimeters.

momentum of a charged particle is directly proportional to the curvature of its trajectory through a magnetic field. The four magnets generate a magnetic field over a volume of 12000 m^3 (defined as the region in which the field exceeds 50 mT). Their configuration is shown in 7.2.

The inner solenoid magnet [95] immerses the ID in a 2.0 T axial magnetic field which runs parallel to the beam axis. A field of this strength is required to bend the most energetic of particles sufficiently for their momenta to be determined. At around 5.3 m in length, 2.4 m in diameter and 4.5 cm thick, the solenoid is a single-layer coil wound with 9 km of aluminium-stabilised superconducting wire. Because the magnet is situated in front of the calorimeters, it is imperative that the radiative thickness of the magnet (0.66 radiation lengths) is minimised. It operates at a nominal current of 7.73 kA and

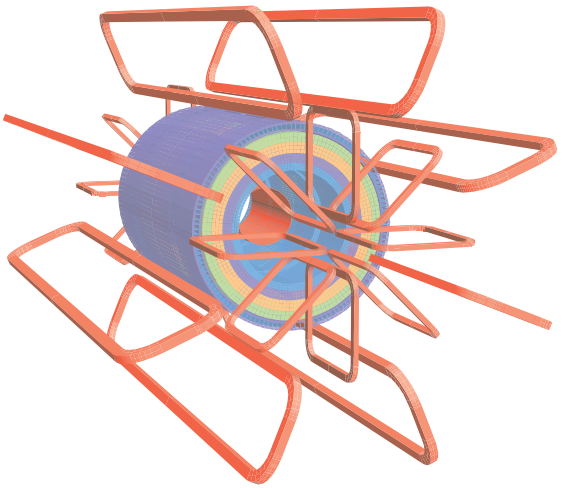


FIGURE 7.2: Diagram of the ATLAS detector magnet system. The barrel toroid magnets, shown in red, surround the smaller endcap toroid magnets, also shown in red, and the inner solenoid magnet. Figure taken from [87].

produces a field with a stored energy of 38 MJ.

Three toroid magnets, one in the barrel region [96] and two smaller endcap magnets [97], provide a toroidal field for the MS. The barrel toroid is 25.3 m in length with a 20.1 m outer diameter, providing a 0.5 T magnetic field storing a total energy of 1.08 GJ. It is constructed from eight separate coils surrounding the calorimeters and the endcap toroids and covers $|\eta| < 1.4$. The endcap toroid magnets, each with eight coils, cover the range $1.6 < |\eta| < 2.7$ and have a field strength of 1.0 T. The field in the transition region, $1.4 < |\eta| < 1.6$, is the result of contributions from the barrel and endcap magnets. This magnet configuration provides a field that is mostly orthogonal to the muon trajectories.

7.4 The Inner Detector

The Inner Detector (ID) [98, 99] is the closest sub-detector system to the beam line, designed to track the early trajectories of charged particles for momentum calculations and locate their primary and secondary vertices with exceptionally high resolution. It is required to deal with a dense environment of tracks - of the order of 1000 particles per collision - arising every 25 ns.

The acceptance in pseudorapidity of the ID is $|\eta| < 2.5$ with full coverage in azimuthal angle, ϕ. The ID comprises four complementary sub-detectors: the Insertable B-Layer (IBL), the Pixel Detector (PD), the Semi-Conductor Tracker (SCT) and the Transition Radiation Tracker (TRT). Their layout is shown in Figure 7.3. The entire ID is immersed in a homogeneous 2 T magnetic field generated by the central solenoid, which is described in Section 7.3.

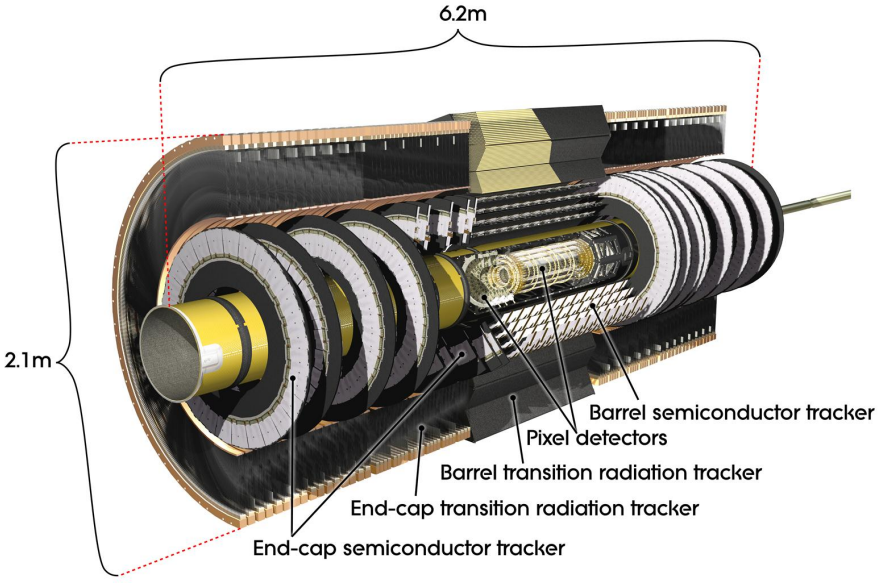


FIGURE 7.3: Schematic depicting a longitudinal view of the ATLAS inner detector, which is the closest sub-detector system to the beam line. Figure taken from [99].

Subdetector	Element size [µm]	Resolution [µm]	Radius of layers in barrel [mm]	Readout channels
IBL	50×250	8×40	33.2	12×10^6
PD	50×400	10×115	50.5, 88.5, 122.5	80×10^6
SCT	80	17	299, 371, 443, 514	6×10^6
TRT	4000	130	554-1082	3.5×10^5

TABLE 7.2: The subsystems which make up the ATLAS ID and their primary characteristics. For the IBL and PD, the sensor element size and intrinsic resolution is given in terms of (R-ϕ,z). For the SCT, the element size quoted is the spacing of the readout strips, while for the TRT, the element size refers to the straw tube diameter. The resolution for the SCT and TRT is reported in terms of (R-ϕ).

Table 7.2 summarises the components which make up the ID [100].

The innermost layers of the ID (the IBL, PD and SCT) utilise high granularity *silicon semiconductor* sensors - in the form of pixels and strip sensors - to provide remarkably precise measurements for tracking, vertexing and b-tagging. Discrete spatial measurements are made when a charged particle deposits energy in the detector, generating electron-hole pairs which are accelerated by an electric field to the nearest electrodes. The TRT relies on *proportional drift tubes*, in which charged particles ionise the gas contained in the tube, causing electrons to drift toward the positively charged wire (anode) at the centre of the tube. The position of the charged particle relative to the wire can be calculated from the speed at which the electrons drift in the gas.

7.4.1 The Pixel Detector

The silicon Pixel Detector (PD) [101], which spans radial distances between 50.5 mm and 150 mm, utilises silicon semiconductor sensors to provide precise measurements for tracking, vertexing and b-tagging. As the innermost layer of the detector (until the addition of the Insertable B-Layer, described in Section 7.4.1.1), its design is subject to strict requirements relating to resolution and radiation hardness.

The PD comprises 1744 identical silicon pixel modules arranged in three concentric barrel layers and two endcaps of three disks each, such that each particle track generates three measurements. Each module contains 46080 pixels, giving a total of over 80 million readout channels in the entire PD, each with a typical size of 50×400 µm^2. The longer dimension is oriented in the z-direction in the barrel and along the radial R-direction in the endcaps. Each pixel is made from 250 µm thick n-type silicon.

7.4.1.1 The Insertable B-Layer

During Run I, the PD was the innermost sub-detector of the ATLAS detector. For Run II, the Insertable B-Layer (IBL) [102] was commissioned. The main motivation of the IBL was to improve the pixel detector performance for Run II despite possible radiation damage to the first layer of the PD and the increasing bandwidth requirements associated with the increasing LHC luminosity. During the first Long Shutdown, the IBL was inserted between the new, smaller-radius beryllium beam pipe and the existing first layer of the PD.

The IBL is positioned at a radius of 33.2 mm from the beam axis and consists of around 12 million pixels with dimensions 50×250 µm^{-2}. These are arranged such that the IBL gives full coverage in azimuthal angle, ϕ, and covers pseudorapidties of $|\eta| < 2.9$. Studies [102] show that the IBL is robust against pile-up and improves the quality of the impact parameter reconstruction and vertexing and, therefore, the b-tagging performance.

7.4.2 The Semiconductor Tracker

The Semiconductor Tracker (SCT) [103], which surrounds the PD and covers the radial region between 299 mm and 560 mm, is a silicon microstrip detector with over six million readout channels.

The SCT is constructed from a barrel region comprising 2112 modules and two endcaps of 988 modules each. In the barrel region, each module is made up of two pairs of sensors, positioned back-to-back and separated by a stereo angle of ± 20 mrad around

the geometrical centre of the sensors. This allows the position of the particle in the z-direction (along the beam pipe) to be constructed. These modules are arranged in four coaxial concentric cylindrical layers around the beam axis. There are three different designs of modules that make up the endcaps, each having two back-to-back sensors, again with a relative rotation of ± 20 mrad. These are arranged in nine disk layers for each of the two endcaps.

The sensors are 285 µm thick and are constructed of high-resistivity n-type bulk silicon with p-type implants. Readout strips are positioned every 80 µm, providing a space-point resolution of about 17 µm in the cross-strip direction and 580 µm longitudinally. For particles originating in the beam interaction region, the SCT aims to provide four space-points (i.e. eight strip measurements).

7.4.3 The Transition Radiation Tracker

The sensitive volume of the Transition Radiation Tracker (TRT) [104], the outermost sub-detector of the ID, covers radial distances from 563 mm to 1066 mm and has a volume of approximately 12 m^3. The detector comprises approximately 3×10^5 proportional drift tubes (referred to as straws) of 4 mm diameter arranged, like the other elements of the ID, in a barrel region and two endcaps. The barrel region comprises approximately 5×10^4 straws, each 144 cm long running parallel to the beam line, with readout at each end of each straw. The endcap straws are 37 cm long and are arranged radially around the beam line.

The polyimide straw tubes act as cathodes and are kept at high negative voltage; a 30 µm diameter gold-plated tungsten wire anode runs through the centre of each tube. The layers of straws are interleaved with polypropylene fibres (in the barrel region) and foils (in the endcaps). When an ultra-relativistic charged particle passes through this material, it emits X-ray transition radiation. This gives high-threshold hits which are used to distinguish the electron from other charged particle tracks. The straws are filled with a gas mixture of 70% xenon (for good X-ray absorption), with CO_2 (27%) and O_2 (3%) to increase the electron drift velocity and for photon-quenching.

7.5 The Calorimeters

Calorimeters are used to measure the energy of both charged and neutral particles as they pass through the detector. ATLAS utilises *sampling* calorimetry, which means that the calorimeters are constructed from alternating layers of a passive absorber, which

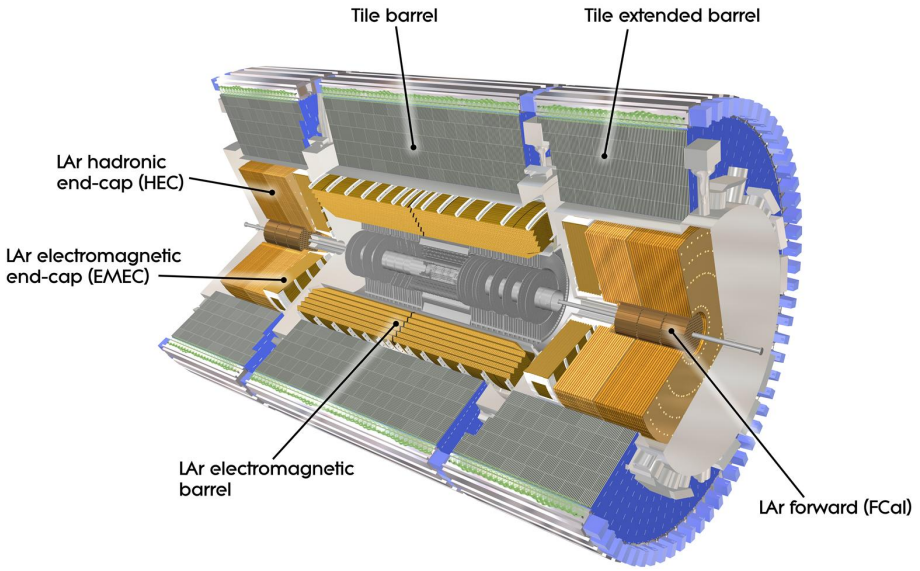


FIGURE 7.4: Schematic depicting the ATLAS Electromagnetic and Hadronic Calorimeters, which surround the ID. Figure taken from [105].

produces the particle showers, and active detector layers which measure the particle energy. As such, a fraction of the total particle energy is deposited in the passive material and is not measured; the overall energy must be deduced from the definite measurements taken in the active detector layers.

ATLAS has two distinct calorimeter subsystems [105], as illustrated in Figure 7.4: the Electromagnetic Calorimeter (ECAL), which measures electrons, photons and the electromagnetically interacting components of jets; and the Hadronic Calorimeters, which measure the strongly-interacting components of jets. These are outlined in Sections 7.5.1 and 7.5.2 respectively.

The calorimeters are each constructed from a barrel region and two endcaps. The calorimeters both cover $|\eta| < 4.9$, with complete coverage in ϕ; high spatial coverage is necessary to provide a reliable calculation of the missing transverse momentum, which is crucial in many physics analyses (including those presented in this thesis). The radial dimension of both calorimeters is such that the electromagnetic and hadronic showers are suitably contained, with minimal punch-through into the surrounding muon system. The total depth of the calorimeters corresponds to around 11 interaction lengths, where one interaction length is the average distance required for the energy of a particle to reduce by a factor of $1/e$ via hadronic interactions. The depth of the barrel region of the ECAL ranges from 22 to 33 radiation lengths, where one radiation length is defined as the average distance required for an electromagnetic particle to lose all but $1/e$ of its energy via bremsstrahlung.

7.5.1 The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) [106] is situated outside of the solenoid magnet that surrounds the ID. The barrel region, covering $|\eta| < 1.475$, is constructed from two identical sections separated by a 4 mm gap at $z = 0$. The two endcaps, covering $1.375 < |\eta| < 3.2$, are constructed from two coaxial wheels. In the transition region between the barrel and endcaps within $1.375 < |\eta| < 1.52$, referred to as the "crack" region, a significant amount of inactive material exists to provide necessary services (such as cooling) to the ID. This region is excluded in many physics analyses that require high-precision measurements of electrons, τ-leptons or photons - including those in this thesis.

The ECAL is constructed from lead absorbers (lead is chosen for its short radiation length) interspaced with liquid Argon (LAr) and kapton electrodes assembled in an accordion shape. The complex geometry ensures complete coverage and symmetry in azimuthal angle, ϕ. The high granularity of the ECAL provides the high energy resolution required and contributes to the identification of jets, photons and leptons due to the shapes of their showers.

7.5.2 The Hadronic Calorimeter

The Hadronic Calorimeters, which surround the ECAL, consist of the tile calorimeter (TileCal) [107], LAr hadronic endcap calorimeter (HEC) and the LAr forward calorimeter (FCal) [106].

The TileCal consists of three barrels: the innermost barrel, which provides coverage up to $|\eta| < 1.0$, and two extended barrels covering $0.8 < |\eta| < 1.7$. Each detector cylinder is divided into 64 independent wedges along the azimuthal direction; these consist of steel absorbers with scintillating tiles (of which there are 5×10^5) as the active material. Scintillating material absorbs the energy of incident particles, emitting photons which are read out by photomultiplier tubes.

To complement the tile calorimeter, the HEC provides hadronic calorimetry over the region $1.5 < |\eta| < 3.2$. The HEC comprises two independent wheels built from 32 wedge-shaped modules, and uses LAr as the active material (due to the high radiation in this region) with copper absorbers.

The hadronic calorimetry is extended to larger pseudorapidities by the FCal, which uses LAr as the active material and covers the region $3.1 < |\eta| < 4.9$. As the FCal modules are located at high η, at a distance of approximately 4.7 m from the interaction point,

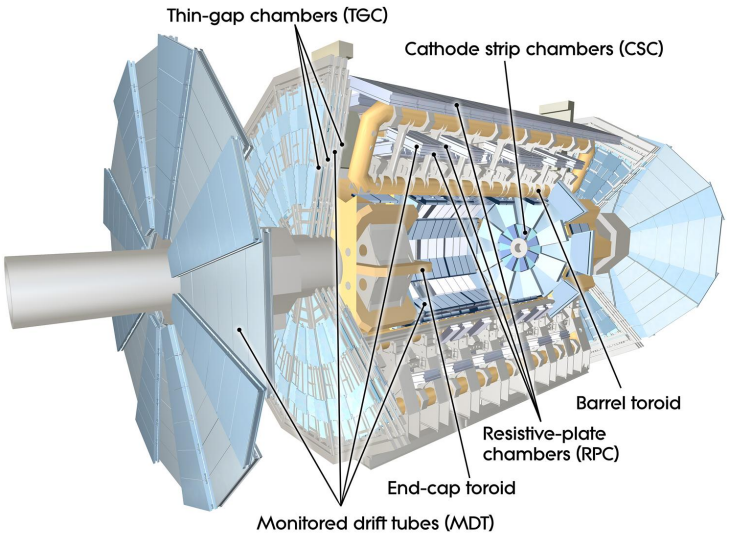


FIGURE 7.5: The ATLAS Muon Spectrometer. Figure taken from [87].

they are exposed to a high particle flux. Therefore, they are designed with particularly small LAr gaps to avoid ion build-up and provide high detector density.

The FCal comprises three cylindrical modules consisting of tubes orientated parallel to the beam pipe. The innermost module uses radiation-hard hollow tubes for electromagnetic calorimetry, whilst the subsequent two modules use tungsten for predominantly hadronic measurements. In total, the LAr calorimeters comprise 1.1×10^5 readout channels.

7.6 The Muon Spectrometer

The ATLAS detector's outermost and largest subdetector is the high-resolution Muon Spectrometer (MS), which is designed to identify and measure the momenta of muons leaving the calorimeters in the range $|\eta| < 2.7$. A diagram of the MS is shown in Figure 7.5. Muons are the only detectable particles that are expected to escape the Hadronic Calorimeter system. Momentum measurements in the MS are facilitated by the magnetic deflection of muon tracks in three large superconducting toroid magnets, described in Section 7.3.

The barrel region is constructed from chambers arranged in three cylindrical layers around the beam axis, whereas the chambers in the endcaps are installed vertically. Over most of the pseudorapidity range, precision tracking is provided by Monitored Drift Tubes (MDTs), with the exception of the forward region, where Cathode Strip Chambers (CSCs) are used for the first layer due to their ability to sustain a higher rate.

The individual MDTs are cylindrical aluminium drift tubes of 30 mm diameter with a central tungsten-rhenium wire to collect ionisation electrons. The average tracking resolution for a single tube is approximately 80 µm in the z-direction; this corresponds to a value of around 35 µm per chamber. There are a total of 1171 MDT chambers, corresponding to more than 3.5×10^5 tubes. The CSCs are multiwire proportional chambers which provide measurements with a resolution of 60 µm in the radial R-direction; there are 7×10^4 readout channels.

The muon trigger system, positioned around the middle and outer layers of the MDT chambers, covers $|\eta| < 2.4$. Resistive Plate Chambers (RPCs) are used in the barrel region, $|\eta| < 1.05$, and Thin Gap Chambers (TGCs) are used in the forward (endcap) region, $1.05 < |\eta| < 2.4$. The RPCs are gaseous parallel electrode-plate detectors which are positioned around the middle and outer layers of the MDTs. There are 3.8×10^5 RPC readout channels. The TGCs are arranged in four layers, three positioned around the second MDT wheel and the fourth in front of the CSC layer, with a total of 4.4×10^5 channels. The muon trigger system is discussed further in Section 7.7.

7.7 The Trigger and Data Acquisition System

The physics processes that the ATLAS experiment aims to detect typically have very low cross-sections. Therefore, to produce sufficient signal events to generate a statistically significant result, a high luminosity is required. With around 40 million bunch crossings per second and up to 80 collisions per bunch crossing, the data collection bandwidth and storage capacity of ATLAS is insufficient to record every event that occurs. The ATLAS trigger system plays the crucial role of selecting high-quality, rare events with interesting signatures which may hint at new physics. The trigger system operates in two stages: a hardware component known as the Level 1 (L1) trigger [108], followed by the software-based High Level Trigger (HLT) [109].

The L1 trigger reduces the event rate to a maximum of 100 kHz using a Central Trigger Processor (CTP), which operates on signals from dedicated hardware in the calorimeter and muon detector systems. The decision time, at under 2.5 µs, is faster than the ID can process events so ID information is omitted. For each data-taking period, the L1 trigger is loaded with a *trigger menu*, a list of up to 256 criteria used to determine whether an event is accepted. The trigger menus are designed to accommodate a broad physics programme, with high acceptance for both BSM searches and SM precision measurements.

The L1 trigger also uses detector information with reduced granularity to identify *Regions of Interest* (RoI) [110] in η and ϕ, which are used as the starting point for the HLT. The accepted events and RoI information are passed to the HLT, which considers the RoIs with full granularity information from the calorimeters and the MS as well as tracking information from the ID. The HLT processes the events using algorithms for particle reconstruction, calibration and identification which are as close as possible to those applied offline. This stage of the trigger reduces the event rate to $\sim$ 1 kHz with an average latency of $\sim$ 0.2 s. These events are stored and processed for later analysis.

7.8 Detector Simulation

Monte Carlo generators, as described in Section 8.2, simulate the proton-proton collisions which occur inside the ATLAS detector. A detector simulation mimics the passage of the produced particles through the experimental apparatus such that the simulated hypothesis can be directly compared with the real data collected by ATLAS.

The software models the trajectories of particles and emulates showering in the calorimeters to compute the energy deposits throughout the detector. This is converted into the same digital data format as that collected from the electric current and voltage signals from the detector. This ensures that the reconstruction chain applied to real data (see Chapter 9) can be applied to simulation.

Physics analyses, such as those in Parts III and IV, require huge samples of events. The total time required for event simulation is dominated by the detector simulation, of which two types are available: *full simulation* and *fast simulation*. The full simulation is used for most samples listed in Chapter 8 and is provided by the GEANT4 particle simulation toolkit [111, 112], which uses a highly detailed description of the detector geometry and physics processes. Almost 80% of the time taken for a full simulation is required to simulate the showering of particles as they traverse the calorimeters. The fast simulation provided by ATLFAST-II [113] reduces the simulation time by more than one order of magnitude by modelling the calorimetry using FASTCALOSIM [114], which is intrinsically less accurate but can be tuned using data. ATLFAST-II is used to generate the signal MC samples required in Part IV.

Chapter 8

Data and Monte Carlo Samples

The analyses described in this thesis use data collected by the ATLAS experiment in 2015 and 2016, as detailed in Section 8.1.

In order to determine whether a signal process is present in proton-proton collision data, precise modelling of all signal and background processes is required such that a comparison can be made. Modelling of physics processes requires several stages of computation to model the hard interaction, parton showering and hadronisation, initial- and final-state radiation and pile-up.

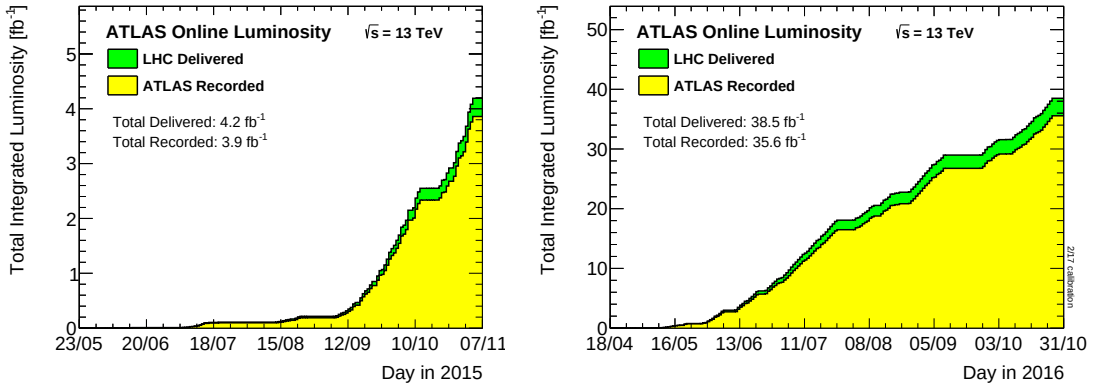
The generators used for the signal and background processes are outlined in Sections 8.3 and 8.4 respectively.

8.1 Data

The data used in both analyses are measurements of proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, collected by the ATLAS detector at the LHC during the 2015 and 2016 running periods. Selected data events are required to have been measured with all relevant components of the ATLAS detector in good working condition; the number of events collected corresponds to an integrated luminosity of 3.2 fb^{-1} in 2015 and 32.9 fb^{-1} in 2016, summing to a total of 36.1 fb^{-1}.

8.2 Simulation of Physics Processes

The simulation of signal and background processes is performed using Monte Carlo (MC) methods over a series of stages of computation.



(a) 2015 (b) 2016

FIGURE 8.1: Cumulative integrated luminosity delivered by the LHC (green) and recorded by the ATLAS experiment (yellow) in (a) 2015 and (b) 2016. The total integrated luminosity used for analysis is slightly less than the total recorded as events are required to have been measured with all relevant components of the ATLAS detector in good working condition.

The hard interaction occurs between partons of two colliding protons; the partons include *valence quarks*, gluons, which mediate the strong interactions between the valence quarks, and *sea quarks*, which are virtual quark-antiquark pairs. The distribution of the total hadron momentum among the partons is described by *parton distribution functions* (PDFs). Specifically, a PDF describes the probability of finding a given constituent parton carrying a given fraction of the proton's total longitudinal momentum. As perturbative QCD cannot be used to calculate the PDFs, they are obtained by fitting observables to experimental data.

The hard interaction itself is simulated by calculating the *matrix element*, $\mathcal{M}$, using Feynman diagrams at either leading order (LO) or next-to-leading order (NLO) in perturbation theory. For increased accuracy, the process can be normalised to next-to-next-to-leading order (NNLO) if required. The intermediate particles are decayed according to their branching ratios.

As well as calculating the hard interaction, it is important to see how the resulting partons evolve. Coloured partons radiate further partons in a process called *showering*. This can be modelled by perturbative QCD until the energy scale of confinement, where the theory breaks down, but given the complex calculations associated with high-multiplicity events, purpose-built parton shower algorithms are employed. The full parton topologies are described by a combination of the matrix element generator and parton shower algorithms. These must be *matched* and *merged* to generate the whole phase space, avoiding double counting. At low energies, these partons combine to form colourless hadrons in a process called *hadronisation*. As with parton showers, hadronisation is non-perturbative and requires phenomenological modelling using specific hadronisation models.

Particles involved in the hard scattering can emit initial- and final-state radiation in the form of gluons and photons. These particles, as well as those originating from other soft scattering interactions, are described as the *underlying event*. All non-primary interactions are collectively referred to as pile-up, which is described in Section 6.2; these are simulated separately and later overlaid on the hard scatter event.

Finally, event generators come with a set of parameters which are *tuned* to match data. The tuned parameters are used to simulate hadronisation and the underlying event, which are based on phenomenological models.

8.3 Signal Monte Carlo Samples

8.3.1 Resonant and Non-Resonant Higgs Boson Pair-Production

The SM non-resonant Higgs pair-production process is simulated, assuming a Higgs boson mass of $m_h = 125.09$ GeV, using an effective field theory (EFT) model. The model is implemented in the MadGraph5_aMC@NLO v2.2.2 generator [115] at next-to-leading order (NLO) using the CT10 parton distribution function (PDF) set [116]. The parton showers and hadronisation are simulated using Herwig++ [117] using the UEEE5 set of tuned parameters (tune) [118]. In the EFT model, computations are performed in the limit $m_t \to \infty$ (where m_t is the mass of the top-quark). The generated events are reweighted to reproduce the m_{hh} spectrum obtained in [37, 119], which fully accounts for the finite mass of the top quark.

The resonant samples are all generated using MadGraph5_aMC@NLO at leading order (LO) and interfaced to the Pythia 8 [120] parton shower model using the A14 tune [121] and the NNPDF23LO PDF set [122]. For both the generic narrow width scalar, H, and the bulk RS graviton, G, an SM Higgs boson mass of $m_h = 125$ GeV is assumed. The width of the heavy scalar is set to $\Gamma = 4$ MeV. The cross-section and width of the RS graviton are taken from [123]. RS graviton signal events are generated with different values of the coupling constant, $k/M_{\text{Pl}} = 1.0, 2.0$; for an RS graviton with $m_X = 1$ TeV and $k/M_{\text{Pl}} = 1.0$, the resonance width is $\Gamma = 55$ GeV.

There are 14 samples with masses $m_X = 260$, 275, 300, 325, 350, 375, 400, 450, 500, 550, 600, 700, 800, 900, 1000 GeV used for the scalar and the RS graviton with $k/M_{\text{Pl}} = 2.0$; a subset of 9 samples with $m_G = 260$, 300, 400, 500, 600, 700, 800, 900, 1000 GeV are used for the RS graviton with $k/M_{\text{Pl}} = 1.0$.

8.3.2 Third-Generation Scalar Leptoquark Pair-Production

The leptoquark signal samples are generated at next-to-leading order (NLO) in QCD with MadGraph5_aMC@NLO v2.4.3 [115], using the leptoquark model described in [124] and the NNPDF30NLO [125] parton distribution functions (PDFs). The parton showering and hadronisation is performed with Pythia 8 using the A14 set of tuned parameters [121].

The leptoquark production cross-sections are taken from calculations of the direct top-squark pair-production [126]; top-squarks and third-generation leptoquarks have the same production modes and both carry mass, colour and spin-0. Uncertainties on the calculations are estimated from factorisation and renormalisation scales, α_s, and PDF variations.

The decay of the leptoquarks is simulated using Madspin [127]. The coupling parameter λ (defined in Section 3.5) is set to $\lambda = 0.3$, giving a width of about 0.2% of the leptoquark mass. Samples are produced for up- and down-type leptoquarks with model parameter $\beta = 0.5$ and for up-type leptoquarks only with $\beta = 1$.

For up-type leptoquarks, there are 21 samples with $\beta = 1$ and masses $m(\mathrm{LQ}_3^u) = 200$, 225, 250, 275, 300, 325, 350, 400, 450, 500, 550, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500 GeV. These are used to produce limits with the assumption that 100% of up-type leptoquarks decay to $b\tau$ pairs. There are a further eight samples each for up- down-type leptoquarks with $\beta = 0.5$ and masses $m(\mathrm{LQ}_3^d) = 400$, 500, 600, 700, 800, 900, 1000, 1100 GeV. These are reweighted to produce samples with varying branching ratio, B, where the applied weight, $w(B)$, is given by

$$w(B) = \left(\frac{B}{\hat{B}}\right)^n \times \left(\frac{1-B}{1-\hat{B}}\right)^{(2-n)}, \tag{8.1}$$

where $\hat{B}$ is the branching ratio of the original MC sample, and n is the number of charged leptons in the event which originate directly from the decaying leptoquarks.

8.4 Background Monte Carlo Samples

The same background processes apply to both searches presented in this thesis; the same MC samples are used for both.

For the generation of $t\bar{t}$ and single top-quarks produced via the Wt- and s-channels, the Powheg-Box v2 generator [128] is used with the CT10 PDF set. For single top-quarks

produced via the electroweak t-channel, the POWHEG-BOX v1 generator is used with the CT10f4 PDF set [116], and the top-quarks are decayed using MADSPIN [127]. The parton shower, fragmentation, and the underlying event are simulated using PYTHIA 6.428 [129] with the CTEQ6L1 PDF sets [130] and the Perugia 2012 tune [131]. For the $t\bar{t}$ production, the cross-section is generated at NNLO+NNLL [132], while for single-top processes the samples are corrected to NLO [133, 134]. For all processes, a value of $m_t = 172.5$ GeV is assumed for the top-quark mass.

Events containing W or Z bosons with associated jets are simulated using the SHERPA 2.2.1 generator [135]. Matrix elements are calculated using the COMIX [136] and OpenLoops [137] matrix element generators; this is then merged with the SHERPA parton shower [138] using the ME+PS@NLO prescription [139]. The NNPDF30NNLO PDF set [125] is used and events are normalized to the NNLO cross-sections [140]. Di-boson and Drell–Yan backgrounds are produced with Sherpa 2.1.1 using the CT10 PDF set and the generator NLO cross-section predictions.

Quark-induced Zh processes, where an SM Higgs is produced in association with a Z boson, are generated with PYTHIA 8.186, using the A14 tune and the NNPDF23LO PDF set, for both decay channels: $qq \to Zh(Z \to \tau\tau, h \to bb)$ and $qq \to Zh(Z \to bb, h \to \tau\tau)$. The gluon-induced Zh process, $gg \to Zh(Z \to \tau\tau, h \to bb)$ [141] is generated with POWHEG-BOX v2 using the CT10 PDF set and and the subsequent parton showering is simulated using PYTHIA 8.186 with the AZNLO tune [142]. The other gluon-induced Zh process, $gg \to Zh(Z \to bb, h \to \tau\tau)$, is difficult to model and is accounted for by scaling up its quark-induced counterpart by 6%. The Zh production cross-sections [143–147] are scaled to NLO+NLL in QCD.

SM Higgs boson production in association with a pair of top-quarks is simulated using MADGRAPH5_AMC@NLO interfaced with PYTHIA 8 to simulate parton showering, with all decays produced inclusively. The cross-section is taken from [28].

For all background processes involving an SM Higgs boson, the Higgs mass is set to $m_h = 125$ GeV. EVTGEN v1.2.0 [148] is used to model the decays of charm and bottom hadron decays for all processes except those simulated with SHERPA. Simulated events are reweighted to match the distribution of the number of inelastic collisions per event (pile-up) in data. All MC samples are passed through the full simulation of the ATLAS detector produced in GEANT4 apart from the leptoquark signal samples, which use ATLFAST-II. All Monte Carlo events are reconstructed using the same software as used for data.

Chapter 9

Object Reconstruction and Identification

Particles traversing the ATLAS detector deposit energy in the layers of sub-detectors, giving rise to electronic signals in any number of the millions of electronic channels in the detector. These are converted, via the complex reconstruction process, into a series of measurements which form physics objects associated with the original particles. Further, not all objects built by the reconstruction algorithms are signal objects. ATLAS employs identification algorithms tailored for each object type to reject background objects with varying efficiency.

This chapter describes the reconstruction and identification, implemented using standardised ATLAS software, of all physics objects relevant to the analyses in this thesis.

9.1 Electrons

9.1.1 Electron Reconstruction

Electrons in the central region of the detector ($|\eta| < 2.47$) are reconstructed from tracks in the ID (see Section 7.4) which are matched with energy deposits in the ECAL (Section 7.5.1).

The first step of the electron reconstruction process involves the reconstruction of a cluster of energy deposits in the ECAL using the *sliding window* algorithm [149]. This method is also relevant for photon and τ-lepton (Section 9.5) reconstruction. The first step involved in sliding window clustering is referred to as *tower building*. The η-ϕ space of the ECAL is divided into $N_\eta \times N_\phi = 200 \times 256$ elements of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$.

The energy deposited in all calorimeter layers of an element is summed to give the tower energy.

The next step involves 'sliding' a rectangular window of size 3×5 towers across the tower grid in steps of $\Delta\eta$ and $\Delta\phi$. If the sum of the transverse energy of all towers in a window is a local maximum and is above a given threshold, $E_{\mathrm{T}}^{\mathrm{thresh}} = 2.5$ GeV, a cluster seed is formed. Finally, clusters are formed around the seeds using a clustering algorithm, as described in [149].

Track reconstruction is composed of two steps: *pattern recognition* and *track fitting* [150]. When a particle traverses the ID, it generates a number of electronic signals or 'hits' in the layers of silicon detectors (pixel and SCT, Sections 7.4.1 and 7.4.2 respectively). The pattern recognition algorithm is employed to identify which hits belong to which tracks. It searches for a track seed with $p_{\mathrm{T}} > 1$ GeV comprising three hits in different layers of the silicon detectors, and attempts to extend this seed to a full track of > 7 hits using the *pion hypothesis* for energy loss [151]. If this fails, it is retried using a modified pattern recognition algorithm, which uses an *electron hypothesis* to account for energy loss at each material surface due to *bremsstrahlung* [152].

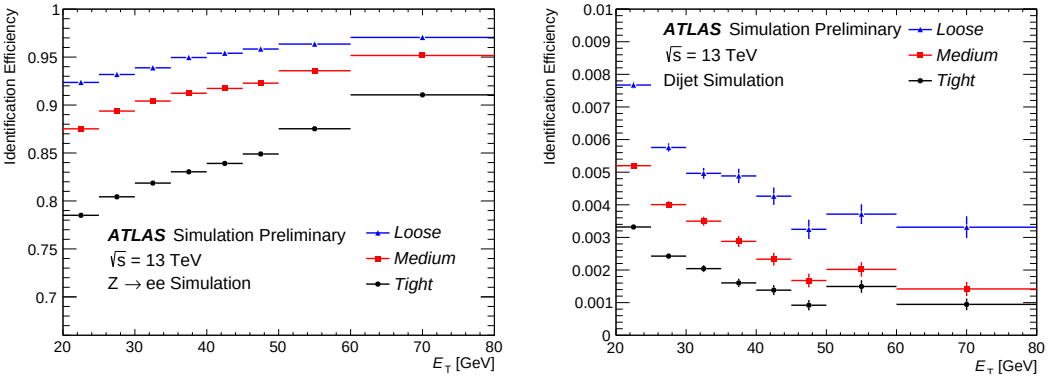
Track candidates are fitted using the ATLAS global χ^2 track fitter [153], where the fit aims to minimise χ^2, which is a function of: the residuals (the difference between a measurement and the track prediction) of the track parameters (impact parameter, direction and momentum) at the vertex; and the scattering angles. An extra term is introduced for the electron hypothesis to account for the increase in χ^2 due to bremsstrahlung losses. Then, tracks are extended into the TRT (see Section 7.4.3), where they are matched with high-threshold (HT) hits associated with transition radiation [151].

Once reconstructed, tracks with $\geqslant 4$ hits are extrapolated and matched with EM clusters to which their proximity in η and ϕ passes one of two series of loose requirements [154]:

1. $\Delta\phi < 0.2$ if the EM cluster is on the side the track is bending towards or $\Delta\phi < 0.05$ on the opposite side, and $\Delta\eta < 0.05$;
2. after rescaling the track momentum to the measured cluster energy, $\Delta\phi < 0.1$ if the EM cluster is on the side the track is bending towards or $\Delta\phi < 0.05$ on the opposite side.

The aim of the second criterion is to recover tracks with large curvature that have undergone significant energy loss.

An electron is reconstructed if at least one track is matched to the seed cluster. The track parameters of these electron candidates are then re-estimated using an electron-specific



(a) Electron identification efficiency (b) Hadron misidentification rate

FIGURE 9.1: The efficiency with which (a) electrons are identified in $z \to ee$ decays and (b) hadrons are misidentified as electrons in di-jet samples, both as a function of transverse energy, E_{T}. These efficiencies are calculated using MC simulations.

track fitter [155]; this fitter produces a better estimate by accounting for bremsstrahlung losses. The final electron reconstruction is performed using these new tracks and tighter requirements for track-cluster matching (the separation in ϕ must be < 0.1, rather than < 0.2).

9.1.2 Electron Identification

Objects built by the electron reconstruction algorithms are not necessarily signal objects, i.e. prompt electrons. Other objects, for example hadronic jets or non-prompt electrons, may produce a signature which is reconstructed as an electron. Therefore, a likelihood technique is employed to select only signal objects [154]. The likelihood method is based on information from the ID and ECAL that discriminates between electrons and background objects (see [150]).

The final discriminant produced by the electron likelihood method is used to define three working points, referred to as (in order of increasing background rejection): *Loose*, *Medium* and *Tight*. The three working points are defined by the selection applied to the likelihood discriminant, where the samples selected by the working points are subsets of one another. The efficiency with which electrons are correctly identified in $Z \to ee$ decays is shown in Figure 9.1a for the three working points. Figure 9.1b shows the rate at which hadrons are misidentified as electrons in di-jet samples.

For both analyses in this thesis, electron candidates are required to pass the *Loose* working point of the identification algorithm, which gives an efficiency of 95%. Additionally, they are required to have $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 2.47$. Electron candidates in the barrel-endcap transition region of the ECAL ($1.37 < |\eta| < 1.52$) are vetoed.

Further, analysis-specific, selection cuts are applied to signal electrons based on the trigger used and a higher electron quality is required. These are described in Section 11.2.

Electrons are also required to have a certain level of 'isolation' from other particles in order to further reject background objects (including light hadrons misidentified as electrons, converted photons and electrons from hadron decays). The analyses in this thesis impose a *Loose* isolation based on $p_{\mathrm{T}}^{\mathrm{cone}}$, i.e. the scalar sum of the transverse momentum of the tracks in a cone of given ΔR around the electron [150]. The value of ΔR is dependent on the p_{T} of the electron – this gives an improvement in performance for electrons from decays of high-p_{T} particles.

9.2 Muons

9.2.1 Muon Reconstruction

The reconstruction of muons utilises the MS ($|\eta| < 2.7$) as well as the ID ($|\eta| < 2.5$). Initially, the reconstruction is performed independently in each sub-detector; the information collected is then combined to form the final muon tracks [156]. The reconstruction in the ID is performed as for any other track; therefore, this section concentrates on the muon reconstruction in the MS and the combined reconstruction.

The initial step of the reconstruction in the MS is to search for hit patterns in each muon chamber (see Section 7.6) to build *segments*. Segments in multiple layers are then matched to form muon track candidates according to their relative positions and angles. At least two matching segments are required to build a track, apart from in the barrel-endcap transition region, where a track may be formed from one high-quality segment. The hits associated with each track candidate are fitted with a χ^2 fit and are accepted if the χ^2 passes the selection criteria. Hits with a large contribution to the χ^2 are removed and the track fit is repeated.

The track candidates from the MS are combined with information from the ID and ECAL using various algorithms. They are categorised into four muon 'types' according to the sub-detectors used in their reconstruction:

- *Combined* muons are formed using a re-fit of the hits from tracks in the MS and ID in the region $|\eta| < 2.5$. This category gives the best rejection of fake muons and has the best momentum resolution.

	$4 < p_{\mathrm{T}} < 20$ GeV		$20 < p_{\mathrm{T}} < 100$ GeV	
Selection	$\epsilon_{\mu}^{\mathrm{MC}}[\%]$	$\epsilon_{\mathrm{Hadrons}}^{\mathrm{MC}}[\%]$	$\epsilon_{\mu}^{\mathrm{MC}}[\%]$	$\epsilon_{\mathrm{Hadrons}}^{\mathrm{MC}}[\%]$
Loose	96.7	0.53	98.1	0.76
Medium	95.5	0.38	96.1	0.17
Tight	89.9	0.19	91.8	0.11
High-p_{T}	78.1	0.26	80.4	0.13

TABLE 9.1: The efficiency with which prompt muons from W boson decays are identified, $\epsilon_{\mu}^{\mathrm{MC}}$, and the rate at which hadrons decaying in flight are misidentified as prompt muons, $\epsilon_{\mathrm{Hadrons}}^{\mathrm{MC}}$, are shown for each of the four working points. The efficiencies are measured using simulated $t\bar{t}$ events for low- and high-p_{T} muons with $|\eta| < 2.5$.

- *Segment-tagged* muons are tracks in the ID that are tagged as muons if, when extrapolated to the MS, they are matched with at least one track segment. Segment-tagged muons are reconstructed in the $|\eta| < 2.5$ region.
- *Calorimeter-tagged* muons are used in the region $|\eta| < 0.1$ which is only partially covered by the MS. A calorimeter-tagged muon is formed when a track in the ID matches with an appropriate energy deposit in the ECAL.
- *Extrapolated* muons are reconstructed based on tracks from the MS only, with a loose requirement that they originate from the IP. Extrapolated muons are used to extend the muon reconstruction range into the region $2.5 < |\eta| < 2.7$, which is not covered by the ID.

9.2.2 Muon Identification

Muon identification is necessary to suppress background, of which the majority originates from charged hadron decays. These muons are often identifiable by a 'kink' in the reconstructed track, resulting in a poor fit quality of the track; there are several variables that provide good discrimination between signal and background muons and are used in the identification procedure, as detailed in [156]. Four muon selection categories are provided: *Loose*, *Medium*, *Tight* and *High-*p_{T}. As well as requirements on variables associated with the muon tracks, these categories use different combinations of the muon types.

Table 9.1 summarises the muon identification efficiencies for prompt muons from W boson decays, alongside the misidentification rate of hadrons decaying in flight, measured using MC-simulated $t\bar{t}$ events. The signal identification and background rejection are generally higher for higher-p_{T} objects.

Muons used in the searches in this thesis are required to have $p_{\mathrm{T}} > 7$ GeV and $|\eta| < 2.7$ and pass the *Loose* identification working point. This criteria uses all four muon types. Further selection cuts are applied to signal muons based on the trigger used and a higher muon quality is required. These are described in Section 11.2.

Muons are also required to be isolated from other detector activity using the *Loose* criteria. Muon isolation follows a similar process to that for electrons, as described in Section 9.1.2.

9.3 Jets

As described in Section 2.2.1, partons cannot exist individually due to colour confinement; quarks and gluons produced in proton-proton collisions must *hadronise*. This results in the creation of collimated showers of particles, referred to as *hadronic jets*, which are the dominant physics objects at the LHC. Jets produce clusters of energy in the hadronic calorimeters and tracks in the ID, from which they can be reconstructed, allowing the momentum and energy of the original parton to be measured.

Jets that originate from a b-quark (henceforth referred to as b-jets) can be differentiated from other types of jets due to their unique signature. Their identification, which is vital for the analyses in this thesis, is detailed in Section 9.4.

9.3.1 Jet Reconstruction

The jets used in the analyses in this thesis are reconstructed using the *anti-k_t* algorithm [157], one of a number of jet-finding algorithms [158], which uses topological clusters of energy deposits in calorimeter cells (known as *topo-clusters*) as input. These jets are known as *calorimeter jets*. Jets may also be reconstructed from tracks; these are known as *track jets*. Track jets are less likely to originate from pile-up activity as only tracks originating from the primary vertex are used. However, because the acceptance of the ID is limited to $|\eta| < 2.5$, most ATLAS analyses rely on calorimeter jets and this section describes only calorimeter jets.

The topological algorithm [149] begins with a seed cell and builds a cluster by iteratively adding neighbouring cells, providing these cells have significant energy relative to the expected noise. There are two main sources of noise in the calorimeter: the readout electronics and pile-up. The pile-up noise is estimated using simulation. The seed cell is required to pass a noise threshold higher than that for the neighbouring cells, and a neighbouring cell may serve as a further seed if it fulfils an intermediate noise

requirement. The use of a lower threshold at the cluster perimeter ensures that jet tails are not discarded. Providing they satisfy a p_{T} requirement of 7 GeV, the clusters are then used to form jets.

The jet reconstruction algorithm chosen for the analyses in this thesis, and the main jet algorithm used in ATLAS, is the anti-k_t algorithm with distance parameter $R = 0.4$ [157]. The anti-k_t algorithm is one of three main sequential combination algorithms, which group particles based on distance parameters, d, between pairs of clusters, i and j, and between cluster i and the beam, B. These parameters are defined as

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}) \frac{\Delta R_{ij}^2}{R^2}, \tag{9.1}$$

$$d_{iB} = k_{ti}^{2p}, \tag{9.2}$$

where $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, ηi and ϕi are the transverse momentum, pseudorapidity and azimuth of cluster i. For the anti-k_t algorithm specifically, $p = -1$.

The sequential algorithms follow an iterative process:

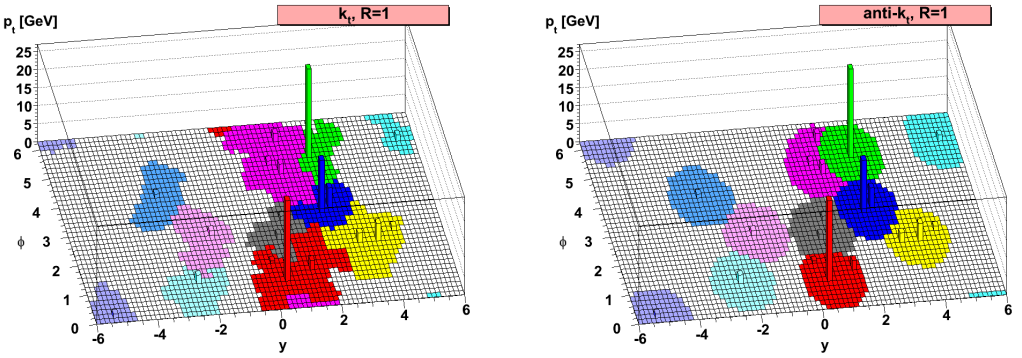
- The two clusters with the minimum d_{ij} are combined into a single *proto-jet*.
- If the minimum value of d_{ij} is d_{iB}, i.e. it is between a cluster and the beam line, the cluster is labelled as a jet and is removed from the list.
- The process continues until the separation between all clusters and proto-jets is $> R$, where R is the required jet radius.

The choice of $p = -1$ (Equation 9.1), which characterises the anti-k_t algorithm, ensures that the jets are formed around the hardest hadrons in the event, giving them an approximately conical shape. This is illustrated in Figure 9.2, which compares jets formed by the anti-k_t algorithm with those formed by the k_t algorithm ($p = 1$ in Equations 9.1 and 9.2).

9.3.2 Jet Calibration and Correction

Various calibration steps are performed to account for detector response and performance.

Since the energy of the calorimeter cells is measured at the electromagnetic scale, the *local cluster weighting* (LCW) calibration is applied to topo-clusters (before they are



(a) k_t algorithm (b) Anti-k_t algorithm

FIGURE 9.2: A sample parton-level event, together with many random soft 'ghosts', clustered using (a) the k_t jet algorithm and (b) anti-k_t jet algorithm. The figures illustrate the region around the hard jets within which the soft ghosts are clustered into that jet. When using the anti-kt algorithm, the hard jets are all approximately circular, with only the softer jets forming more complex shapes. Taken from [157].

used to form jets) to account for the differences in the detector response to hadronic and electromagnetic showers. The LCW calibration is in the form of a number of weights taken from MC simulation of single-pion events [159].

The *jet energy scale* (JES) calibration is a p_{T}- and η-dependent correction which ensures that the reconstructed jet energy matches that of simulated truth jets [160]. The JES calibration accounts for: the differences in response of the hadronic and electromagnetic calorimeters; *out-of-cone effects*, i.e. the particles that were not reconstructed in the jet; leakage of particles outside the calorimeter; and energy loss in dead material.

Furthermore, jets which may arise from pile-up events are vetoed using the *jet vertex tagger* (JVT) [161]. This is a multivariate combination of two track-based variables specifically developed to separate hard-scatter jets from pile-up jets:

- The *jet vertex fraction* (JVF) is used to quantify the likelihood of a jet having originated from the primary vertex [162]. It is defined as the fraction of the summed track p_{T} for all tracks matched to a given jet and associated with the primary vertex, relative to the summed p_{T} for all tracks matched to the jet (where a track is matched to the jet if, when extrapolated from the ID to the calorimeter, it is within a cone of size ΔR around the jet). The JVT uses a version of the JVF corrected for the number of reconstructed primary vertices in the event.

- The variable $R_{p_{\mathrm{T}}}$ is defined as the scalar sum of the p_{T} of the tracks that are associated with the jet and originate from the hard-scatter vertex, divided by the calibrated jet p_{T}. For pile-up jets, very little p_{T} is expected from the hard-scatter vertex.

A *jet cleaning* selection is applied in order to veto any 'fake' jets, which arise from non-collision background events, such as cosmic rays, or from detector effects. Finally, all jets in the analysis are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.4$.

9.4 *b*-Jets

Jets originating from the hadronisation of *b*-quarks are a key component in many high precision measurements and new physics searches performed by the ATLAS experiment, including those in this thesis. Whilst *b*-jet identification is challenging due to the large backgrounds from *c*- and light-jets (where light-jets are jets from *u*-, *d*- or *s*-quarks or gluons), efficient identification algorithms can *tag* *b*-jets by exploiting their unique features:

- The heavy hadrons from which *b*-jets arise have a long lifetime ($\tau \simeq 1.5$ ps, $c\tau \simeq 450$ µm), allowing them to travel a measurable distance through the detector before decaying. This distance can be seen in Figure 9.3, where it is labelled L_{xy}. As such, the point at which the *b*-hadron decays, known as the *secondary vertex*, is significantly displaced relative to the primary vertex. It is also possible to reconstruct the *tertiary vertex*, where the *b*-quark decays.
- The large mass of *b*-jets leads to decay products with a larger transverse momentum relative to the jet axis than would be expected for jets from light partons.
- Finally, *b*-hadrons have a large branching ratio for semileptonic decays, resulting in the presence of leptons in *b*-jets.

9.4.1 *b*-Tagging Algorithms

For the purposes of this thesis, the identification of *b*-jets is based on three distinct *b*-tagging algorithms, the outputs of which are combined in a multivariate discriminant known as MV2c10. MV2c10 utilises a BDT in order to give the best separation between *b*-, *c*- and light- jets.

The ATLAS *b*-tagging algorithms used in Run II are:

- impact parameter based algorithms (IP2D, IP3D), which exploit the fact that the tracks within the *b*-jet do not point to the primary vertex;

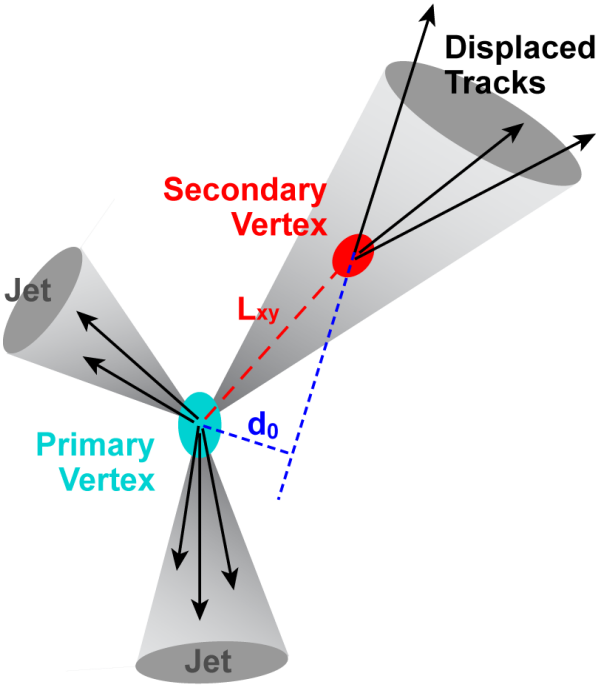


FIGURE 9.3: Diagram of an event with a b-jet. The b-hadron decay occurs at the secondary vertex at a distance L_{xy} from the primary vertex. The distance labelled d_0 is the transverse impact parameter; this has a large value for b-hadrons. Taken from [163].

- secondary vertex finding (SV1), which reconstructs a displaced secondary vertex within the jet;
- decay chain multi-vertex fit (JetFitter), which aims to reconstruct the full b-hadron decay chain;
- the multivariate discriminant (MV2c10), which combines the above three algorithms with relevant kinematic variables in a BDT.

9.4.1.1 Impact Parameter Based Algorithms

The impact parameter algorithms, IP2D and IP3D, utilise the long lifetime of hadrons containing a b-quark - see Figure 9.3. The distance labelled d_0 is known as the transverse impact parameter and is defined as the distance of closest approach of the track to the primary vertex in the transverse $(r - \phi)$ plane. The longitudinal impact parameter, $z_0 \sin\theta$, is defined as the distance of the track to the primary vertex in the longitudinal plane at the point of closest approach in the transverse plane. The large positive impact parameters that are characteristic of b-jets are used to differentiate them from other jets. Secondary vertices placed behind the primary vertex relative to the direction of the jet are usually attributed to background.

Events must pass the following selection to be considered by the impact parameter algorithms:

- track $p_{\mathrm{T}} > 1$ GeV,
- $|d_0| < 1$ mm and $|z_0 \sin\theta| < 1.5$ mm, and
- $\geqslant 7$ silicon hits.

The impact parameter algorithms are based on a log-likelihood ratio (LLR) method. IP2D uses the transverse impact parameter significance, d_0/σ_{d_0}, where σ_{d_0} is the uncertainty on d_0. IP3D also includes the longitudinal impact parameter significance, $z_0 \sin\theta/\sigma_{z_0 \sin\theta}$. Probability density functions (PDFs) for these are calculated from MC simulations on a per-track basis, individually for each of the b-, c- and light-jets, such that the algorithm defines, for example, the probability of a given track belonging to a b-jet [164].

The track weight, w^{track}, is the ratio of the probabilities of, for example, the b- and light-flavour jet hypotheses [165]:

$$w_{bu}^{\mathrm{track}} = \frac{p_b}{p_u}, \tag{9.3}$$

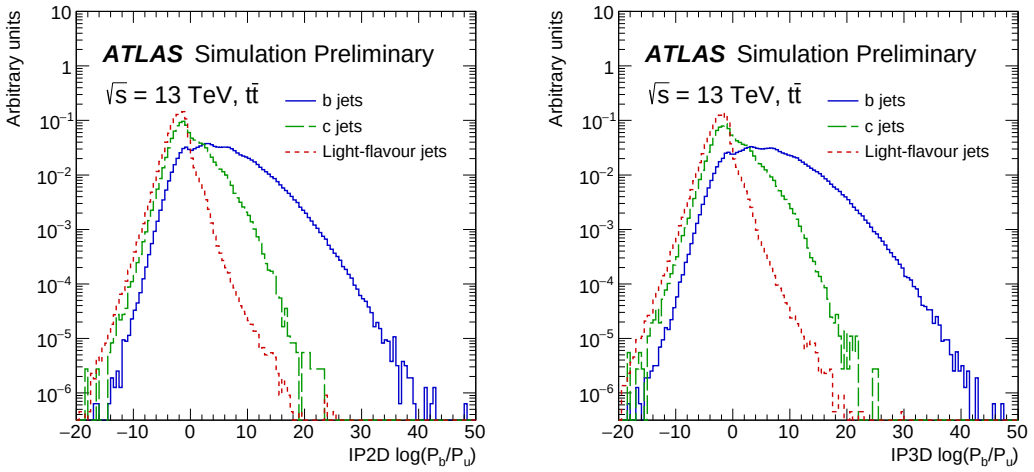
where p_b is the probability of a track belonging to a b-jet, and p_u is the probability of a track belonging to a light-jet [164]. Similarly, the probability of a track belonging to a c-jet is represented by p_c. Next, the jet weight, w^{jet}, is calculated as the sum of the logarithms of the track weights for all tracks in the jet. Again, to separate b-jets from light-jets:

$$w_{bu}^{\mathrm{jet}} = \sum_{i=1}^{N} \log w_{bu}^{\mathrm{track}}, \tag{9.4}$$

LLRs calculated using the ratio of b- and light-jet hypotheses are shown in Figure 9.4 using both the IP2D and IP3D algorithms. Additional LLR discriminants are constructed according to ratios of the b- and c-jet, and c- and light-jet hypotheses.

9.4.1.2 Secondary Vertex Finding

The purpose of the secondary vertex finding algorithm, SV1, is to reconstruct the secondary vertex formed by the decay products of the b-hadron, including those resulting from the subsequent c-hadron decay.



(a) IP2D (b) IP3D

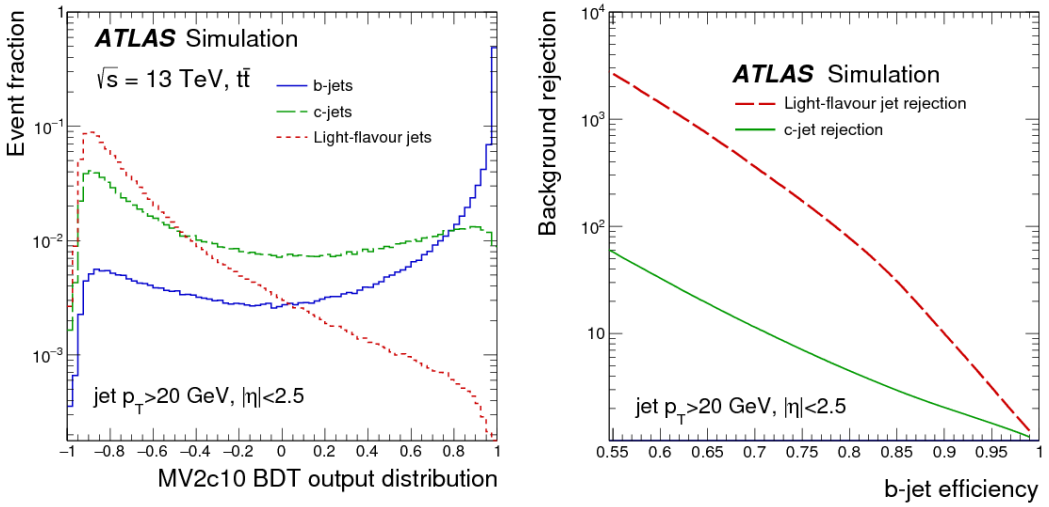
FIGURE 9.4: The log-likelihood ratio (LLR) for (a) the IP2D and (b) the IP3D b-tagging algorithms for b-jets (solid blue), c-jets (dashed green) and light-jets (dotted red) in $t\bar{t}$ events. The LLR discriminants shown here are calculated as a ratio of the b- and light-jet hypotheses. Taken from [164].

The first step is to reconstruct two-track vertices, testing all track pairs within the jet. Any vertices likely to originate from the decay of a long-lived particle (such as a K_s or Δ), photon conversions, or hadronic interactions with the detector material, are rejected [164]. The tracks that remain are then used to fit a new vertex, with outliers being removed.

The algorithm then uses properties associated with this vertex to discriminate between b-jets and light-jets. These include the vertex mass, the transverse decay length, and the energy fraction, defined as the energy from the tracks in the displaced vertex relative to all tracks reconstructed within the jet.

9.4.1.3 Multi-Vertex Fit

The decay-chain multi-vertex reconstruction algorithm, JetFitter [166], tries to reconstruct the full b-hadron decay chain by taking advantage of the characteristic topological structure of weak b- and c-hadron decays that occur within the jet. The algorithm uses a Kalman filter [167] to find a single line upon which the primary vertex and the b- and c-hadron decay vertices lie and the positions of the vertices on it. The benefit of this approach is that it can resolve the b- and c-hadron vertices even if there is only one track attached to each of them, resulting in improved separation between b-jets and other jets.



(a) MV2c10 (b) Rejection factors

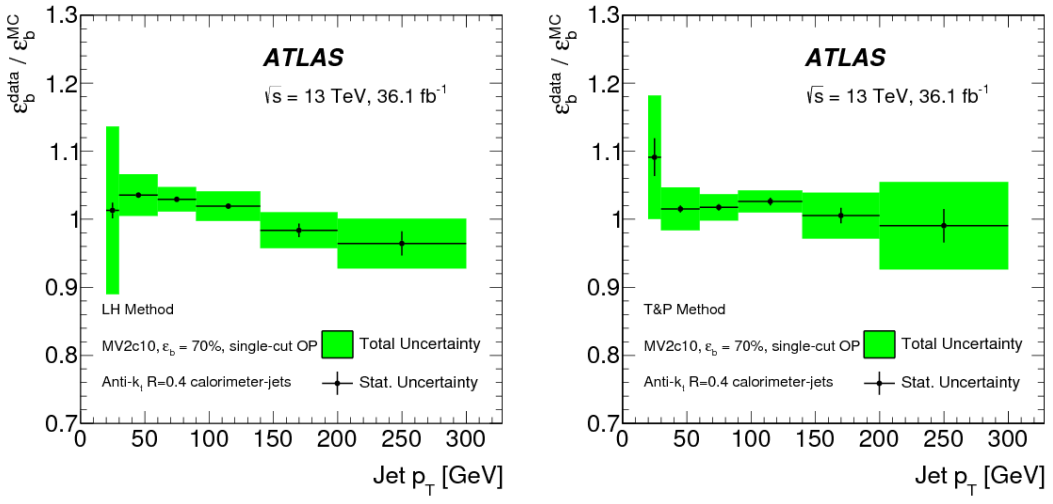
FIGURE 9.5: (a) The MV2c10 BDT output for b-jets (solid blue), c-jets (dashed green) and light-jets (dotted red) in simulated $t\bar{t}$ events. (b) The light-jet (red dashed) and c-jet (solid green) rejection factors as a function of the b-jet tagging efficiency associated with the MV2c10 algorithm. Taken from [168].

9.4.1.4 Multivariate Discriminant

The outputs from the three b-tagging algorithms (IP2D, IP3D, SV1 and JetFitter) are propagated, alongside basic kinematic properties (p_{T} and η) of the jets, to the MV2 BDT discriminant [168]. Combining the three basic algorithms optimises the separation achievable between b-, c- and light-jets.

The input variables included in the BDT training are listed in [169]. The version of the MV2 discriminant used in the analyses in this thesis, MV2c10, is trained on simulated $t\bar{t}$ events with b-jets as signal, and light- and c-jets as background (the number of c-jets is 10% that of light-jets). The performance of the BDT is then tested on a statistically independent $t\bar{t}$ sample, as shown in Figure 9.5a. Figure 9.5b shows the light- and c-jet rejection factors as a function of the b-jet tagging efficiency, where the rejection factors are defined as the inverse of the efficiency of mistakenly tagging a light- or c-jet as a b-jet.

Four working points are defined, corresponding to 60%, 70%, 77% and 85% b-jet tagging efficiencies in simulated $t\bar{t}$ events. These are achieved by a set of selection criteria placed upon the MV2c10 output distribution. In the analyses included in this thesis, the 70% working point is used, which gives rise to a rejection factor of around 381 (12) against light- (c-) jets.



(a) LH method (b) T&P method

FIGURE 9.6: Data-to-simulation scale factors as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ using (a) the likelihood method and (b) the tag-and-probe method. Taken from [170].

9.4.2 *b*-Tagging Efficiency Calibration

As alluded to in Section 9.4.1.4, the performance of the *b*-tagging algorithm is characterised by the *b*-jet *tagging efficiency*, ε_b, i.e. the probability of tagging a *b*-jet, and the *mistagging efficiency*, ε_c and $\varepsilon_{\mathrm{LF}}$, i.e. the probabilities of mistakenly identifying a *c*- or light-jet as a *b*-jet, respectively. These are calculated by applying the MV2 discriminant to a simulated $t\bar{t}$ sample. However, calibration is required to account for differences between data and simulation.

Each of the four working points, corresponding to *b*-jet tagging efficiencies of 60%, 70%, 77% and 85%, are calibrated separately for *b*-, *c*- and light-jets (the light-jet calibration is discussed in Chapter 10). The calibrations are provided in the form of data-to-simulation efficiency scale factors as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ and, if relevant, $|\eta^{\mathrm{jet}}|$, calculated according to

$$\mathrm{SF} = \frac{\varepsilon^{\mathrm{data}}}{\varepsilon^{\mathrm{sim}}}, \tag{9.5}$$

where $\varepsilon^{\mathrm{data}}$ is the efficiency measured in data and $\varepsilon^{\mathrm{sim}}$ is the efficiency predicted by simulation. These scale factors can then be used in physics analysis on a per-jet basis to correct the rate of events after applying a *b*-tagging requirement (i.e. one of the working points mentioned above).

Figure 9.6 shows the data-to-simulation efficiency scale factors for *b*-jets as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ for two different calibration methods (as described in [170]).

9.4.3 Truth-Tagging

The analyses in this thesis categorise events according to the number of b-tagged jets, e.g. the two b-tag category may be used as the signal region and the one b-tag category as a validation region. When using the *direct tagging* method, jets that fail the b-tagging requirements are removed. MC-simulated backgrounds with low b-jet multiplicity have few events that survive these requirements, leading to badly estimated background processes.

The *truth-tagging* method aims to keep all of the MC events generated whilst also maintaining the correct normalisations and shapes determined by the b-tagging selection. To achieve this, each event is weighted by the probability that a given number of jets in the event are tagged.

For an event with three jets and a requirement that two jets are b-tagged, the weight assigned to the event, w, is given by

$$w = \varepsilon_1\varepsilon_2(1-\varepsilon_3) + \varepsilon_1(1-\varepsilon_2)\varepsilon_3 + (1-\varepsilon_1)\varepsilon_2\varepsilon_3, \tag{9.6}$$

where the efficiency with which jet i is b-tagged is given by ε_i. The b-tagging efficiency is parametrised as a function of the flavour, p_{T} and η of the given jet, as well as the process type.

The two jets to be assigned as b-jets are randomly selected based on their tagging efficiency, such that the probability for jets 1 and 2 to be selected as the b-jets is given by

$$P(1 \cap 2) = \frac{\varepsilon_1\varepsilon_2(1-\varepsilon_3)}{w}. \tag{9.7}$$

This method is applied to all samples except the signals, single-top and top-quark pair-production, and $Z \to \tau\tau + (bb, bc, cc)$ (where a Z boson, produced in association with any combination of two b- and c-jets, decays to a pair of τ-leptons).

9.5 Hadronically-Decaying τ-Leptons

As the heaviest lepton in the SM, the τ-lepton decays before reaching the detector volume and can only be observed through its decay products. Hadronic decays of τ-leptons usually result in either one or three charged pions, some neutral pions and a

τ-neutrino (or τ-antineutrino). This results in an experimental signature of a collimated calorimeter shower with either one or three associated tracks.

9.5.1 τ_{had} Reconstruction

Jets formed using the anti-k_t algorithm with a distance parameter of $R = 0.4$ (see Section 9.3.1), satisfying $p_{\text{T}} > 10$ GeV and $|\eta| < 2.5$, are used as seeds for the τ_{had} reconstruction algorithm [171, 172]. Candidates in the barrel-endcap transition region of the calorimeter ($1.37 < |\eta| < 1.52$) are vetoed.

In events with multiple simultaneous interactions, the chosen primary vertex is not necessarily that from which the τ_{had} originates. The *τ-vertex* (TV) *association algorithm* is employed to select the τ_{had} production vertex from the primary vertex candidates. It takes as input all τ_{had} candidate tracks in a cone of $\Delta R < 0.2$ around the jet seed direction. The TV is the vertex to which the greatest fraction of the sum of the p_{T} of the tracks is matched.

Track selection in the τ_{had} reconstruction requires tracks to fulfill the following: $p_{\text{T}} > 1$ GeV, at least two associated hits in the pixel detector, and at least seven hits (in total) in the pixel and the SCT detectors. Furthermore, the distance of closest approach of the track to the TV must satisfy: $|d_0| < 1.0$ mm in the transverse plane and $|\Delta z_0 \sin\theta| < 1.5$ mm longitudinally. Of the tracks that pass these criteria, those in the *core region* ($\Delta R < 0.2$) are used in the classification of the associated τ_{had} as a function of its number of associated tracks. The tracks in the *isolation region* ($0.2 < \Delta R < 0.4$) are used alongside those in the core region to produce the variables used in the τ_{had} identification.

The distribution of the number of tracks associated to τ_{had} candidates is shown, separately for 1- and 3-prong decays, in Figure 9.7. The number of prongs is susceptible to underestimation due to tracking inefficiency, or overestimation due to tracks from photon conversions passing the track selection criteria.

9.5.2 τ_{had} Identification

The visible experimental signature associated with hadronically-decaying τ-leptons is very similar to that of quark- or gluon-initiated jets. The τ_{had} reconstruction process provides little rejection of the jet background; the main rejection is performed by the τ_{had} identification algorithms.

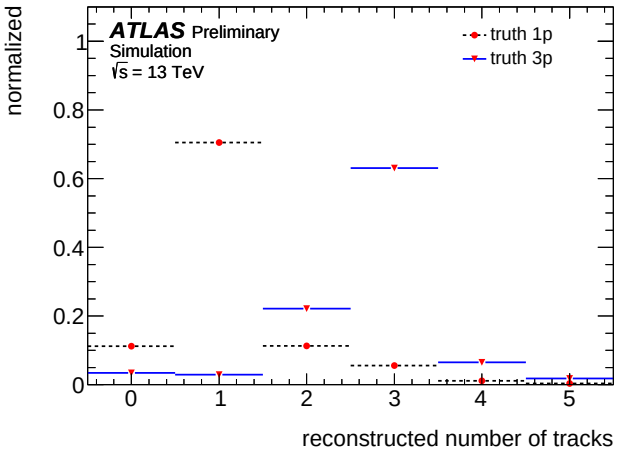


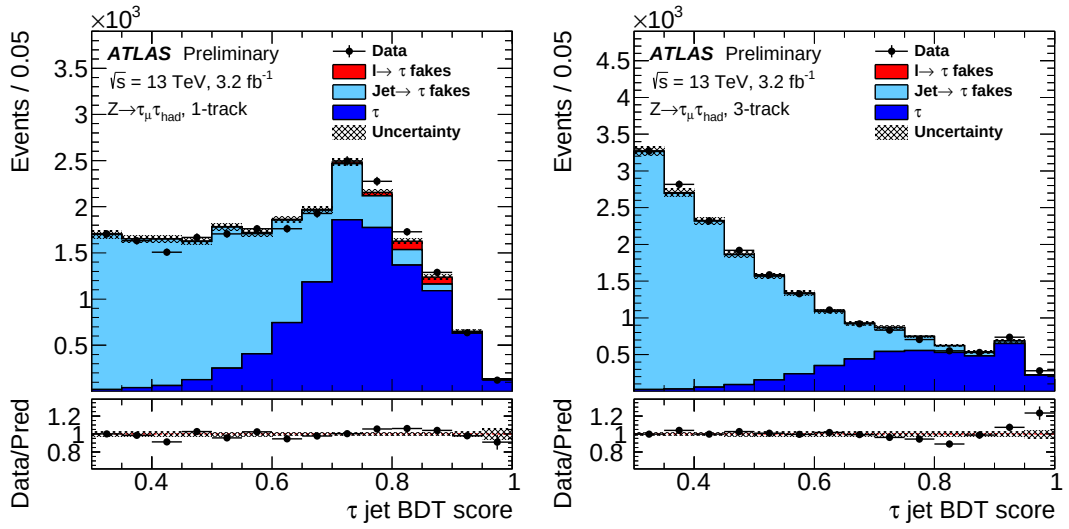
FIGURE 9.7: The number of tracks reconstructed for τ_{had} candidates with true 1-prong (dot, dashed error bar) and 3-prong (triangle, solid error bar) decays. Taken from [172].

The τ_{had} identification process uses information based upon the tracks and topo-clusters in the core and isolation regions defined in Section 9.5.1. Multivariate BDT discriminants are employed to capitalise on the differences observed between signal and background in the longitudinal and lateral shower shape and the number of pions. The variables used as input for the BDTs in Run II are listed in [172]. Separate BDTs are trained for 1- and 3-prong decays; these are shown in Figure 9.8. The BDT score is transformed such that it is independent of the p_{T} of the true-τ_{had} and the event pile-up conditions.

Three working points, labelled *Loose*, *Medium* and *Tight*, are provided, corresponding to increasing τ_{had} identification efficiencies. For the analyses in this thesis, τ_{had} candidates are required to pass the *Medium* working point (which has signal efficiencies of 0.55 for one-prong τ_{had} candidates and 0.4 for three-prong [172]) and have $p_{\text{T}} > 20$ GeV, $|\eta| < 2.5$, one or three tracks, and unit charge.

9.5.3 Fake Hadronic τ-Leptons

Fake hadronic τ-leptons are jets that have been misidentified as hadronic τ-leptons. To provide regions enriched in events where a jet fakes a τ_{had}, an anti-ID-τ_{had} selection is defined. An anti-ID-τ_{had} is a τ-candidate that has failed the *Medium* identification requirement but has a BDT score greater than 0.35. The BDT score requirement ensures that the properties of the jet (or anti-ID-τ_{had}) resemble those of a real τ_{had} and ensures that the composition of quark- and gluon-initiated jets is similar to that of the signal region.



(a) 1-prong (b) 3-prong

FIGURE 9.8: The BDT discriminant output distribution for (a) 1-prong and (b) 3-prong τ_{had} candidates, with only the statistical uncertainty shown. Taken from [173].

9.6 Missing Transverse Momentum

The transverse component of the momentum of a pair of colliding protons can be assumed to be zero. From the law of conservation of momentum, this should also be the case for the decay products. A deviation of the total transverse momentum from zero can indicate the existence of a non-interacting particle. In the SM, the only particles that do not deposit energy in the detector are neutrinos, as they are weakly interacting leptons which do not undergo the strong or electromagnetic forces (although some theories beyond the SM predict the existence of additional weakly-interacting particles).

The negative vector sum of the transverse momentum collected in the detector (within the region $|\eta| < 4.9$) is known as the *missing transverse momentum* ($\boldsymbol{p}_{\text{T}}^{\text{miss}}$ and its modulus $E_{\text{T}}^{\text{miss}}$),[1] from which the existence of an 'invisible' particle can be inferred [174]. The reconstruction of $E_{\text{T}}^{\text{miss}}$ takes into account all reconstructed, calibrated physics objects, as well as *soft-event* signals consisting of reconstructed tracks originating from the primary vertex but not associated with any other hard objects (electrons, muons, τ-leptons or jets) contributing to the $E_{\text{T}}^{\text{miss}}$ definition.

As the reconstruction procedures for the different types of object are independent of one another, it is likely that one detector signal will be used in the reconstruction of more than one object – resulting in double counting of the signal in the $E_{\text{T}}^{\text{miss}}$ calculation. This is avoided by the *signal ambiguity resolution*. The $E_{\text{T}}^{\text{miss}}$ reconstruction summation is performed in sequence such that the object types are added with varying priority.

[1] Masses are neglected in calculating the missing transverse momentum; therefore, the missing transverse energy is defined as the magnitude of the missing transverse momentum.

Electrons enter the reconstruction first, with subsequent particles being fully rejected should they share a calorimeter signal with an object that has already been included. Since muons are reconstructed from ID and MS tracks alone, they do not overlap with the other particles in the calorimeter. The summation is performed as

$$\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{\substack{\text{selected}\\\text{electrons}}} \boldsymbol{p}_{\mathrm{T}}^{e} - \sum_{\substack{\text{accepted}\\\text{photons}}} \boldsymbol{p}_{\mathrm{T}}^{\gamma} - \sum_{\substack{\text{accepted}\\\tau-\text{leptons}}} \boldsymbol{p}_{\mathrm{T}}^{\tau_{\mathrm{had}}} - \sum_{\substack{\text{selected}\\\text{muons}}} \boldsymbol{p}_{\mathrm{T}}^{\mu} - \sum_{\substack{\text{accepted}\\\text{jets}}} \boldsymbol{p}_{\mathrm{T}}^{\mathrm{jet}} - \sum_{\substack{\text{unused}\\\text{tracks}}} \boldsymbol{p}_{\mathrm{T}}^{\mathrm{track}}, \tag{9.8}$$

where the final term involves the soft-event signals, i.e. the reconstructed tracks which have not already been included in any objects [174].

9.7 Missing Mass Calculator

Because of the presence of multiple neutrinos in τ-lepton decays, it is difficult to accurately reconstruct the mass of a heavy resonance decaying to a di-τ final state. The *Missing Mass Calculator* (MMC) method [175] provides a full reconstruction of the event topology. The advantages of this method include an improvement of the resolution of the invariant mass of the di-τ system and a low failure rate in finding a solution.

Calculating the di-τ invariant mass requires solving an underconstrained system of equations for six to eight unknowns, depending on the number of neutrinos in the final state: the x-, y- and z-components of the invisible momentum carried by the neutrinos for each of the two τ-leptons in the event, and the invariant mass of the two neutrinos from any leptonic τ-lepton decays. The kinematic constraints involved are the x- and y-components of the missing p_{T}, and the visible masses of both τ-lepton candidates.

Although it is not possible to find an exact solution based on the available information, not all solutions are equally likely. The concept of the MMC is to reconstruct the momentum for each neutrino pair using all of the known constraints and perform a scan over the undetermined variables. Each possibility is then weighted by its probability, as determined by the τ-lepton decay topologies and missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) resolution. A histogram of the final discriminant, the mass of the di-τ system, is filled with the weighted scan points, such that the most probable value is determined as the histogram maximum.

The efficiency of the MMC is defined as the ratio between the number of events for which a solution can be determined and the total number of events. This is typically $\sim 99\%$ for signal events [29]; in a small fraction of cases the true values of the scanned variables or the $E_{\mathrm{T}}^{\mathrm{miss}}$ are outside the scan range.

Objects	$\Delta R <$	Further requirements	Priority
e, jet	0.2	—	e
jet, e	0.4	—	jet
μ, jet	0.4	< 3 tracks with $p_{\mathrm{T}}^{\mathrm{track}} > 500$ GeV	μ
jet, μ	0.4	$\geqslant 3$ tracks with $p_{\mathrm{T}}^{\mathrm{track}} > 500$ GeV	jet
e, μ	0.2	e and μ share an ID track	μ
μ, τ	0.2	—	μ
e, τ	0.2	—	e
τ, jet	0.2	—	Priority given to: 1. *Medium* τ 2. b-tagged jet 3. Anti-ID τ 4. Light jet

TABLE 9.2: The procedure for the removal of overlapping objects based on their proximity in ΔR. For each pair of objects, one object must take priority and is kept whilst the other is removed.

9.8 Overlap Removal

Signals in the detector can be included in the reconstruction of more than one physics object, resulting in double-counting of detector measurements. An overlap removal (OR) procedure is applied to objects that have a geometric overlap in ΔR below a certain threshold; the order in which objects receive priority and the ΔR thresholds are defined in Table 9.2.

Chapter 10

Calibration of the Light-Flavour Jet b-Tagging Efficiency

As described in Section 9.4.1.4, the b-tagging efficiency and c- and light-jet mistagging efficiencies are calibrated to account for differences between data and simulation. Each of the four working points, corresponding to b-jet tagging efficiencies of 60%, 70%, 77% and 85%, are calibrated separately.

There are two methods by which the light-jet calibration (i.e. the calibration of the mistagging efficiency for light-jets) is performed by ATLAS: the *negative-tag* [176] method and the *adjusted Monte Carlo* method (hereafter referred to as adjusted-MC).

The adjusted-MC method was developed for Run II as a 'bottom-up' approach, where the tracking variables are adjusted in MC simulations to match data and compared with the nominal MC sample to achieve the calibration scale factor. It was first introduced for the analysis of data taken in 2015 and 2016 using ATLAS software release 20; this is documented in [176]. The author contributed to this result before adapting the framework for a new release of ATLAS software (release 21). Section 10.2 outlines the adjusted-MC method using figures from the documented calibration. Section 10.3 shows the initial results of the calibration using the new release.

10.1 Negative-Tag Method

The negative-tag method, which was introduced in Run I [165], uses data enriched in light-jets and applies the b-tagging algorithms described in Section 9.4.1 with some of the criteria reversed. Tracks associated with light-jets are expected to exhibit a signed impact parameter that is symmetric around zero, whereas the impact parameter of b-

and c-jets will have a larger tail at large positive values due to their longer lifetime. The IP2D and IP3D algorithms (Section 9.4.1.1) utilise this effect to discriminate between the different flavours of jets. The negative-tag method approximates the light-jet mistagging efficiency as the b-tagging efficiency found when running the IP2D and IP3D algorithm after reversing the sign of the impact parameter of the jet tracks.

10.2 Adjusted-MC Method

The adjusted-MC method uses MC samples of multi-jet events from strong interaction processes generated using Pythia 8.186 [120] with the the A14 tune [121] and the NNPDF23LO PDF set [122]. The alternative simulated samples used to provide the generator uncertainty are generated using Herwig++ [117] with the CTEQ6L1 PDF sets [130] and the UEEE5 tune [118]. All MC samples are passed through the full simulation of the ATLAS detector produced in GEANT4 [111, 112].

The adjusted-MC method proceeds as follows:

1. Accurate measurements are made of the track parameters in data and simulation.
2. The MC sample is adjusted to match the performance observed in data and the b-tagging algorithm is applied.
3. The light-jet mistagging efficiency is calculated for the nominal and adjusted MC samples and the light-jet calibration scale factor, $\mathrm{SF}_{\mathrm{LF}}$, is calculated according to

$$\mathrm{SF}_{\mathrm{LF}} = \frac{\varepsilon_{\mathrm{LF}}^{\mathrm{adjusted}}}{\varepsilon_{\mathrm{LF}}^{\mathrm{nominal}}}, \tag{10.1}$$

   where $\varepsilon_{\mathrm{LF}}^{\mathrm{adjusted}}$ is the light-jet mistagging efficiency for the adjusted simulation and $\varepsilon_{\mathrm{LF}}^{\mathrm{nominal}}$ is that for the nominal simulation.
4. Some effects cannot be replicated using MC simulation. Such effects are accounted for by applying a systematic uncertainty.

Effects which contribute to the discrepancies between data and simulation and are accounted for in the adjusted-MC calibration are:

- The resolution of the track impact parameters. The impact parameters of charged particle tracks with $p_{\mathrm{T}} = 1$ GeV ($p_{\mathrm{T}} = 20$ GeV) are measured by the ATLAS tracking system with a resolution of order 100 µm (10 µm). This is limited by

a convolution of numerous effects, such as: the intrinsic single-hit resolution, the alignment of the tracking components, multiple-scattering inside the detector material and the accuracy of track reconstruction algorithms. These are difficult to model and the track impact parameters in MC simulation require additional tuning to match data [165].

- The shapes of the distributions of the track impact parameters. These are assumed to be Gaussian but, in fact, include non-Gaussian tails. These tails occur for a number of reasons which are difficult to reproduce in simulation, such as: reconstruction issues, contamination from poor-quality tracks and the presence of secondary particles due to hadronic interactions with the material.

- The probability of reconstructing fake tracks using the *Loose* selection criteria. This accounts for a significant discrepancy between data and simulation [177].

- The fraction of long-lived strange-hadrons, such as K_s and Λ_s, in Pythia 8 simulation. This appears to be underestimated by 30% [178].

- The impact of the detector material on the rate of photon conversions and secondary hadronic interactions [179]. This is accurate to within 10% but has a large impact on the MV2c10 distribution for light-jets.

- The choice of parton shower model. The mistagging efficiency is calculated using an alternative sample produced using Herwig++.

Other effects were investigated and found to be negligible. The correction that is applied to the impact parameter resolution of the MC, $\sigma^{\text{Corr.}}(d_0)$, is calculated according to

$$\sigma^{\text{Corr.}}(d_0) = \sqrt{(\sigma^{\text{Data}}(d_0))^2 - (\sigma^{\text{MC}}(d_0))^2}, \tag{10.2}$$

where $\sigma^{\text{Data}}(d_0)$ is the d_0 impact parameter resolution in data and $\sigma^{\text{MC}}(d_0))^2$ is the impact resolution in MC. The track IP resolution is *unfolded*, in both data and simulation, using an iterative deconvolution of the primary vertex resolution, as explained in [180]. The same calculation is performed for the z_0 impact parameter resolution.

Figure 10.1 shows the effect of some of these variations on the MV2c10 distribution alongside the nominal MC sample. This is shown for the release 20 analysis. Of all variations, smearing the impact parameters, d_0 and z_0, results in the largest deviation of the MV2c10 distribution from nominal.

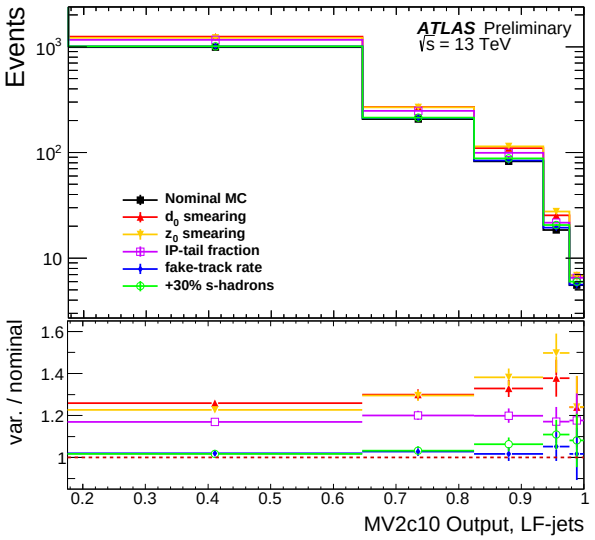


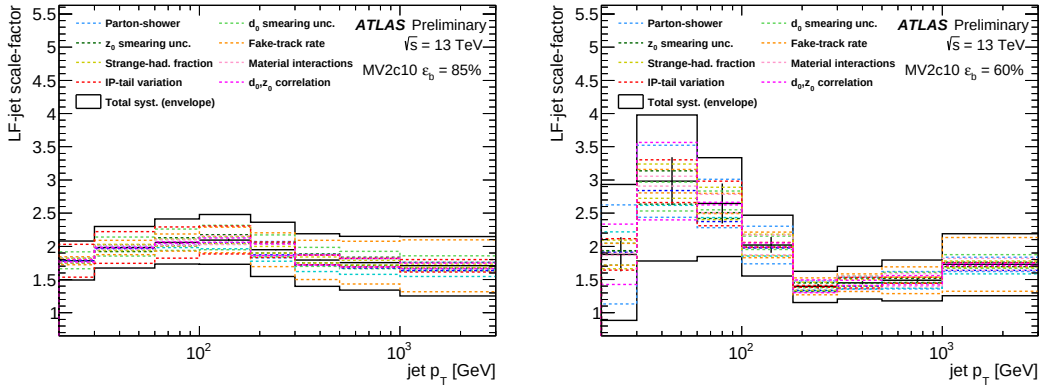
FIGURE 10.1: The MV2c10 distribution for light-jets for the nominal simulation (black) and for several variations: d_0 smearing (red), z_0 smearing (orange), impact parameter tail correction (violet), fake track rate variation (blue), and strange-hadron fraction variation (green). The bottom panel displays the ratios of the different variations with the nominal. Taken from [176].

For each of the variations, i, listed above, a light-jet mistagging efficiency, $\varepsilon_{\mathrm{LF}}^{\mathrm{adjusted},i}$, can be deduced. Assuming that each of the variations are uncorrelated, the calibration scale factor is given by

$$\begin{aligned}\mathrm{SF}_{\mathrm{LF}} = \prod_i \frac{\varepsilon_{\mathrm{LF}}^{\mathrm{adjusted},i}}{\varepsilon_{\mathrm{LF}}^{\mathrm{nominal}}} &= \prod_i \mathrm{SF}_{\mathrm{LF}}^{\mathrm{adjusted},i} \\ &= \mathrm{SF}_{\mathrm{LF}}^{d_0\text{-smearing}} \cdot \mathrm{SF}_{\mathrm{LF}}^{z_0\text{-smearing}} \cdot \mathrm{SF}_{\mathrm{LF}}^{\text{fake tracks}} \cdot \mathrm{SF}_{\mathrm{LF}}^{s\text{-hadrons}} \cdot \mathrm{SF}_{\mathrm{LF}}^{\text{IP tails}}.\end{aligned} \tag{10.3}$$

The uncertainties associated with this calculation are considered as such:

- The statistical uncertainty associated with the MC samples is considered in the calculation of the total systematic uncertainty.
- Uncertainties on the impact parameter resolution measurements originate from the unfolding process and from track-density effects. A further uncertainty is calculated from the difference between the impact parameter resolution measurement using di-jet events and that using $Z \to \mu\mu$ and minimum-bias events.
- The uncertainty arising from correlations between the d_0 and z_0 smearing is considered by comparing the results of separate d_0 and z_0 smearing with simultaneous d_0 and z_0 smearing.



(a) 85% efficiency WP (b) 60% efficiency WP

FIGURE 10.2: The calibration scale factors, calculated using the adjusted-MC method, for (a) the 85% efficiency working point and (b) the 60% efficiency working point, as a function of jet p_{T}. The scale factors and their statistical errors are represented by the black crosses, and the black continuous lines are the total systematic uncertainty (added in quadrature). The various systematic uncertainties are represented by dashed coloured lines. Taken from [176].

- Tests of the impact parameter tail correction using an alternative parametrisation of the d_0-distribution and a different track selection, and the analysis of the composition of the non-prompt track component in simulation, have shown consistent results. Regardless, a conservative uncertainty on the impact parameter tail correction is applied; this is equal to doubling or removing $\mathrm{SF}_{\mathrm{LF}}^{\mathrm{IP\ tails}}$.

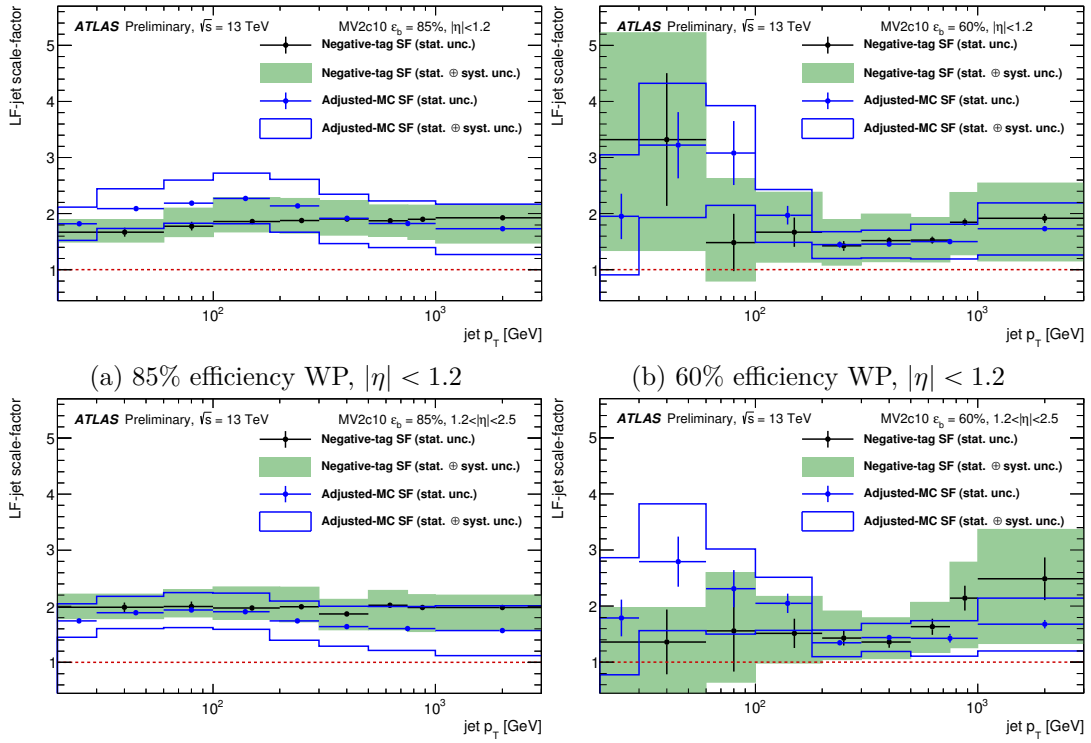
- The uncertainty on the rate of fake tracks is calculated from the difference in the measurements using the tight and loose track selections.

- To account for the uncertainty on the correction of the fraction of strange hadrons, a conservative uncertainty of $\pm 30\%$ is applied to $\mathrm{SF}_{\mathrm{LQ}}{}^{s\text{-hadrons}}$.

- The rate of interactions with the detector material in data and MC are in agreement within an uncertainty of 10%. This uncertainty is included as it has a large impact on the MV2c10 distribution of light-jets.

For each systematic variation, Equation 10.3 is re-evaluated and compared with the nominal $\mathrm{SF}_{\mathrm{LF}}$ to yield the associated uncertainty. The parton shower uncertainty is calculated by evaluating Equation 10.3 using an alternative MC sample produced using the HERWIG++ generator and comparing the results with the nominal calibration.

Figure 10.2 shows the calibration scale factors, calculated according to Equation 10.3, for the 85% and 60% working points. The contributions of each systematic uncertainty to the overall uncertainty are illustrated by the dashed lines.



(a) 85% efficiency WP, $|\eta| < 1.2$

(b) 60% efficiency WP, $|\eta| < 1.2$

(c) 85% efficiency WP, $1.2 < |\eta| < 2.5$

(d) 60% efficiency WP, $|\eta| < 1.2$

FIGURE 10.3: The calibration scale factors as a function of jet p_{T}, calculated using the adjusted-MC method (blue points with blue uncertainty bands) and the negative-tag method (black points with solid green uncertainty bands). Scale factors are shown for: (a) the 85% efficiency working point in the central region of the detector; (b) the 60% efficiency working point in the central region of the detector; (c) the 85% efficiency working point in the forward region of the detector; and (d) the 60% efficiency working point in the forward region of the detector. Taken from [176].

10.2.1 Comparison of Calibration Methods

The light-jet mistagging efficiencies obtained using the adjusted-MC and negative-tag methods are compared in Figure 10.3, separately for the central ($|\eta| < 1.2$) and forward ($1.2 < |\eta| < 2.5$) regions of the detector for the 85% and 60% efficiency working points. Again, these scale factors are performed using software release 20. The scale factors produced by the two methods are in good agreement within the systematic uncertainties. The systematics are of comparable size for the two methods, though for the 85% working point, the uncertainties derived for the negative-tag method are smaller than for the adjusted-MC method. For the 60% working point, the uncertainties associated with the results of the adjusted-MC method are smaller due to the larger b- and c-jet contamination in the negative-tag sample [176]. This is especially evident at high $p_{\mathrm{T}}^{\mathrm{jet}}$, where the negative-tag suffers from low statistics in the data sample. This region is important for searches for new physics involving high-p_{T} decays.

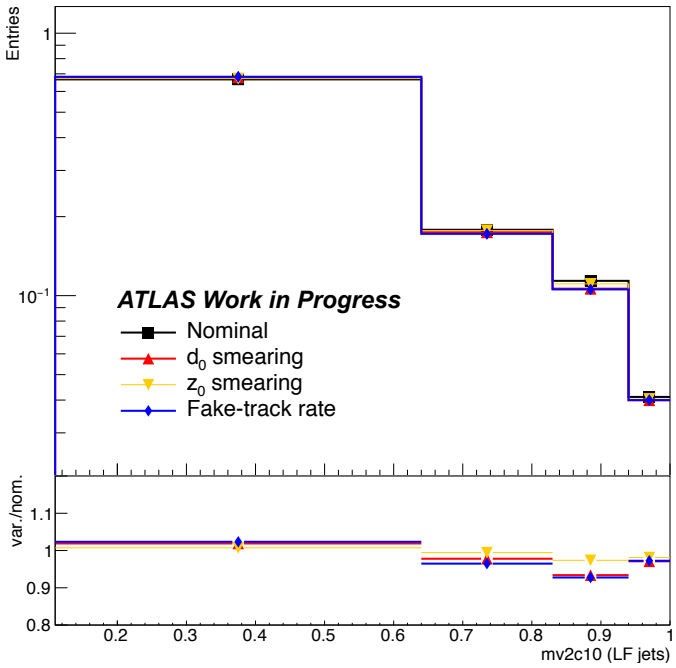


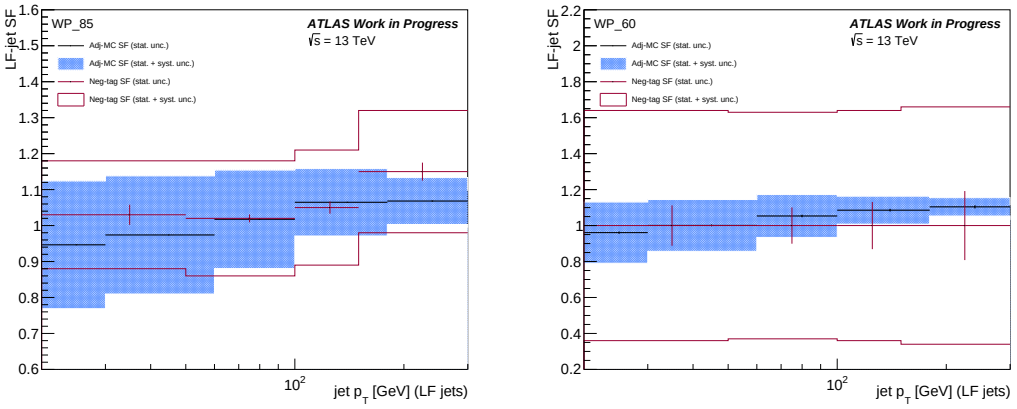
FIGURE 10.4: The distribution of MV2c10 for light-jets in the nominal sample and for each variation used in the calibration. The bottom panel shows the ratio of each variation to the nominal.

10.3 Adjusted-MC Calibration in ATLAS Software Release 21

For the 2017 and 2018 data-taking and the corresponding simulation, a number of improvements were made to the ATLAS event reconstruction and flavour tagging algorithms with respect to the 2016 environment [181]. The pixel clustering algorithm was improved and a more realistic charge deposition model was introduced for the reconstruction of tracks in the ID, leading to an improvement in the tracking performance. Furthermore, the JetFitter algorithm, which performs a topological decay reconstruction along the jet axis, was updated, providing improved light-jet rejection.

Figure 10.4 shows the MV2c10 distribution for light-jets for the nominal simulation and for several variations: d_0 smearing, z_0 smearing and the fake track rate variation. In comparison with Figure 10.1, the track impact parameter smearing (both d_0 and z_0) yields a far smaller variation of the MV2c10 distribution.

Figure 10.5 shows the preliminary light-jet calibration scale factors as a function of jet p_{T} for the 85% and 60% working points for the adjusted-MC and negative-tag methods. The adjusted-MC results are derived using variations of the track impact parameter resolution and probability of reconstructing fake tracks; samples for the other variations listed in Section 10.2 are not available at the time of writing. The systematic uncertainty is estimated by applying a 100% uncertainty to each variation. The scale factors are



(a) 85% efficiency WP (b) 60% efficiency WP

FIGURE 10.5: The calibration scale factors as a function of jet p_{T}, calculated using the adjusted-MC method (black points with blue filled uncertainty bands), compared with those calculated using the negative-tag method (dark pink points with dark pink uncertainty bands). These are shown for (a) the 85% working point and (b) 60% efficiency working point.

reduced compared to Figure 10.3 for both methods. This is due to the improved tracking performance and light-jet rejection.

Part III

Search for Resonant and Non-Resonant Higgs Boson Pair-Production Decaying to a $b\bar{b}\tau^+\tau^-$ Final State

Chapter 11

Trigger and Event Selection

The ATLAS trigger system, described in Section 11.1, is a two-level system consisting of a hardware-based trigger (Level 1) and a software-based trigger (Level 2) used to select events with characteristics which make them interesting for physics analysis. Events that pass the triggers selected for the analysis are subject to a number of further analysis-specific requirements to ensure that the particles involved and their measurements are consistent with the final state of the signal process.

11.1 Trigger and Data Cleaning

The $\tau_{\text{lep}}\tau_{\text{had}}$ analysis uses events that pass one of the single lepton triggers (SLTs) or, should they fail the SLTs, the lepton-plus-τ_{had} triggers (LTT). Events are analysed separately based on the type of trigger they are selected by. The contribution of the author focused on analysing the SLT events; therefore, only the SLTs are outlined here.

Several different SLTs are used, all of which use the lowest p_{T} thresholds for each run period and are unprescaled,[1] in order to maximise acceptance. The SLTs are divided into electron and muon channels. Events with an electron are accepted if they pass one of the following requirements:

- at least one electron with $p_{\text{T}} > 24$ GeV satisfying the *Medium* identification criteria and *Loose* isolation requirements (in later data-taking periods the p_{T} threshold is raised to 26 GeV and identification changed to *Tight*), or
- at least one electron with $p_{\text{T}} > 60$ GeV satisfying the *Medium* identification criteria but with no isolation requirement, or

[1] Prescaling reduces the trigger rates by randomly vetoing events that pass the trigger.

- at least one electron with $p_{\mathrm{T}} > 120 - 140$ GeV, depending on the data-taking period, satisfying the *Loose* identification criteria.

In the muon channel, events are required to have:

- at least one muon with $p_{\mathrm{T}} > 24 - 26$ GeV, depending on the data-taking period, satisfying the *Loose* isolation requirements, or
- at least one muon with $p_{\mathrm{T}} > 50$ without any other requirements.

The trigger requirements vary with the data-taking period; the p_{T} threshold and identification requirements were increased to keep the rate of data consistent whilst the instantaneous luminosity increased. The triggers and their corresponding requirements are summarised in Table 11.1.

All electrons and muons reconstructed offline are required to be matched to the trigger objects. Events are further required to contain at least one primary vertex, where a primary vertex is defined as a point of proton-proton collision with at least two associated tracks, each with $p_{\mathrm{T}} > 400$ MeV [182]. Should more than one primary vertex be present in the event, the one with the largest $(p_{\mathrm{T}}^{\mathrm{track}})^2$ is chosen.

11.2 Event Selection

An *event selection* is applied to all events that pass the trigger selection in order to select a final state containing an electron or muon, ℓ, a hadronic τ-lepton, at least two jets and missing transverse momentum.

Events are subject to the following requirements (using the object definitions described in Chapter 9):

- exactly one electron passing the *Tight* identification criteria or exactly one muon passing the *Medium* identification criteria (with the additional requirement $|\eta^{\mu}| < 2.5$), with p_{T} required to be 1 GeV higher than the trigger threshold corresponding to the data-taking period of the event, such that the trigger efficiency reaches the plateau where the MC modelling improves;
- exactly one hadronic τ-lepton with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.3$;
- at least two jets, where the leading (subleading) jet fulfils $p_{\mathrm{T}} > 45$ (20) GeV;
- no other electrons or muons (as defined in Sections 9.1 and 9.2) in the event;

Lepton Type	Name	p_{T} Threshold [GeV]	Identification	Isolation
2015 Data				
Electron	HLT_e24_lhmedium_iloose_L1EM20VH	24	*Medium*	*Loose*
	HLT_e60_lhmedium	60	*Medium*	—
	HLT_e120_lhloose	120	*Loose*	—
Muon	HLT_mu24_imedium	24	*Medium*	—
	HLT_mu50	50	—	—
2016 Period A-D3				
Electron	HLT_e24_lhmedium_ivarloose	24	*Medium*	*Loose*
	HLT_e60_lhmedium_nod0	60	*Medium*	—
	HLT_e140_lhloose_nod0	140	*Loose*	—
Muon	HLT_mu24_imedium	24	*Medium*	—
	HLT_mu50	50	—	—
2016 Period D4 - end of 2016				
Electron	HLT_e26_lhtight_ivarloose	26	*Tight*	*Loose*
	HLT_e60_lhmedium_nod0	60	*Medium*	—
	HLT_e140_lhloose_nod0	140	*Loose*	—
Muon	HLT_mu26_imedium	26	*Medium*	—
	HLT_mu50	50	—	—

TABLE 11.1: The SLTs used for each data-taking period and their p_{T}, identification and isolation requirements. The naming convention for each trigger is as follows: 'HLT' (indicating a High Level Trigger), followed by the lepton type ('e' or 'mu' for an electron or muon respectively), the p_{T} threshold, the identification working point (prefixed with 'lh' to indicate a likelihood-based trigger), and finally the isolation working point (prefixed with 'i'). The 'nod0' suffix indicates that no transverse impact parameter (d_0) cuts are required and 'ivarloose' refers to a variable-size cone isolation requirement. A data-taking *period* refers to a set of data acquired under similar operating conditions.

- the light lepton and hadronic τ-lepton must have opposite-sign electric charge;
- the invariant mass of the di-τ system, calculated using the MMC (as detailed in Section 9.7), $m_{\tau\tau}^{\text{MMC}} > 60$ GeV.

Following the above selection, events are categorised according to the multiplicity of b-tagged jets in the event. The signal region requires exactly two b-tagged jets, whilst events with zero or one b-tags are used for validating the background modelling (events with zero b-tags use the two jets with highest p_{T} and events with one b-tag consider the b-tagged jet and the highest-p_{T} non-tagged jet).

Chapter 12

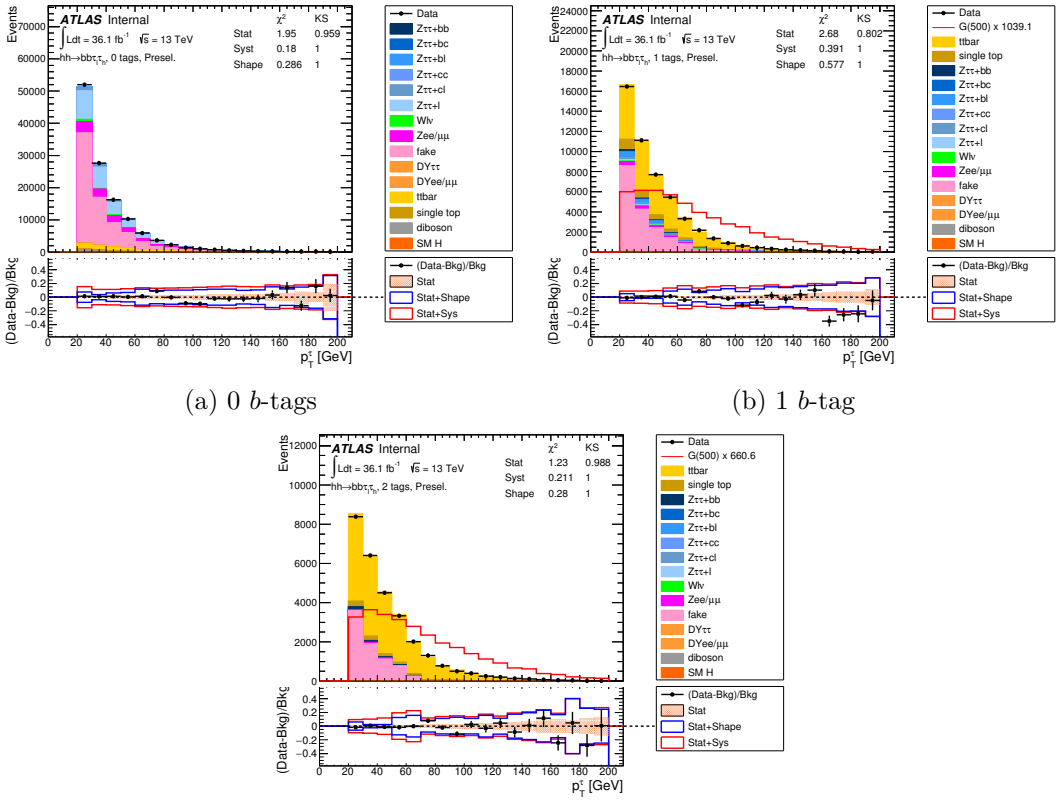
Background Estimation

This chapter summarises the methods used to estimate the major background processes and their associated systematic uncertainties. Background processes are any processes that have similar final state kinematics to those expected for the signal process. The presence of a signal is inferred from an excess of events above the predicted background; therefore, the background processes must be carefully modelled with suitable uncertainties applied.

Figure 12.1 is the transverse momentum of the hadronically-decaying τ-lepton, $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$, for all of the relevant background processes, in the zero, one and two b-tag regions. In all three regions the background distribution is consistent with the data, which are shown as black points with error bars.

The background processes considered in both analyses are listed below.

- The dominant background process is the pair-production of top-quarks, referred to as $t\bar{t}$ production, which constitutes over 90% of the total background in the two b-tag signal region. Any $t\bar{t}$ processes that include a real hadronically decaying τ-lepton are modelled using MC simulation, as described in Section 12.1, and are labelled 'ttbar' in Figure 12.1.
- Processes where a jet is misidentified as a hadronically decaying τ-lepton are not well modelled in MC simulation; therefore, they are modelled using a data-driven method and are shown as a single background labelled 'fake' in Figure 12.1 (see Section 12.2). These processes are $t\bar{t}$, W + jets, and QCD multi-jet.
- Single top production, which includes the s- and t-channels and Wt associated production, is labelled as 'single top' in Figure 12.1.



(a) 0 b-tags

(b) 1 b-tag

(c) 2 b-tags

FIGURE 12.1: Distribution of the transverse mass of the hadronically-decaying τ-lepton, $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$, for all of the relevant background processes, in the (a) the zero b-tag region, (b) the one b-tag region and (c) the two b-tag region. The RS graviton signal with $m_G = 500$ GeV is shown as a red line. The bottom pane represents the deviation of the data from the background estimation, with the red envelope representing the statistical and systematic uncertainties.

- $Z \to \tau\tau$ processes, where a Z boson, which is produced in association with any combination of b-, c- and light-flavour jets, decays to a pair of τ-leptons, are labelled as 'Z → ττ+ bb, bc, bl, cc, cl, l' in Figure 12.1 (see Section 12.3).

- Processes where a W boson, which is produced in association with jets, decays to a lepton and neutrino, are referred to as 'Wlν' in Figure 12.1.

- $Z \to \ell\ell$ processes where a Z boson, produced in association with jets, decays to a pair of light leptons, are labelled 'Zee/μμ' in Figure 12.1 (see Section 12.4). The $Z \to \mu\mu$ process is negligible because a muon is unlikely to deposit enough energy in the calorimeters to be misidentified as a τ_{had}.

- Drell-Yan processes where a pair of leptons are pair-produced, are included in Figure 12.1 as 'DYττ' and 'DYee/μμ'.

- Di-boson processes involving combinations of two Z and W bosons, i.e. WW, ZW and WW, are referred to in Figure 12.1 as 'diboson'.

- Processes involving a single SM Higgs boson, including the associated production of a Z boson or $t\bar{t}$ pair and a Higgs (referred to as Zh or tth production), are shown as a single background labelled 'SM H' in Figure 12.1.

Background processes where a lepton is misidentified as a τ_{had}, which account for less than 5% of the total background, are estimated using MC simulation. All backgrounds that are not covered in the following sections are modelled using MC simulation, as outlined in Chapter 8.

12.1 Top-quark pair-production decaying to a real τ_{had}

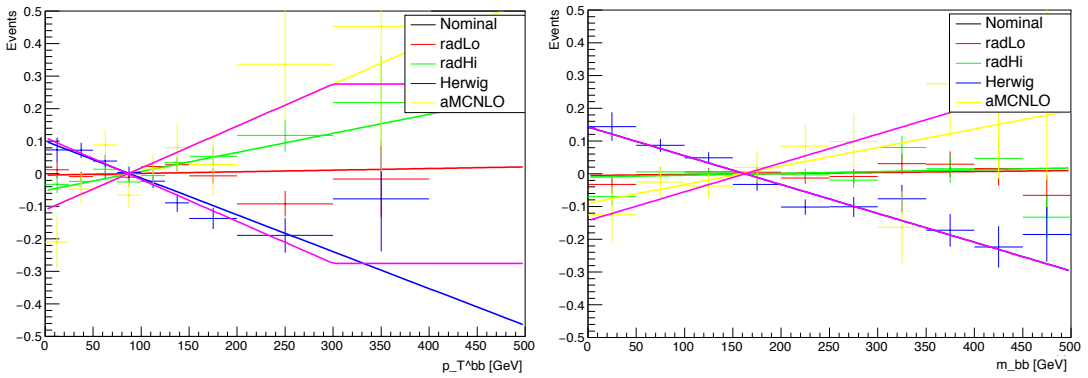
As described in Chapter 8, the MC simulation for the $t\bar{t}$ process is generated using Powheg+Pythia6. This is then normalised to data; the normalisation is allowed to vary in the final fit (which is described in Chapter 5), constrained by the low-score region of the BDT distribution (the final discriminant). For the background-only fit, this gives an overall normalisation factor of 1.06 ± 0.13. For each signal hypothesis, a different fit is performed and a different value of the normalisation is obtained (though all normalisation factors are consistent within their uncertainties).

12.1.1 Uncertainties on top-quark processes with a real τ_{had}

Generator-level uncertainties, including those associated with the shower radiation and hadronisation model, are applied to the real $t\bar{t}$ processes.

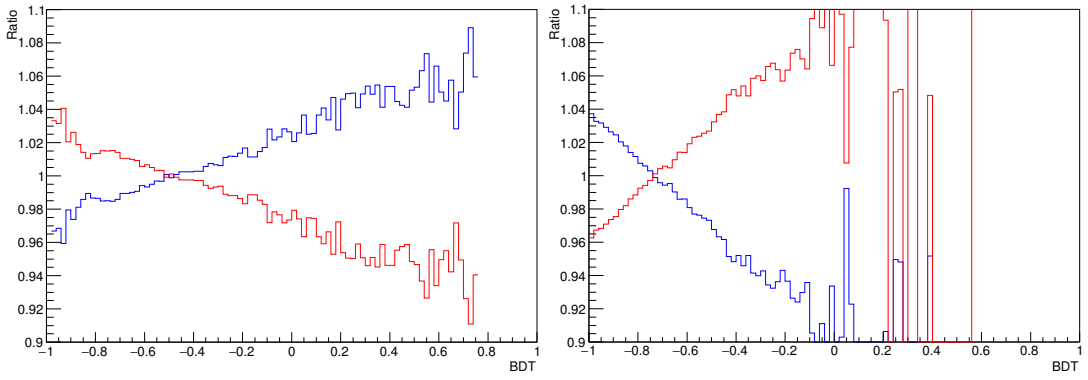
The uncertainty on the $t\bar{t}$ normalisation is calculated by applying the following variations:

- the Powheg+Pythia6 factorisation and renormalisation scales are varied by factors of one half (two) and low (high) radiation variations of the Perugia 2012 tune are used, corresponding to less (more) parton shower radiation;
- the Powheg+Pythia6 parton shower fragmentation model is compared with Powheg+Herwig++;
- the hard scatter simulation, generated with Powheg-Box and showered using Herwig++, is compared with a sample generated with aMC@NLO and showered using Herwig++.



(a) p_{T}^{bb} (b) m_{bb}

FIGURE 12.2: Shape variations of the $t\bar{t}$ background as a function of (a) p_{T}^{bb} and (b) m_{bb}. 'radHi' and 'radLo' are variations resulting in high and low radiation, respectively; 'Herwig' refers to the variation of the fragmentation model from POWHEG+PYTHIA6 to POWHEG+HERWIG; and 'aMCNLO' is the variation of the hard scatter simulation from that generated with POWHEG and showered using HERWIG++ to a sample generated with aMC@NLO. The overall uncertainty is shown by the purple line.



(a) $m_G = 260$ GeV (b) $m_G = 900$ GeV

FIGURE 12.3: Shape variations of the BDT distribution when applying the $t\bar{t}$ systematics as functions of p_{T}^{bb} and m_{bb} for an RS graviton signal with (a) $m_G = 260$ GeV and (b) $m_G = 900$ GeV.

The discrepancies between the nominal sample and these variations are parametrised as functions of p_{T}^{bb} and m_{bb}, as shown in Figure 12.2. These variables were chosen as they exhibited the largest variations. The overall uncertainty is a symmetrised envelope of the individual effects; in p_{T}^{bb} the largest effect is the parton shower model variation and in m_{bb} it is the hard scatter variation. The uncertainties in p_{T}^{bb} and m_{bb} are then propagated to the BDT distribution to give an overall shape systematic as a function of the BDT output variable. Figure 12.3 shows how the BDT distribution varies from the nominal when applying the systematics in both p_{T}^{bb} and m_{bb} for an RS graviton signal with (a) $m_G = 260$ GeV and (b) $m_G = 900$ GeV.

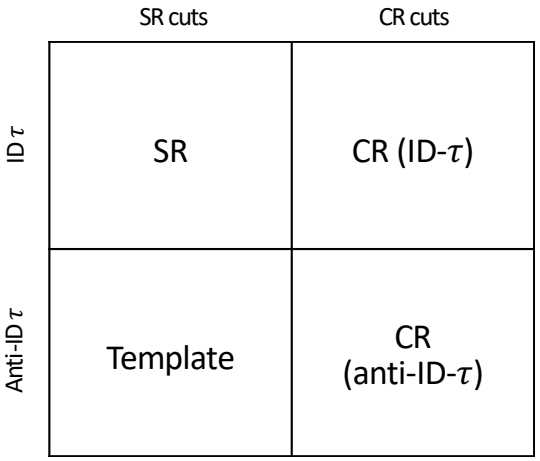


FIGURE 12.4: Diagram of the four regions relevant for the fake factor calculation. The fake factor is calculated in the control region (CR), which is defined separately for each background process, by taking the ratio of the number of events in the ID-τ_{had} region to the number of events in the anti-ID-τ_{had} region. This is then applied to the template region in order to estimate the signal region (SR).

12.2 Backgrounds where a jet is misidentified as a hadronic τ-lepton

Background processes where a jet is misidentified as a τ_{had} (also referred to as fake-τ_{had} processes) are estimated using a data-driven technique due to the poor reliability of MC modelling.[1] In the two b-tag signal region, the dominant fake-τ_{had} source is $t\bar{t}$ production, with small contributions from QCD multi-jet and W + jets processes. The fake-τ_{had} contribution in the one b-tag region comprises an approximately equal split of $t\bar{t}$ and W + jets events, while the majority of the zero b-tag fake-τ_{had} background comes from W + jets.

In order to estimate the number of events in the signal region where a jet fakes a τ_{had}, a *template region*, enriched in fake-τ_{had} events, is constructed and its events weighted by a *fake factor*. The template region, shown on the bottom left of Figure 12.4, is formed by applying the signal region selection, but requiring the event to pass the anti-ID selection defined in Section 9.5.3, i.e. requiring the event to include one τ_{had}-candidate which has failed the *Medium* τ_{had} identification requirement but has a τ-identification BDT score greater than 0.35.

The fake factor is the ratio of the number of fake-τ_{had} events in the signal region to the number of fake-τ_{had} events in the template region; this calculation is performed using events in the *control region*, which is shown on the right-hand-side of Figure 12.4. Three separate control regions are defined, each enriched in one of the three background

[1] The fake-τ_{had} contribution to the $t\bar{t}$ background is modelled using simulation when training the BDTs and using the data-driven method elsewhere.

processes, such that a separate fake factor can be calculated for each process. This is because the fake-τ_{had} composition is different for each background process; W + jets and $t\bar{t}$ processes are dominated by quark-initiated jets and QCD multi-jet processes are dominated by gluon-initiated jets. The control regions are:

- **QCD multi-jet:** The isolation criteria for electrons and muons are inverted, i.e. electrons and muons are required to fail the *Loose* isolation working points. The fake factor is calculated separately for events with zero and one b-tagged jet(s), though these are similar. The fake factor calculated in the one b-tag region is used for events with two b-tagged jets, due to the lack of events in the two b-tag region.
- ***tt̄*:** Events are required to have $m_{\text{T}}^{W} > 40$ GeV and two b-tagged jets.[2]
- ***W*+ jets:** Events are required to have $m_{\text{T}}^{W} > 40$ GeV and exactly zero b-tagged jets.

Here, m_{T}^{W} is defined as the transverse mass of the light lepton and $E_{\text{T}}^{\text{miss}}$, labelled as such because, for the $t\bar{t}$ and W + jets backgrounds, it corresponds to the transverse mass of the W boson. It is defined as

$$m_{\text{T}}^{W} = \sqrt{2p_{\text{T}}^{\ell}E_{\text{T}}^{\text{miss}}(1 - \cos\Delta\phi)}, \tag{12.1}$$

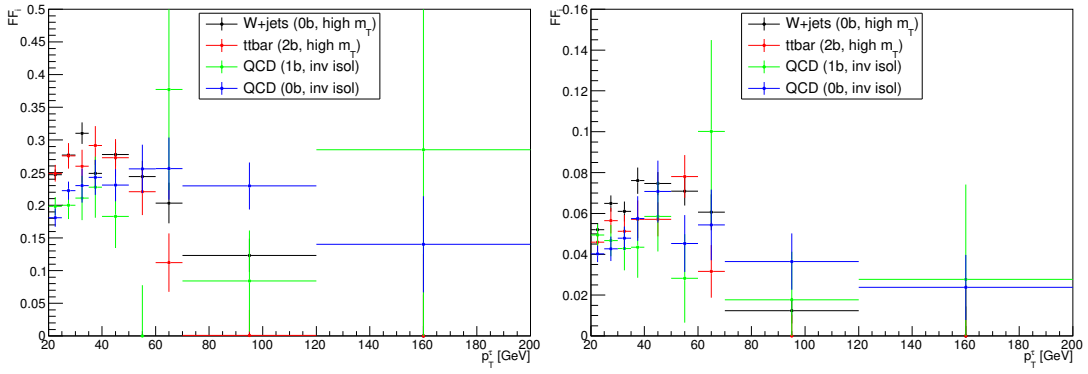
where p_{T}^{ℓ} is the transverse momentum of the lepton and $\Delta\phi$ is the angle between the lepton and $E_{\text{T}}^{\text{miss}}$. For the $t\bar{t}$ and W + jets backgrounds, a Jacobian edge in the m_{T}^{W} distribution is expected at the W boson mass due to the presence of a lepton and neutrino from the decay of a W boson.

The fake factor, FF_i, for each background process, i, is calculated as

$$\text{FF}_i = \frac{N_{\text{ID}}^{\text{CR},i}}{N_{\text{anti-ID}}^{\text{CR},i}}, \tag{12.2}$$

where $N_{\text{ID}}^{\text{CR}}$ refers to the number of events in the background-specific control region (shown on the top right of Figure 12.4), and $N_{\text{anti-ID}}^{\text{CR}}$ is the number of events passing the control region requirements in a region enriched in fake-τ_{had} decays (shown on the bottom right of Figure 12.4); both numbers have the number of events from non-fake sources and fakes from other minor sources (those that are not estimated using the data-driven approach, such as Z + jets) subtracted using MC-simulated events. If more than

[2] Whilst this control region is a subset of the signal region, the events in the high-m_{T}^{W} region are expected to be at low BDT score (see Chapter 13) and, therefore, do not have a large influence.



(a) One-prong (b) Three-prong

FIGURE 12.5: The fake factors, FF_i, for the QCD multi-jet, $t\bar{t}$ and $W+$jets processes, as a function of $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ for (a) one-prong τ_{had} decays and (b) three-prong τ_{had} decays. The fake factor for QCD multi-jet is calculated separately for events with zero and one b-tagged jets (both are shown in the figure) in a control region requiring the electron and muon isolation criteria to be inverted. The $W+$jets and $t\bar{t}$ fake factors are calculated in the zero and two b-tag regions, respectively, with the additional requirement of $m_{\mathrm{T}}^{W} > 40$ GeV.

one anti-ID-τ_{had} is present in the event, one is chosen at random. The fake factors are parametrised as a function of $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ and are calculated separately for one- and three-prong τ_{had} decays (see Section 2.4.1.1 for an explanation of τ-lepton decays).

The individual fake factors calculated for the QCD multi-jet, $t\bar{t}$ and $W+$jets processes (labelled $\mathrm{FF}_{\mathrm{QCD}}$, $\mathrm{FF}_{t\bar{t}}$ and FF_{W} respectively) are shown in Figure 12.5 for both one- and three-prong τ_{had} decays. The $t\bar{t}$ and $W+$jets fake factors are generally consistent within their statistical errors and the QCD multi-jet fake factors are similar for zero and one b-tags. The $t\bar{t}$ and $W+$jets fake factors are larger than those for QCD multi-jet processes; this is due to the higher fraction of quark-initiated jets in $t\bar{t}$ and $W+$jets processes than in QCD multi-jet processes. Quark-initiated jets are typically more narrow and produce fewer hadrons than gluon-initiated jets, and are thus more likely to pass the τ_{had} identification. Furthermore, the fake factors are larger (for all three background sources) for one-prong events than for three-prong events because a one-prong fake τ-lepton candidate is more likely to pass the τ-lepton identification than a three-prong candidate.

The three fake factors are combined to form the *combined fake factor*, $\mathrm{FF}_{\mathrm{comb}}$, which is defined as

$$\mathrm{FF}_{\mathrm{comb}} = \mathrm{FF}_{\mathrm{QCD}} \times r_{\mathrm{QCD}} + \mathrm{FF}_{t\bar{t}} \times r_{t\bar{t}} \times (1 - r_{\mathrm{QCD}}) + \mathrm{FF}_{W} \times (1 - r_{t\bar{t}}) \times (1 - r_{\mathrm{QCD}}), \quad (12.3)$$

where r_{QCD} is the proportion of the fake-τ_{had} background for which QCD multi-jet processes are responsible, and $r_{t\bar{t}}$ is the proportion of the remaining background ($W+$jets

and $t\bar{t}$ events) that is made up of $t\bar{t}$ events.

Because the zero b-tag region is dominated by the W + jets background, the $t\bar{t}$ contribution is treated as negligible, i.e. $r_{t\bar{t}} = 0$. In the two b-tag region, the $t\bar{t}$ background is dominant and the W + jets contribution is treated as negligible, i.e. $r_{t\bar{t}} = 1$. In the one b-tag region, separate fake factors are derived for the W + jets and $t\bar{t}$ backgrounds and the size of their contributions relative to one another are determined using MC predictions. Because the fake factors for $t\bar{t}$ and W + jets are compatible within their statistical errors (as shown in Figure 12.5), the sizes of the backgrounds relative to one another are inconsequential.

The ratio r_{QCD} is defined as the fraction of fake-τ_{had} events in the template region that are from QCD multi-jet processes. This is calculated, again as a function of $p_{\text{T}}^{\tau_{\text{had}}}$, according to

$$r_{\text{QCD}} = \frac{N_{\text{QCD}}^{\text{data}}}{N^{\text{data}} - N_{\text{true-}\tau_{\text{had}}}^{\text{MC}}}, \tag{12.4}$$

where the numerator here, the number of multi-jet events in data, is given by

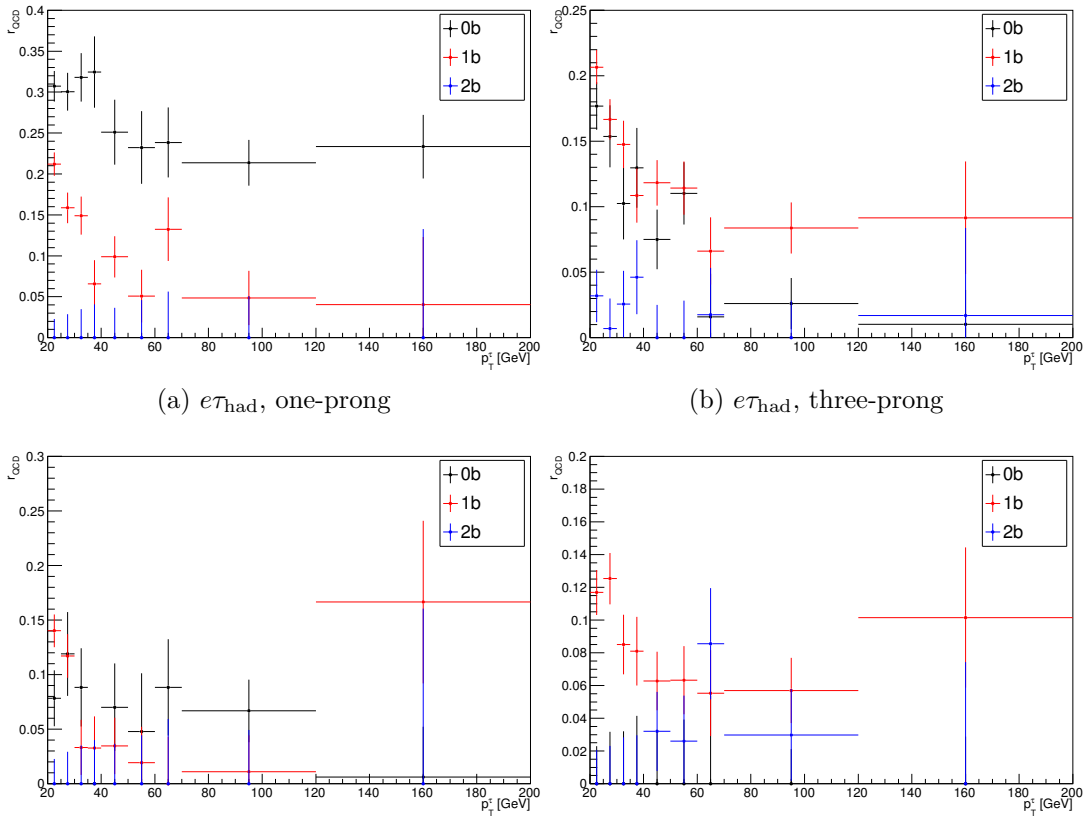
$$N_{\text{QCD}}^{\text{data}} = N^{\text{data}} - N_{\text{true-}\tau_{\text{had}}}^{\text{MC}} - N_{\text{fake-}\tau_{\text{had}}}^{\text{MC}}, \tag{12.5}$$

i.e. the total number of MC-simulated events, including both true- and fake-τ_{had} contributions (i.e. $t\bar{t}$ and W + jets), is subtracted from the number of events in the full data sample in the anti-ID-τ_{had} region (QCD multi-jet is the only non-negligible background for which there is no MC simulation). The ratio r_{QCD} is calculated separately according to the number of b-tagged jets, as well as the type of light lepton to which the τ_{lep} decays (e/μ) and the number of prongs in the τ_{had} decay (one or three).

The values of r_{QCD} are shown in Figure 12.6 for each decay channel. As expected, for one-prong τ_{had} decays, r_{QCD} decreases as the number of b-tagged jets in the event increases. This trend is not observed in the three-prong r_{QCD} distribution but this can be attributed to the lower numbers of three-prong decays. For all four decay channels, the two b-tag r_{QCD} values are consistent with zero. This is because the $t\bar{t}$ background is dominant.

The combined fake factor is used to achieve the final background sample:

$$N_{\text{fake-}\tau_{\text{had}}}^{\text{estimated, SR}} = \left(N_{\text{anti-ID}}^{\text{data, SR}} - N_{\text{anti-ID}}^{\text{true-}\tau_{\text{had}}\text{ MC, SR}} \right) \times \text{FF}_{\text{comb}}, \tag{12.6}$$



(a) $e\tau_{\text{had}}$, one-prong (b) $e\tau_{\text{had}}$, three-prong

(c) $\mu\tau_{\text{had}}$, one-prong (d) $\mu\tau_{\text{had}}$, three-prong

FIGURE 12.6: The values of r_{QCD} as a function of $p_{\text{T}}^{\tau_{\text{had}}}$ for (a) the $e\tau_{\text{had}}$ channel with a one-prong τ_{had} decay, (b) the $e\tau_{\text{had}}$ channel with a three-prong τ_{had} decay, (c) the $\mu\tau_{\text{had}}$ channel with a one-prong τ_{had} decay and (d) the $\mu\tau_{\text{had}}$ channel with a three-prong τ_{had} decay. In each plot, r_{QCD} is shown for the regions with zero, one and two b-tagged jets; these are labelled '0b', '1b' and '2b' respectively.

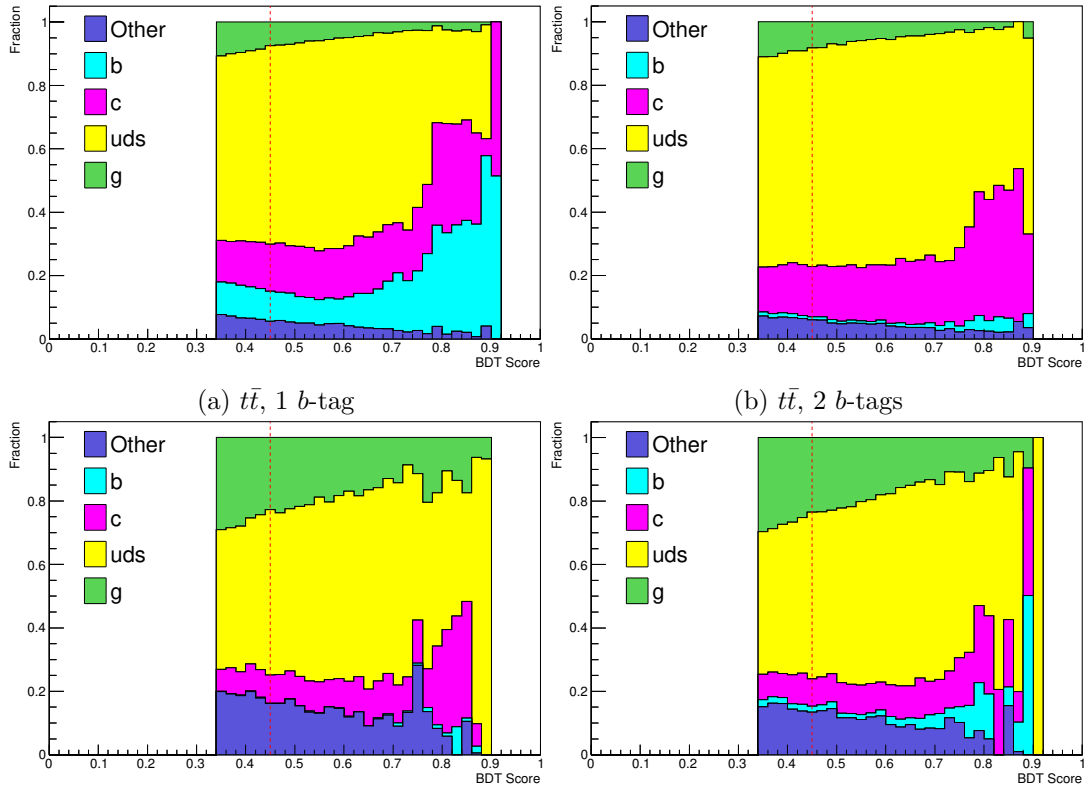
i.e. all MC-simulated true-τ_{had} events are subtracted from the data in the template region, which is similar to the signal region but is enriched in fake τ-leptons, and this is multiplied by the fake factor. The template region is constructed from data events that pass the preselection requirements but include an anti-ID-τ_{had}.

The fake-τ_{had} background modelling is validated in the two b-tag same-sign region, as shown in Figure 12.14 in Section 12.6.

12.2.1 Uncertainties on backgrounds where a jet fakes a τ_{had}

12.2.1.1 Quark and gluon composition variation

A systematic uncertainty is required to account for the difference in the quark and gluon flavour composition of jets misidentified as hadronically decaying τ-leptons in the signal region and the fake-τ_{had} enriched region (the template region - see Section 12.2).



(a) $t\bar{t}$, 1 b-tag

(b) $t\bar{t}$, 2 b-tags

(c) W + jets, 0 b-tags

(d) W + jets, 1 b-tag

FIGURE 12.7: The composition of jets misidentified as hadronically-decaying τ-leptons as a function of the τ-identification BDT score for (a) $t\bar{t}$ events with one b-tagged jet, (b) $t\bar{t}$ events with two b-tagged jets, (c) W + jets events with no b-tagged jets and (d) W + jets events with one b-tagged jet. The plots show the fraction of events in which the fake τ_{had} originates from each of the following: a gluon ('g'), a light quark ('uds'), a c-quark ('c'), a b-quark ('b') or otherwise ('other', mainly pile-up jets). The red line marks the composition of jets at a τ-identification BDT score of 0.45, which is approximately (it is p_{T}-dependent) the minimum value required for a jet to be positively identified as a τ_{had}.

Figure 12.7 shows the fraction of MC-simulated fake-τ_{had} events where the fake-τ_{had} originates from gluons, light-flavour quarks, c-quarks and b-quarks, as a function of the τ-identification BDT score. The origin of these jets varies significantly with the τ-identification BDT score.

The uncertainty is estimated in the two b-tag same-sign validation region, which requires the event to contain a τ_{had} and lepton with same-sign electric charge, plus two b-tagged jets.

The fake-τ_{had} composition in the two b-tag signal region is studied as a function of all input variables used in the analysis BDT (listed in Section 13.1). The greatest variation is observed as a function of m_{hh}; this is shown in Figure 12.8a. Figure 12.8b shows the fake-τ_{had} composition as a function of m_{hh} in the two b-tag same-sign validation region, which exhibits a similar composition to the opposite-sign (signal) region.

The non-closure in the number of fake-τ_{had} events predicted by the fake factor method in the two b-tag same-sign validation region is also most evident as a function of m_{hh}. This is demonstrated in Figure 12.8c, which shows the ratio of the number of fake-τ_{had} events in data (i.e. the number of MC-simulated non-fake events, $N^{\text{MC}}_{\text{true-}\tau_{\text{had}}}$, subtracted from the total number of data events, N^{data}) to the number of fake-τ_{had} events estimated using the fake factor method, $N^{\text{estimated}}_{\text{true-}\tau_{\text{had}}}$:

$$\frac{N^{\text{data}} - N^{\text{MC}}_{\text{true-}\tau_{\text{had}}}}{N^{\text{estimated}}_{\text{true-}\tau_{\text{had}}}} - 1, \tag{12.7}$$

where unity is subtracted from the ratio so that the resulting values are centred around zero.

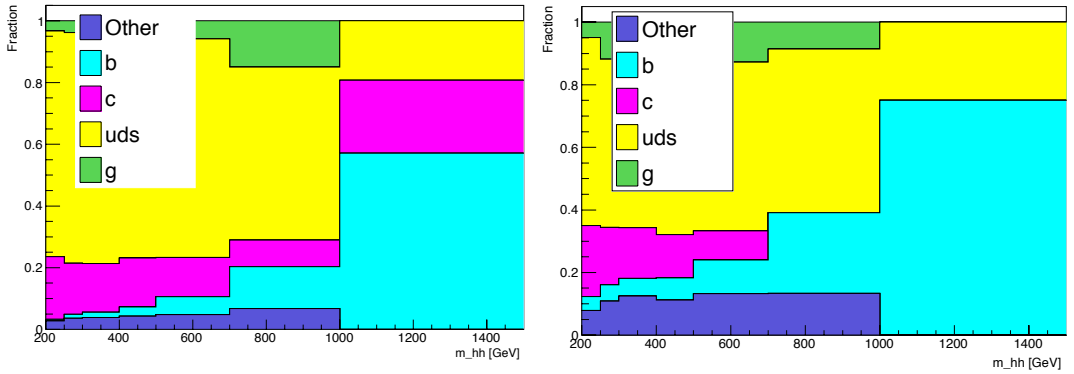
Because m_{hh} exhibits the greatest variation in composition and the greatest non-closure in the same-sign validation region, the uncertainty is parametrised as a function of m_{hh} and symmetrised to provide both the up and down uncertainties. This is illustrated by the green line in Figure 12.8c. The up and down uncertainties are propagated to the analysis BDT discriminant.

12.2.1.2 True-τ_{had} contamination of template region

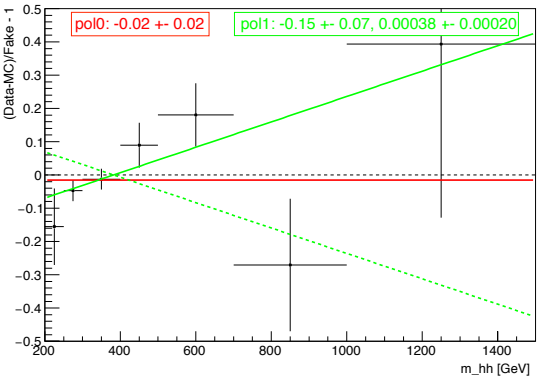
The contamination of the template region from the true-τ_{had} component is varied up and down by propagating the detector-related (see Section 14.1) and theoretical (see Section 12.1.1) uncertainties on the true-τ_{had} $t\bar{t}$ MC sample which is subtracted from data. Smaller backgrounds that are subtracted from data (single top, electroweak, di-boson and SM Higgs) are subject to a conservative variation of $\pm 50\%$.

Figures 12.9a and 12.9b show the estimated shape variation of the true-τ_{had} $t\bar{t}$ MC sample due to detector-related uncertainties as a function of p^{τ}_{T}, using the same binning as for the fake factors. The difference between the resulting distribution of the fake-τ_{had} background and the nominal fake-τ_{had} estimation is symmetrised and used as a shape uncertainty. Any detector-related uncertainties that lead to negligible variations are discounted.

Figures 12.9c and 12.9d show the variation of the true-τ_{had} $t\bar{t}$ MC sample due to theoretical uncertainties. These are calculated using the additional MC samples listed in Section 12.1.1. Again, the variations are propagated through every step of the fake factor method and symmetrised to give an uncertainty on the fake-τ_{had} background estimation. Statistical uncertainties are estimated by varying the fake factors within their statistical



(a) Fake-τ_{had} composition in signal region (b) Fake-τ_{had} composition in same-sign region



(c) Non-closure of fake factor method in same-sign validation region

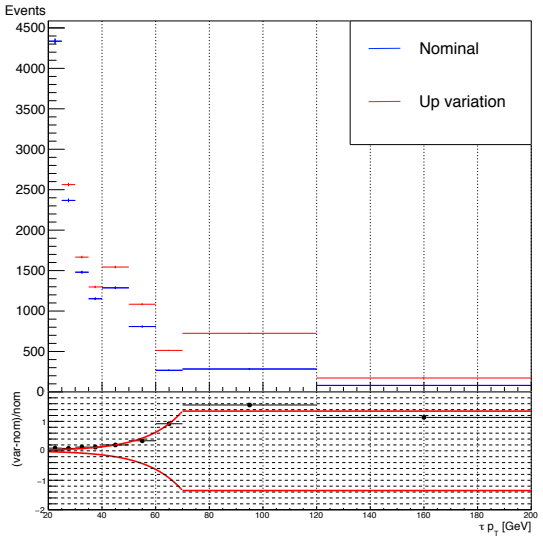
FIGURE 12.8: Composition of jets misidentified as hadronic τ-leptons as a function of m_{hh} in (a) the two b-tag opposite-sign signal region and (b) the two b-tag same-sign validation region. The ratio of data to simulation as a function of m_{hh} in the two b-tag same-sign region is shown in (c). In (c), the ratio is fitted and symmetrised to provide up and down systematic variations (as illustrated by the green line).

errors; in the case of the MC generator variations, the statistical uncertainties dominate and no systematic uncertainty is applied.

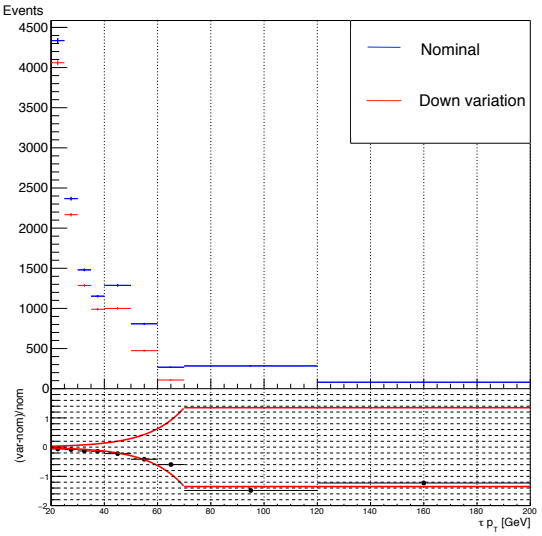
12.2.1.3 Extrapolation of fake factors from control regions to signal region

A systematic uncertainty is applied to account for the extrapolation of the fake factors from the high-m_{T} control regions to the signal region. The $t\bar{t}$ and W + jets fake factors are reproduced using MC simulations in the signal region and in the high-m_{T} region ($m_{\text{T}} > 40$ GeV) in order to study the difference between the fake factors in the two regions.

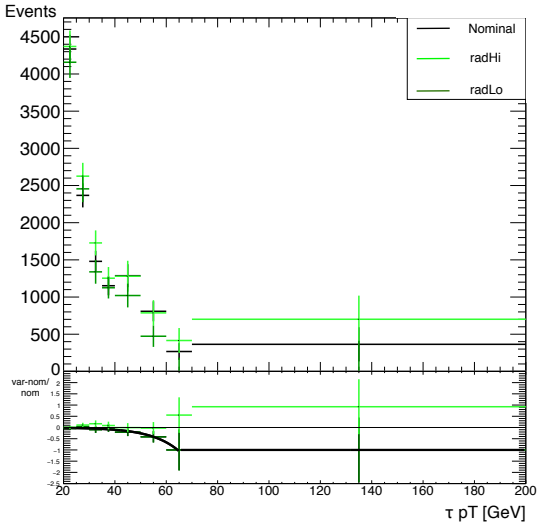
The fake factors for the two regions are shown in Figure 12.10 for one- and three-prong decays. The difference between the two fake factors is applied as up and down variations to the nominal data-derived fake factor, and the $t\bar{t}$ background estimate is rederived for each variation (r_{QCD} is also rederived). The difference between the background estimates



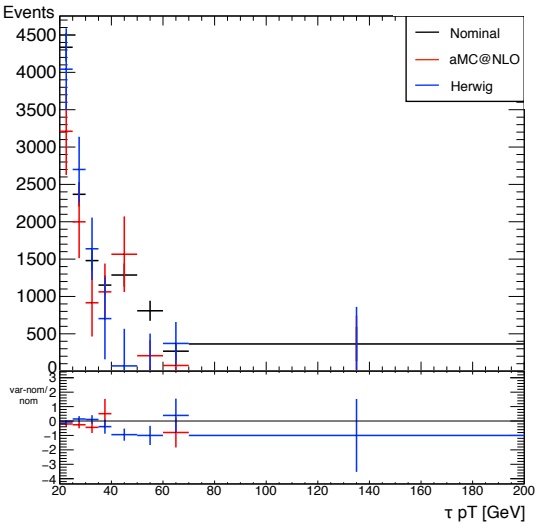
(a) Detector uncertainties 'up' variation



(b) Detector uncertainties 'down' variation

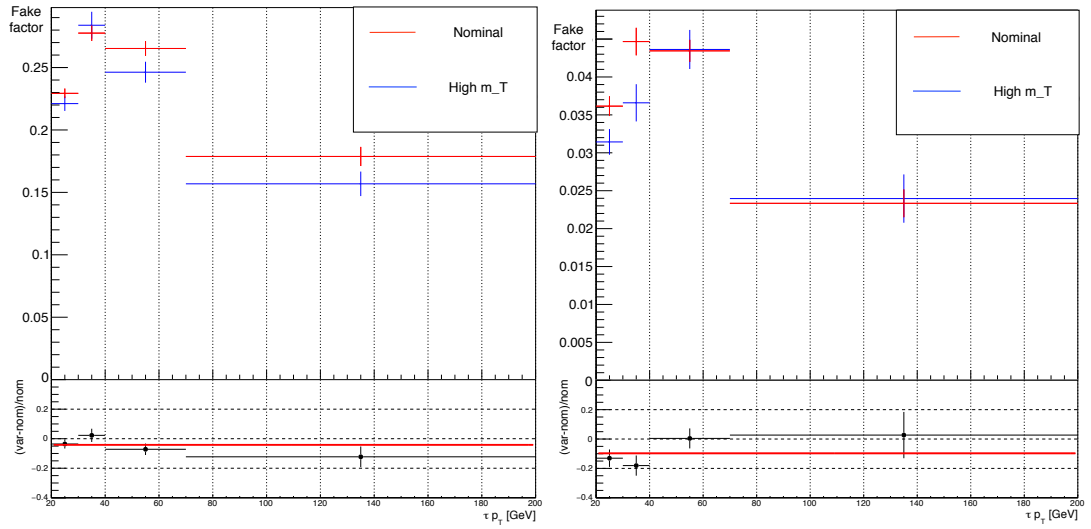


(c) Varied radiation of additional jets



(d) Varied hard scatter simulation ('aMC@NLO') and fragmentation model ('Herwig')

FIGURE 12.9: The estimated shape variation of the fake background as a function of p_{T}^{τ} , arising from the following variations of the true-τ_{had} $t\bar{t}$ MC sample: (a) the 'up' variation of detector-related uncertainties; (b) the 'down' variation of detector-related uncertainties; (c) low and high radiation variations, referred to as 'radHi' and 'radLo'; (d) the variation of the hard scatter simulation (labelled 'aMC@NLO', simulated using the aMC@NLO generator as a replacement for POWHEG) and fragmentation model (labelled 'Herwig', calculated by replacing the POWHEG+PYTHIA6 fragmentation model with POWHEG+HERWIG).



(a) One-prong (b) Three-prong

FIGURE 12.10: Fake factors derived for the $t\bar{t}$ background in the two b-tag region using MC simulated data in the signal region (labelled 'nominal', shown in red) and in the high-m_{T} region (blue) for (a) one-prong τ_{had} decays and (b) three-prong τ_{had} decays.

in each region is found to have no shape-dependence and, therefore, the systematic uncertainty is applied as a flat normalisation of 0.9.

Similarly, the systematic covering the extrapolation of the QCD multi-jet fake factor from the inverse lepton isolation control region to the signal region is calculated by comparing the nominal fake factor with a fake factor calculated in an alternative control region defined by requiring the τ_{had} and lepton to have same sign charge (but pass the isolation requirements).

12.2.1.4 Statistical uncertainty

Finally, the statistical uncertainty on the fake factor is taken into account. The fake factors are varied up and down by their statistical uncertainty (all bins are varied together) and the variations of the fake-τ_{had} estimate are included as a normalisation uncertainty.

12.3 $Z \to \tau\tau$ + heavy-flavour jets

The cross-section for Z boson production in association with a pair of heavy-flavour jets (combinations of b- and c-jets, i.e. $Z \to \tau\tau + bb,\ bc,\ cc$) is known to be poorly predicted by MC generators; a mismodelling of the normalisation of this process is observed in a number of ATLAS analyses [34]. Therefore, MC simulations for these processes are normalised to data. Similarly to the $t\bar{t}$ normalisation described in Section 12.1, the

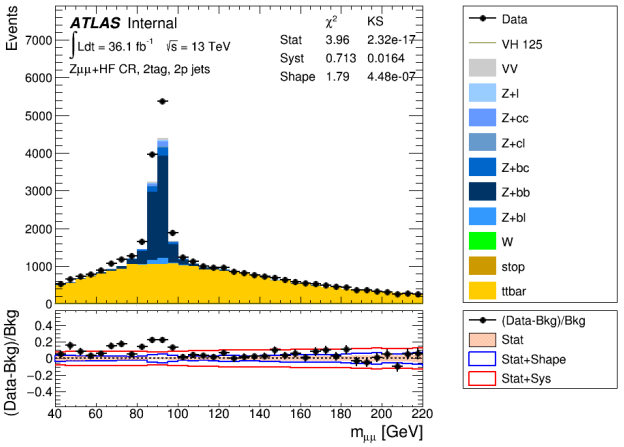


FIGURE 12.11: Distribution of $m_{\mu\mu}$ in the $Z \to \mu\mu$ + heavy-flavour control region. The Z boson production mismodelling is evident in the data-to-MC ratio in the bottom pane.

normalisation is allowed to vary in the final fit, as described in Section 5.1. The normalisation is constrained by a control region of $Z \to \mu\mu$ + heavy-flavour data events (the production cross-section of Z bosons should be independent of their decay mode). The event selection for this control region closely follows that for the signal region, but instead of two τ-leptons, it requires exactly two muons with $p_{\mathrm{T}}^{\mu} > 27$ GeV and di-muon invariant mass $81 < m_{\mu\mu} < 101$ GeV. In order to veto the contribution from SM $Vh(h \to bb)$ production, it is required that $m_{bb} < 80$ GeV or $m_{bb} > 140$ GeV.

Figure 12.11 shows the distribution of $m_{\mu\mu}$ in the $Z \to \mu\mu$ + heavy-flavour control region. The Z boson processes are shown in blue, forming a peak in $m_{\mu\mu}$ around the Z boson mass ($m_{\mu\mu} \sim 90$ GeV) where there is a clear deficit of MC-simulated background events.

The normalisation is applied to all $Z \to \tau\tau + bb, bc, cc$ and $Z \to \ell\ell + bb, bc, cc$ (where ℓ represents muons and electrons) processes. Separate normalisation factors are derived for each signal region fit (i.e. for each signal hypothesis and mass). For the background-only fit in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ signal region, a normalisation factor of 1.34 ± 0.16 relative to the nominal MC prediction is obtained.

The post-fit yields in the $Z \to \mu\mu$ + heavy-flavour control region are shown in Table 12.1; the total background yield and data yield are consistent.

12.3.1 Uncertainties on $Z \to \tau\tau$ + heavy-flavour (bb, bc, cc) processes

Generator-level uncertainties, as listed below, are applied to the $Z \to \tau\tau$ + heavy-flavour jets backgrounds. The variation from the nominal sample generated by each uncertainty

Sample	Post-fit yield
$Z \to \ell\ell$ + bb, bc, cc	8420 ± 550
$t\bar{t}$	3950 ± 510
Other	520 ± 180
Total background	12900 ± 110
Data	12897

TABLE 12.1: Post-fit yields in the $Z \to \mu\mu$ + heavy-flavour control region.

is parametrised as a function of m_{bb} and p_{T}^{bb}, as shown in Figure 12.12, and propagated through the analysis to give a shape uncertainty on the signal/background BDT output distribution.

- The uncertainties on the $Z \to \tau\tau$ + heavy-flavour jets backgrounds relating to the choice of renormalisation and factorisation scales are derived by varying the scales up and down, either together or independently, by a factor of two using the event weights included in the SHERPA 2.2.1 samples. The overall shape uncertainty related to the choice of scale is an envelope of these variations. The individual effects of these variations on the $Z \to \tau\tau$ + heavy-flavour jets modelling are shown in Figures 12.12a and 12.12b. For example, 'MUR0.5_MUF0.5' is the difference between the nominal $Z \to \tau\tau$ + heavy-flavour background and that yielded from varying the renormalisation and factorisation scales down by a factor of two; 'MUR1_MUF2' is the difference between the nominal background and that with only the factorisation scale varied up by a factor of two.

- The uncertainties related to the choice of PDF prescription are evaluated using the event weights included in the samples; these include 100 variations on the nominal NNPDF3.0 PDF set and central values for two alternative PDF sets, MMHT2014NNLO68CL [183] and CT14NNLO [184]. The intra-PDF uncertainty for NNPDF3.0 is estimated as the standard deviation of the 101 variations, while an additional uncertainty is formed from the envelope of the differences between NNPDF3.0 and the two alternatives. The individual effects of these variations on the $Z \to \tau\tau$ + heavy-flavour jets modelling are shown in Figures 12.12c and 12.12d.

- The nominal SHERPA 2.2.1 samples are compared to alternative samples generated using MADGRAPH5_aMC@NLO v2.2.2 interfaced with the PYTHIA 8.186 parton shower model (these use the A14 tune together with the NNPDF2.3LO PDF set and the EvtGen v1.2.0 program is used for properties of the bottom and charm hadron decays). The individual effects of these variations on the $Z \to \tau\tau$ + heavy-flavour jets modelling are shown in Figures 12.12e and 12.12f.

The total theoretical uncertainties on the $Z \to \tau\tau$ + heavy-flavour jets background, parametrised as a function of m_{bb} and p_{T}^{bb}, are shown as thick black lines in Figure 12.12. The dominant effect on p_{T}^{bb} is the renormalisation and factorisation scale uncertainty, while the dominant effect on m_{bb} is the choice of generator; the PDF uncertainties are negligible by comparison. Therefore, the total uncertainties are taken as symmetrised envelopes of these two categories of variations.

12.3.2 Acceptance uncertainties on $Z \to \tau\tau$ + heavy-flavour (bb, bc, cc) and $t\bar{t}$ processes

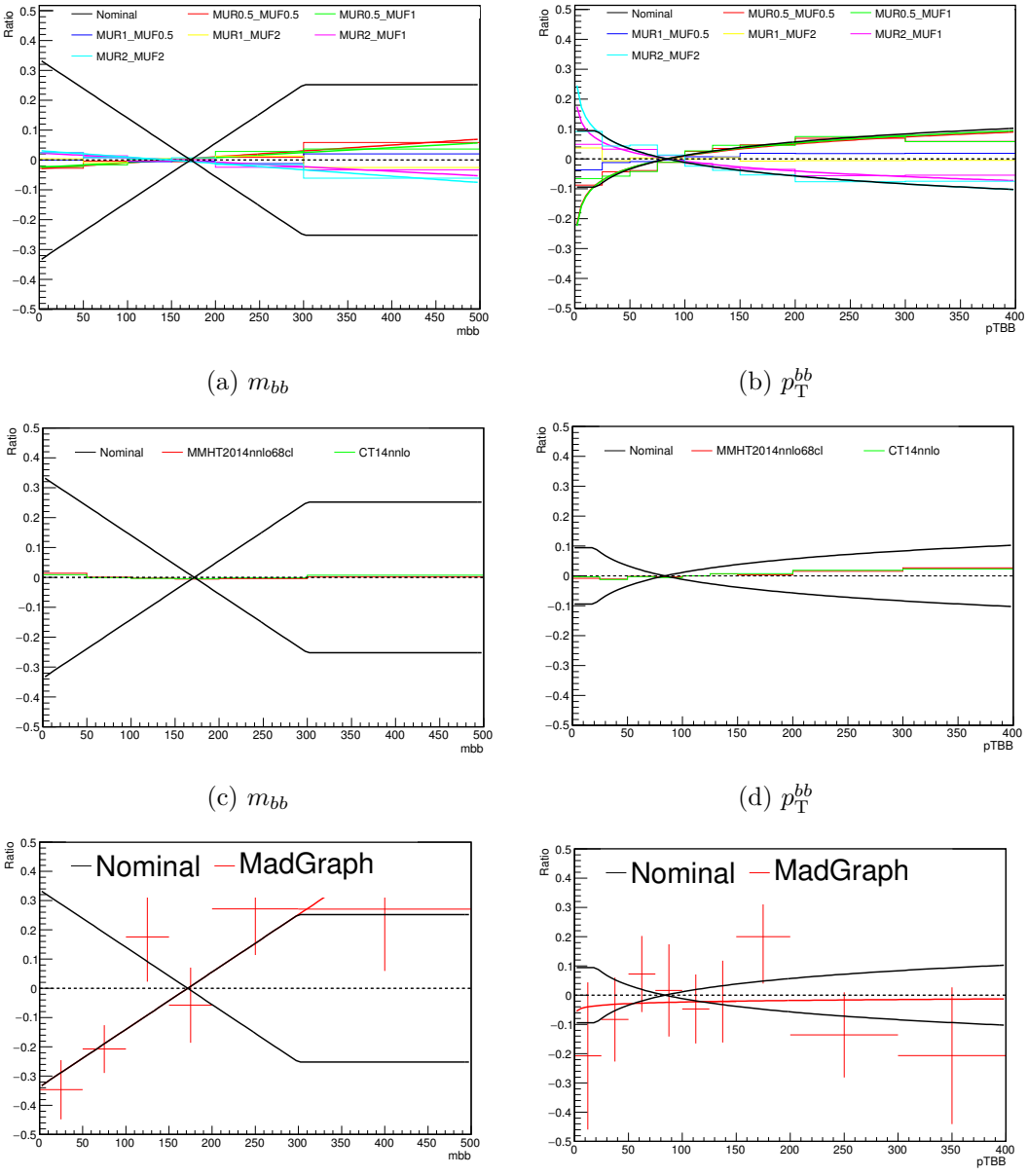
Uncertainties are applied when extrapolating the $t\bar{t}$ and $Z \to \tau\tau$ + heavy-flavour normalisations between the signal region and the $Z \to \mu\mu$ control region to account for differences in acceptance between the regions.

- The $t\bar{t}$ normalisation (described in Section 12.1) is derived in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ signal region (in the low BDT score region of the BDT distribution) and extrapolated to the $Z \to \mu\mu$ + heavy-flavour control region, where it is allowed to vary from the derived value within an uncertainty of $^{+8}_{-9}\%$. This uncertainty is estimated by calculating the ratio of the MC yields between the $Z \to \mu\mu$ + heavy-flavour control region and the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ signal region in each $t\bar{t}$ variation sample (see the bulleted list in Section 12.1.1), and comparing these with the ratio calculated between the two regions in the nominal sample. This quantifies the error in the relative rate of $t\bar{t}$ production in the two regions predicted by the nominal MC simulation.
- The $Z \to \tau\tau$ + heavy-flavour normalisation is derived in the $Z \to \mu\mu$ + heavy-flavour control region and extrapolated to the signal region, where it is allowed to vary from the derived value within an uncertainty of $\pm 29\%$. This uncertainty is estimated using a similar method to the $t\bar{t}$ extrapolation uncertainty, using the variations described in Section 12.3.1 (varied renormalisation and factorisation scales, PDF set and MC generator).

12.4 $Z \to ee$ + jets

The normalisation derived for the $Z \to \tau\tau$ + heavy-flavour jets background (discussed in Section 12.3) is applied to the $Z \to ee$ + heavy-flavour jets contributions.

The contribution to the Z boson background from Z+ light-flavour jets is small; it mainly originates from $Z \to \tau\tau$ + light-flavour jets, which does not include fake τ_{had} events and



(a) m_{bb} (b) p_{T}^{bb}

(c) m_{bb} (d) p_{T}^{bb}

(e) m_{bb} (f) p_{T}^{bb}

FIGURE 12.12: Variation of $Z \to \tau\tau$ + heavy-flavour jets background as a function of m_{bb} (left) and p_{T}^{bb} (right) when varying: the renormalisation and factorisation scales (top); the PDF set (middle); and the MC generator (bottom). The total uncertainty applied to the $Z \to \tau\tau$ + heavy-flavour background is shown as a thick black line; in m_{bb} this is taken as a symmetrised envelope of the renormalisation and factorisation scale uncertainties, and in p_{T}^{bb} this is taken as a symmetrised envelope of the MC generator uncertainties. The PDF uncertainties are negligible by comparison.

is therefore well-modelled by MC simulation. However, the $Z \rightarrow ee$ + light-flavour jets background includes events where an electron fakes a τ_{had}. This may be mismodelled by MC simulation and is therefore checked in data by constructing a further control region to check the normalisation. The control region is defined by applying the standard event preselection and additionally requiring:

- $m_{\text{T}}^{W} < 40$ GeV,
- $E_{\text{T}}^{\text{miss}}$ ϕ centrality > 0 (this is defined in Section 13.1),
- 81 GeV $< m_{\tau\tau}^{\text{visible}} <$ 101 GeV,
- either zero or one b-tagged jets.

The first two cuts are applied to reduce the $t\bar{t}$ background, whilst the third selects the region around the Z boson mass.

In the zero b-tag control region, a normalisation factor of 0.81 ± 0.06 (stat) is found, while the normalisation factor in the one b-tag control region is found to be consistent with unity. Hence, no normalisation is applied to the central value of the $Z \rightarrow ee$ + jets background; a conservative uncertainty equal to the difference between unity and the zero b- tag normalisation factor (19%) is applied to account for any mismodelling in the rate of electrons faking τ_{had} decays.

12.5 Uncertainties on other background processes

For all other background processes estimated from MC simulation (see Chapter 8 for a summary of the MC samples used), the theoretical uncertainties on the cross-sections are taken into account.

The uncertainties on the $Z \rightarrow \tau\tau$ + light-flavour (lc, ll) and di-boson production processes are found to be 5% and 6% respectively, when combining the PDF, α_s and scale uncertainties in quadrature (as outlined in [185]). For true W + heavy-flavour jets processes, a 30% uncertainty is assigned, as estimated in [186] from generator comparisons, the refactorisation and renormalisation scale, PDF and α_s variations.

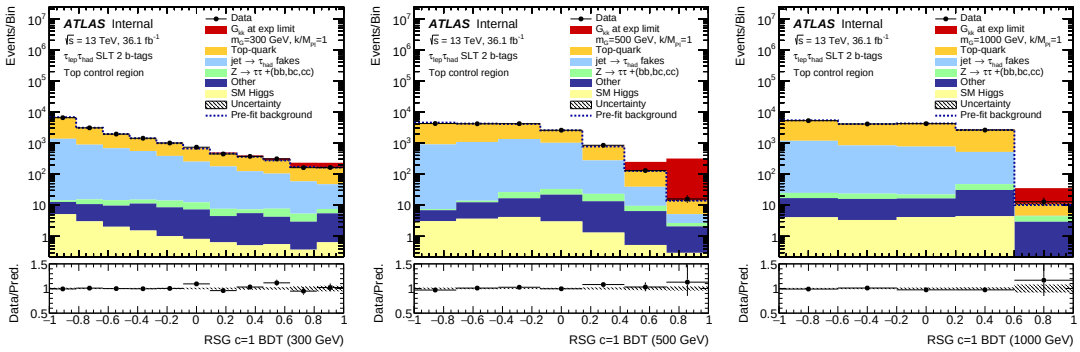
A normalisation uncertainty of 28% is applied to the Zh background component, in line with the latest experimental uncertainty on this measurement [187]. Similarly, a 30% uncertainty is applied to the top-quark associated Higgs production (ttH) background [188].

12.6 Validation of background modelling

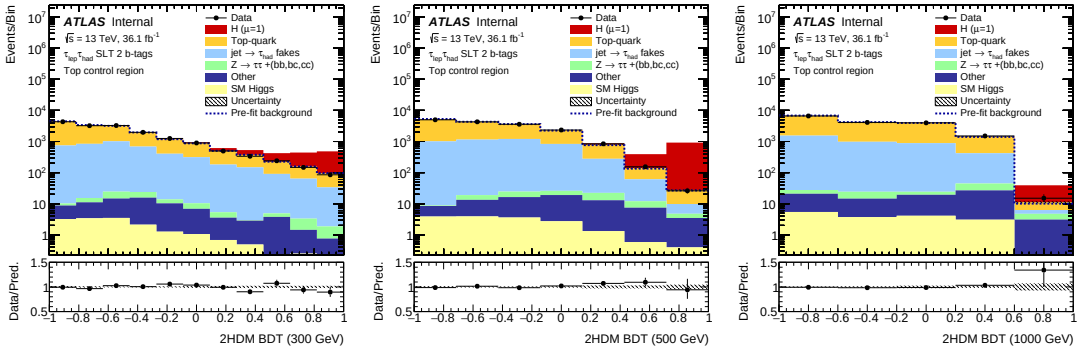
The background modelling is validated by comparing the estimated background with data in specially constructed validation regions. The modelling is studied in analysis BDT distributions for a selection of signal hypotheses. The analysis BDTs are described fully in Chapter 13. The validation regions are as follows:

- The real-τ_{had} $t\bar{t}$ background is validated in the high-m_{T}^{W} region, which is constructed by applying the signal region selection outlined in Section 11.2 with the two b-tag requirement and an additional requirement of $m_{\text{T}}^{W} > 40$ GeV.
- The fake-τ_{had} backgrounds, estimated using the data-driven fake-factor method, are validated in the two b-tag same-sign region, constructed by applying the signal region selection requirements outlined in Section 11.2, but requiring the τ_{had} and lepton to have same-sign (instead of opposite-sign) electric charge. The fake factor is calculated using the method described in Section 12.2 but all calculations are performed using same-sign events.

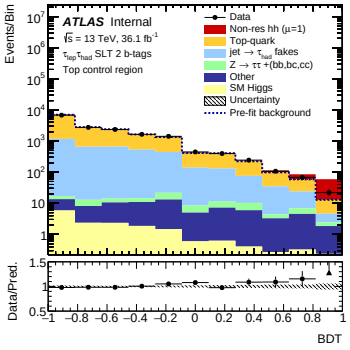
Post-fit BDT distributions in the two b-tag high-m_{T}^{W} validation region are shown in Figure 12.13 for: an RS graviton signal with $k/M_{\text{Pl}} = 1.0$; a generic heavy scalar signal; and an SM non-resonant signal. Post-fit BDT distributions in the two b-tag same-sign validation region are shown in Figure 12.14. In both validation regions, the data is well modelled by the MC simulation within uncertainties.



(a) RS graviton $k/M_{\mathrm{Pl}} = 1.0$, $m_G = 300$ GeV (b) RS graviton $k/M_{\mathrm{Pl}} = 1.0$, $m_G = 500$ GeV (c) RS graviton $k/M_{\mathrm{Pl}} = 1.0$, $m_G = 1000$ GeV

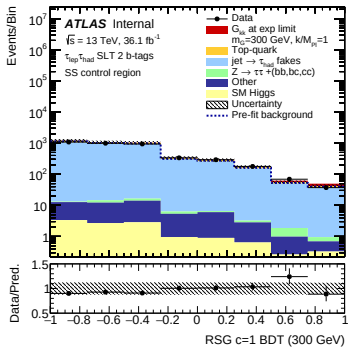


(d) Scalar, $m_H = 300$ GeV (e) Scalar, $m_H = 500$ GeV (f) Scalar, $m_H = 1000$ GeV

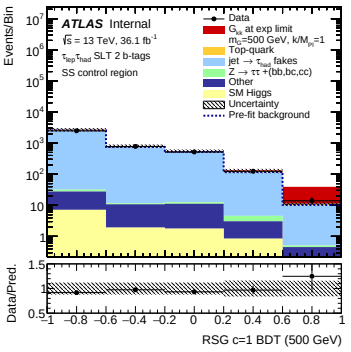


(g) SM non-resonant

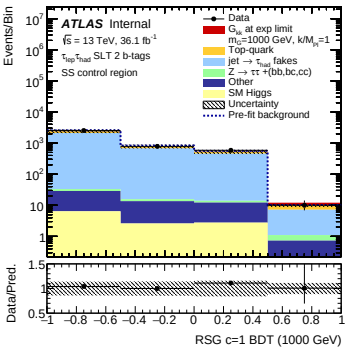
FIGURE 12.13: Post-fit BDT distributions in the high-m_{T}^{W} validation region for: an RS graviton signal with $k/M_{\mathrm{Pl}} = 1.0$ and (a) $m_G = 300$ GeV, (b) $m_G = 500$ GeV, and (c) $m_G = 1000$ GeV; a generic scalar signal with (d) $m_H = 300$ GeV, (e) $m_H = 500$ GeV, and (f) $m_H = 1000$ GeV; and (g) a SM non-resonant di-Higgs signal.



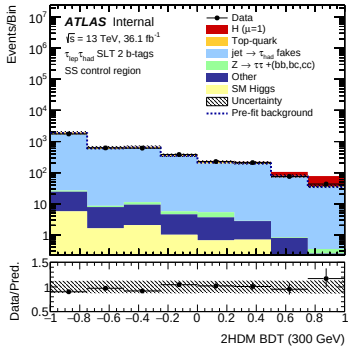
(a) RS graviton $k/M_{\mathrm{Pl}} = 1.0$, $m_G = 300$ GeV



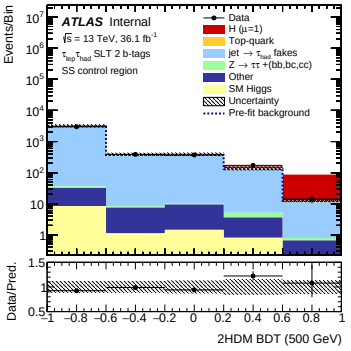
(b) RS graviton $k/M_{\mathrm{Pl}} = 1.0$, $m_G = 500$ GeV



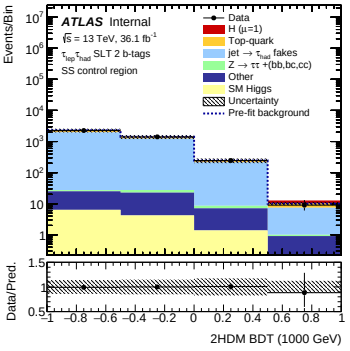
(c) RS graviton $k/M_{\mathrm{Pl}} = 1.0$, $m_G = 1000$ GeV



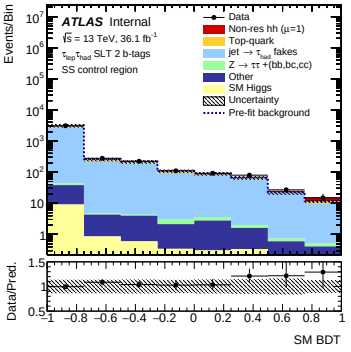
(d) Scalar, $m_H = 300$ GeV



(e) Scalar, $m_H = 500$ GeV



(f) Scalar, $m_H = 1000$ GeV



(g) SM non-resonant

FIGURE 12.14: Post-fit BDT distributions in the same-sign validation region for: an RS graviton signal with $k/M_{\mathrm{Pl}} = 1.0$ and (a) $m_G = 300$ GeV, (b) $m_G = 500$ GeV, and (c) $m_G = 1000$ GeV; a generic scalar signal with (d) $m_H = 300$ GeV, (e) $m_H = 500$ GeV, and (f) $m_H = 1000$ GeV; and (g) a SM non-resonant di-Higgs signal.

Chapter 13

BDT Training

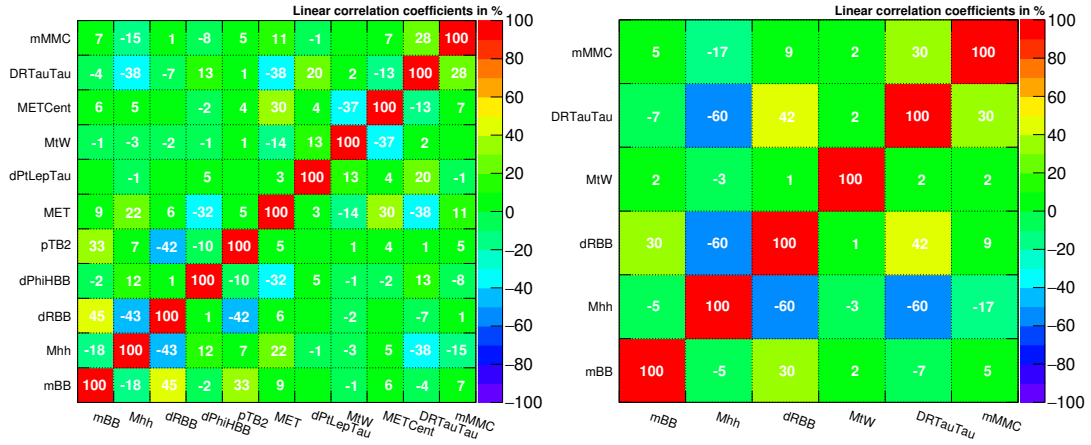
Variables describing the kinematics of events can be used to provide varying degrees of discrimination between signal and background. Whilst some of these can be used to place cuts on the data, removing significant amounts of background, others exhibit only minor separation power. By combining these variables in a BDT, it is possible to exploit the correlations between the variables to achieve increased sensitivity to the signal process relative to using a cuts-based analysis, producing a single distribution for which signal and background are well-separated. The BDT score distribution is used as the discriminant in the statistical fit for hypothesis testing.

The theory behind BDTs is described in more detail in Chapter 4. This chapter documents the strategy used for BDT training and implementation.

13.1 Variables

A number of kinematic variables providing discrimination between the signal and background distributions were identified and used to produce a baseline BDT training. These variables were required to be well modelled by MC simulation.

The inclusion of large numbers of variables in a BDT can lead to overtraining; variables that provide little improvement to the analysis sensitivity or have high correlation with another variable should be removed. Due to the varying kinematics expected for the different mass hypotheses, the separation power exhibited by a variable is dependent on the signal mass. However, for simplicity, the same variables are used for both resonant signal hypotheses over the full signal mass range. For the non-resonant search, the BDT was optimised separately such that the variables used are a subset of those used for the resonant searches.



(a) RS graviton $k/\bar{M}_{\mathrm{Pl}} = 1.0$, $m_G = 500$ GeV (b) SM non-resonant signal

FIGURE 13.1: The correlation indices for each pair of variables used to train BDTs for (a) an RS graviton $k/\bar{M}_{\mathrm{Pl}} = 1.0$ and $m_G = 500$ GeV, and (b) a SM non-resonant signal. These indices are all given as percentages. The correlation is below 60% for every pair of variables.

TMVA derives a ranking of the variables used by the decision trees (calculated for the full decision tree forest) by counting how often each variable is used to split decision tree nodes, and weighting each split occurrence by the *separation gain-squared* it achieves and the number of events in the node. The separation gain achieved when a single parent node is split into two child nodes is given by the change in the gini index; see Section 4.3.

The lowest-ranked of any pairs of variables with a high correlation index ($> 60\%$) across the full signal mass range were removed, as well as any variables with a consistently low importance ranking. For every new list of input variables, the BDTs were retrained. By observing the impact on the receiver operating characteristic (ROC) curves, which illustrate the efficiencies with which signal and background events are correctly classified, and propagating the changes to the expected limits, the performance of the BDTs were assessed. Variables were removed one-by-one and the resulting BDT was tested; only those for which their removal resulted in a deterioration of the BDT performance were used in the final training.

Figure 13.1 shows the correlation indices for each pair of variables used to train BDTs for (a) an RS graviton $k/\bar{M}_{\mathrm{Pl}} = 1.0$ and $m_G = 500$ GeV, and (b) a SM non-resonant signal. No pairs of variables are highly correlated.

The variables used by the BDTs are defined below:

- m_{bb} is the invariant mass of the di-b-jet system.

- $m_{\tau\tau}^{\text{MMC}}$ is the invariant mass of the di-τ system, calculated using the MMC (see Section 9.7).

- m_{hh}, the invariant mass of the di-Higgs system, is reconstructed from the di-τ and di-b-jet masses, $m_{\tau\tau}^{\text{MMC}}$ and m_{bb}. In order to improve the mass resolution, the four-momenta of the di-τ and di-b-jet systems are scaled by $m_h/m_{\tau\tau}^{\text{MMC}}$ and m_h/m_{bb} respectively, where $m_h = 125$ GeV is the value of the Higgs mass used in the simulation.

- $p_{\text{T}}^{b_2}$ is the transverse momentum of the sub-leading b-jet.

- $\Delta R(\ell, \tau_{\text{had}})$ is the ΔR separation between the light lepton and the visible decay products of the hadronic τ-lepton.

- $\Delta p_{\text{T}}(\ell, \tau_{\text{had}})$ is the difference in p_{T} between the light lepton and the visible decay products of the hadronic τ-lepton.

- $\Delta R(b_1, b_2)$ is the ΔR separation between the two b-jets.

- $\Delta\phi(h, h)$ is the $\Delta\phi$ between the two reconstructed Higgs candidates, where the direction of the di-τ system is deduced from the MMC fit.

- $E_{\text{T}}^{\text{miss}}$ is the missing transverse momentum of the event (defined in Section 9.8).

- m_{T}^{W} is the transverse mass of the light lepton and $E_{\text{T}}^{\text{miss}}$and is defined in Section 12.2.

- $E_{\text{T}}^{\text{miss}}$ ϕ centrality quantifies the position in ϕ of the $E_{\text{T}}^{\text{miss}}$ with respect to the visible decay products of the τ-leptons:

$$E_{\text{T}}^{\text{miss}} \ \phi \text{ centrality} = \frac{A + B}{\sqrt{A^2 + B^2}}, \tag{13.1}$$

  where A and B are given by

$$A = \frac{\sin(\phi_{E_{\text{T}}^{\text{miss}}} - \phi_{\tau_{\text{had}}})}{\sin(\phi_{\ell} - \phi_{\tau_{\text{had}}})}, \quad B = \frac{\sin(\phi_{\tau_{\text{had}}} - \phi_{E_{\text{T}}^{\text{miss}}})}{\sin(\phi_{\ell} - \phi_{\tau_{\text{had}}})}. \tag{13.2}$$

Table 13.1 lists the input variables used to train BDTs for both resonant and non-resonant signals.

Distributions of the variables used to train the BDTs for the resonant signals, after performing the full fitting procedure with a background-only hypothesis, are shown in Figure 13.2. The RS graviton signal with $k/\bar{M}_{\text{Pl}} = 1.0$ and $m_G = 500$ GeV is plotted as a red line to show how the signal and background differ for each variable.

Variable	Resonant	Non-resonant
m_{hh}	✓	✓
$m_{\tau\tau}^{\text{MMC}}$	✓	✓
$\Delta R(l,\tau)$	✓	✓
m_{bb}	✓	✓
m_{T}^{W}	✓	✓
$\Delta R(b_1,b_2)$	✓	✓
$p_{\text{T}}^{b_2}$	✓	
$\Delta p_{\text{T}}(l,\tau)$	✓	
$E_{\text{T}}^{\text{miss}}$ ϕ centrality	✓	
$\Delta\phi(h,h)$	✓	
$E_{\text{T}}^{\text{miss}}$	✓	

TABLE 13.1: The variables used to train the BDTs for the resonant and non-resonant signal hypotheses. These are ordered according to the ranking produced by TMVA (from most to least important) when training a BDT using the RSG $m_G = 400$ GeV signal hypothesis. The variable importance varies with mass; variables that have little discriminating power for higher masses rank highly for the $m_G = 400$ hypothesis. Therefore, more variables are used in the BDT for the resonant searches than for non-resonant.

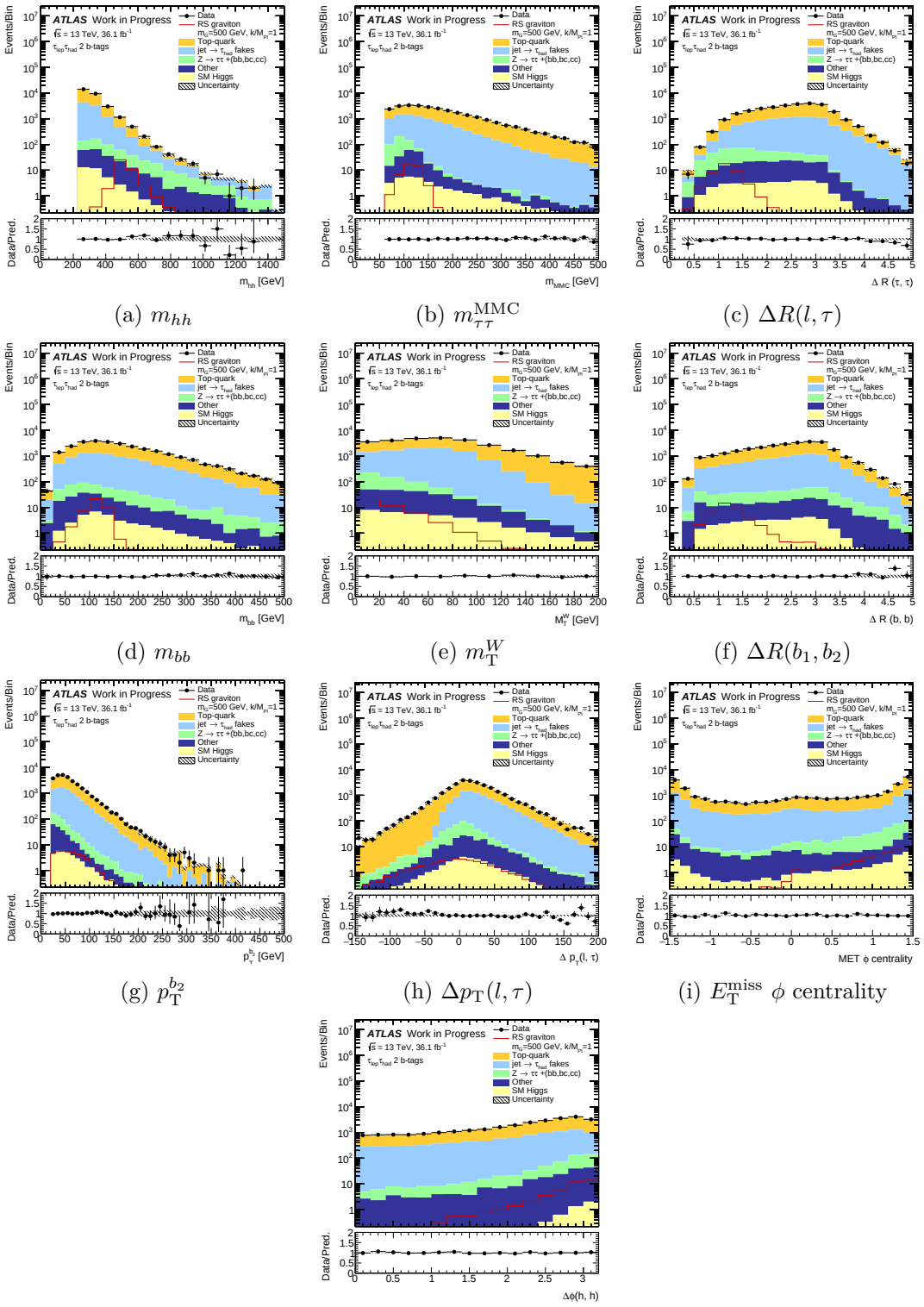
Figure 13.3 shows the variables used to train the BDTs for the non-resonant signal. These are shown after performing a background-only fit, with the non-resonant signal overlaid in red.

13.2 Training

Separate BDTs were trained for:

- an RS graviton, G, with $k/\bar{M}_{\text{Pl}} = 1.0$, for nine mass points between 260 GeV and 1 TeV;
- an RS graviton, G, with $k/\bar{M}_{\text{Pl}} = 2.0$, for 14 mass points between 260 GeV and 1 TeV;
- a generic narrow-width heavy scalar, H, for 14 mass points between 260 GeV and 1 TeV;
- non-resonant SM Higgs pair-production.

Each BDT is trained using signal MC samples weighted by their predicted cross-sections and truth-tagged (see Section 9.4.3) in order to improve statistics. These are trained against the dominant $t\bar{t}$ background, with both real and fake τ_{had} components represented by MC simulation. Events are required to pass the selection criteria outlined in Section 11.2 and contain exactly two b-tagged jets.



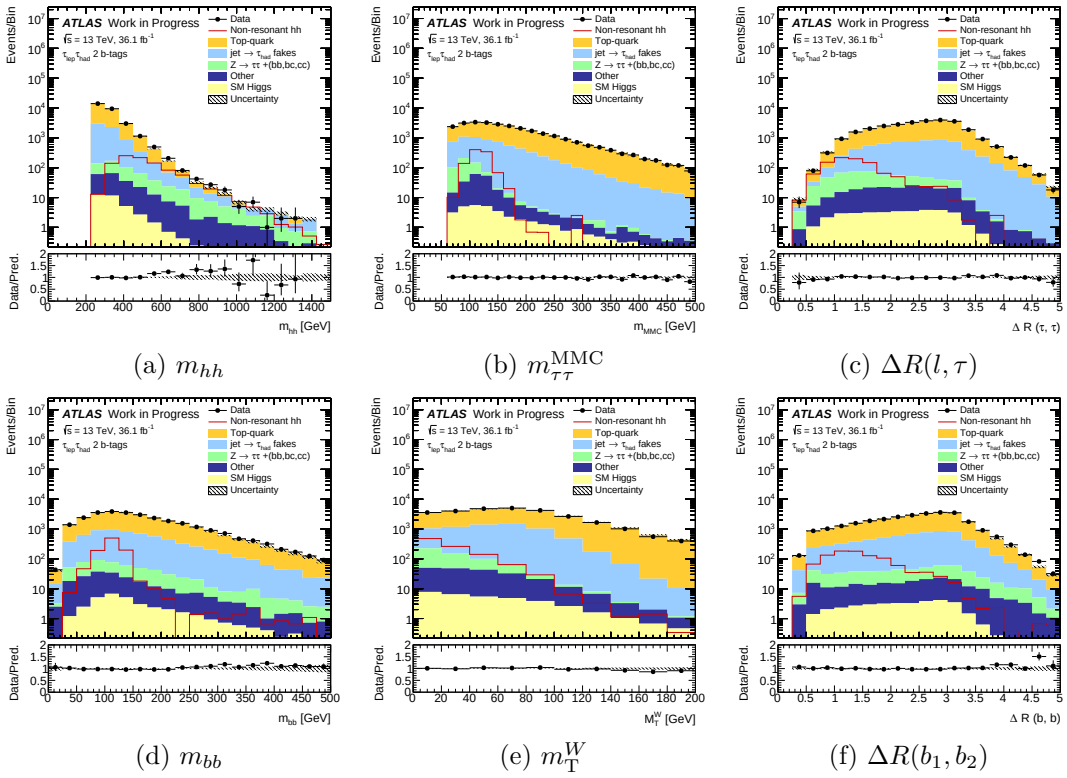
(a) m_{hh} (b) $m_{\tau\tau}^{\mathrm{MMC}}$ (c) $\Delta R(l,\tau)$

(d) m_{bb} (e) m_{T}^{W} (f) $\Delta R(b_1,b_2)$

(g) $p_{\mathrm{T}}^{b_2}$ (h) $\Delta p_{\mathrm{T}}(l,\tau)$ (i) $E_{\mathrm{T}}^{\mathrm{miss}}$ ϕ centrality

(j) $\Delta\phi(h,h)$

FIGURE 13.2: Distributions of the variables used to train the resonant BDTs, shown after performing the background-only fit (see Section 5.1). The RS graviton signal with $k/\bar{M}_{\mathrm{Pl}} = 1.0$ and $m_G = 500$ is shown as a red line.



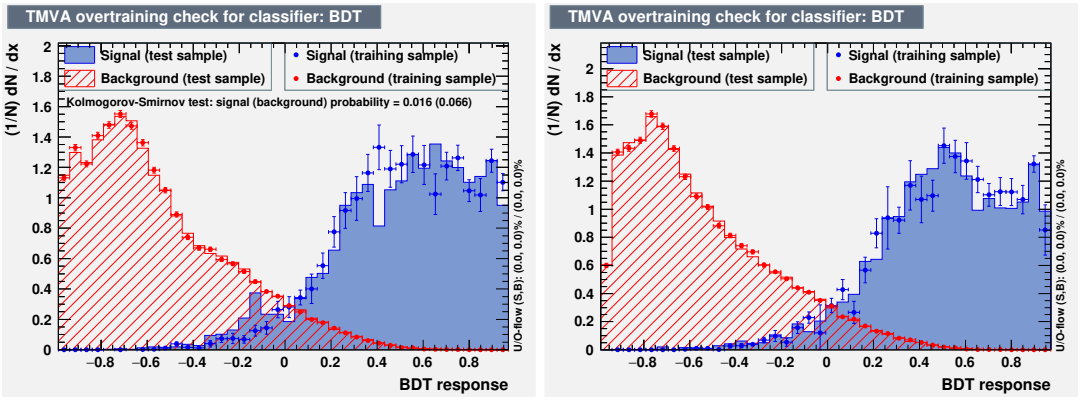
(a) m_{hh} (b) $m_{\tau\tau}^{\text{MMC}}$ (c) $\Delta R(l, \tau)$

(d) m_{bb} (e) m_{T}^{W} (f) $\Delta R(b_1, b_2)$

FIGURE 13.3: Distributions of the variables used to train the non-resonant BDTs, shown after performing the background-only fit (see Section 5.1). The non-resonant Standard Model signal is shown as a red line.

The BDTs are trained using a two-fold method for validation, as outlined in Section 4.3, such that two classifiers (referred to as odd and even) are applied to data; the odd classifier is applied to even event numbers and the even classifier is applied to odd. This method is used to test for and reduce overtraining. Figure 13.4 is used to illustrate the two-fold method; it shows the BDT classifiers for the non-resonant SM hypothesis. Figure 13.4a is the even BDT classifier and Figure 13.4b is odd. The test samples are represented by solid histograms, whereas the training samples are shown as points with error bars.

The BDT is overtrained if there is a significant deviation between the training and test distributions for either signal or background, or both, for either the even or odd classifier. Some examples of these distributions are shown in Figure 13.4 and 13.5. The Kolmogorov-Smirnov (KS) goodness-of-fit test value (which quantifies the similarity between the training and test BDT distribution) is shown on each plot. An overtrained BDT would exhibit a large discrepancy between the training and test samples, a 'spiky' profile, and a small KS probability. In Figure 13.5, the KS probability increases with mass for both the RS graviton signal and the scalar signal; this is because the number of MC signal events increases with mass. No significant overtraining is observed for any signal hypothesis.



(a) 'Even' training (b) 'Odd' training

FIGURE 13.4: The distributions of BDT scores for the 'even' and 'odd' training data sets and their corresponding test sets: (a) is the 'even' BDT training, tested on 'odd'; (b) is the 'odd' training, tested on 'even'. Signal is shown in blue and background in red. The test samples are represented by solid histograms whereas the training samples are both shown as points with error bars.

Furthermore, it can also be concluded from Figure 13.5 that the BDT discriminates between signal and background more successfully with increasing mass. This is because the signal kinematics at low mass are more similar to those of the $t\bar{t}$ background.

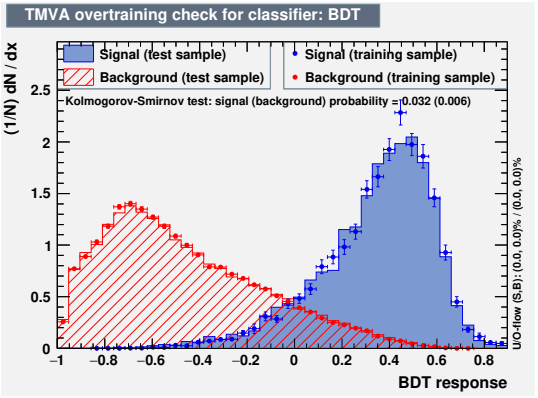
13.2.1 Sensitivity to intermediate masses

Figure 13.6 shows the signal m_{hh} distributions for the scalar signal at a number of different mass points in the range $260 \leqslant m_H \leqslant 1000$ GeV. The m_{hh} peak is shown to be well resolved at low mass. This effect, coupled with the high performance of the BDT in isolating the signal for which it is trained, results in a BDT that is insensitive to signal masses that fall between those used for the BDT training. This loss of sensitivity is most evident at low signal mass.

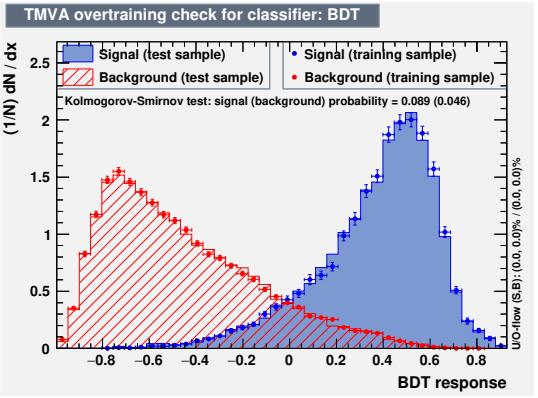
To counteract this loss of sensitivity and avoid missing a potential signal, each resonant BDT is trained with three signal MC samples: the nominal signal sample, plus the two with masses either side of the nominal mass. For example, the BDT training for a scalar signal with $m_H = 500$ GeV is performed using a merged signal sample comprising the samples generated with $m_H = 450$ GeV, $m_H = 500$ GeV and $m_H = 550$ GeV. The m_{hh} distribution of this merged sample has a broader peak, resulting in increased sensitivity to intermediate masses.

There are two methods used to test the analysis sensitivity to intermediate mass points:

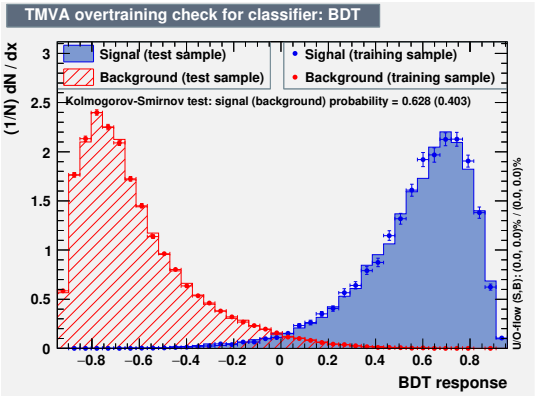
1. A *signal injection* test involves 'injecting' a simulated signal sample at a given mass into the background simulation. A signal-plus-background fit (see Section 5.1) is



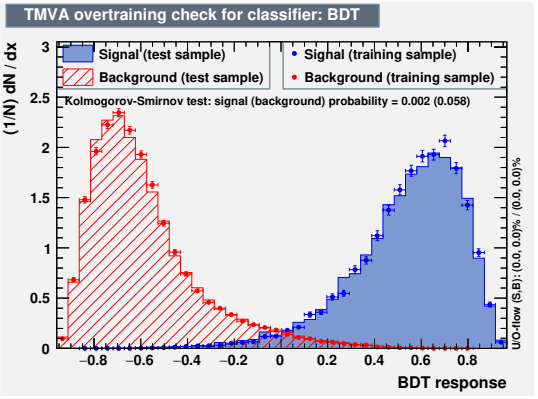
(a) RS graviton 300 GeV



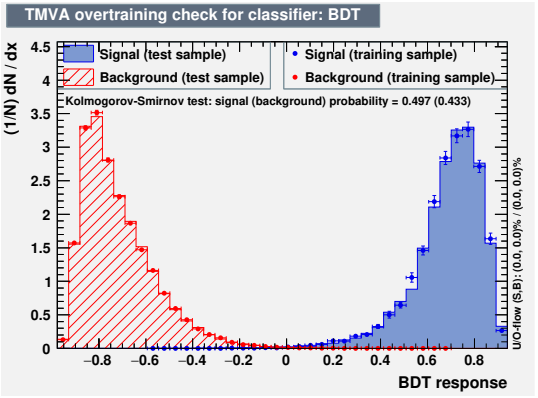
(b) Scalar 300 GeV



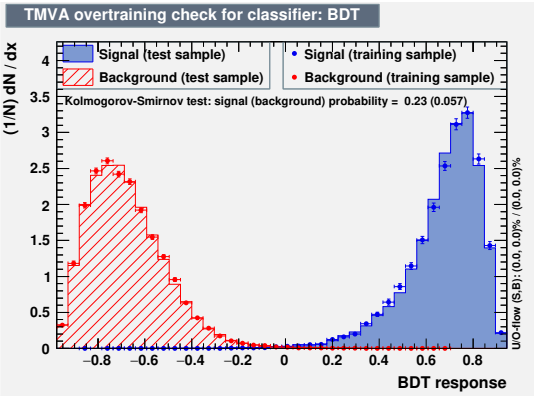
(c) RS graviton 500 GeV



(d) Scalar 500 GeV



(e) RS graviton 1 TeV



(f) Scalar 1 TeV

FIGURE 13.5: The distributions of BDT scores (for one training and test set) for (a) the RS graviton signal with $m_G = 300$ GeV; (b) the scalar signal with $m_H = 300$ GeV; (c) the RS graviton signal with $m_G = 500$ GeV; (d) the scalar signal with $m_H = 500$ GeV; (e) the RS graviton signal with $m_G = 1$ TeV and (f) the scalar signal with $m_H = 1$ TeV. The lower-mass signals suffer from lower statistics than the higher mass; therefore, they exhibit a slight discrepancy between the training and test distributions and a smaller KS probability. Because the signal kinematics for lower-mass signals are more similar to the background, the BDTs are less able to discriminate between signal and background.

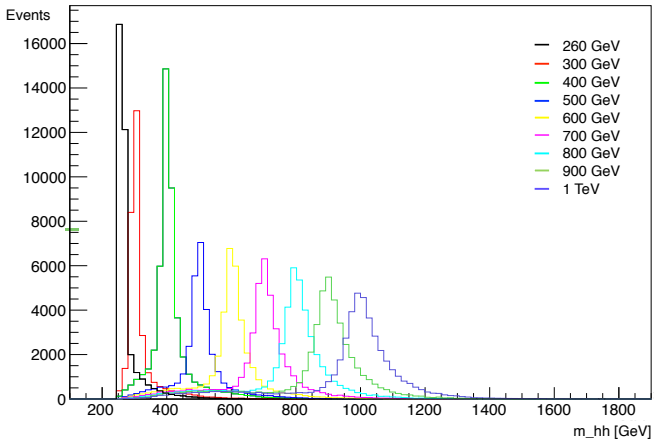
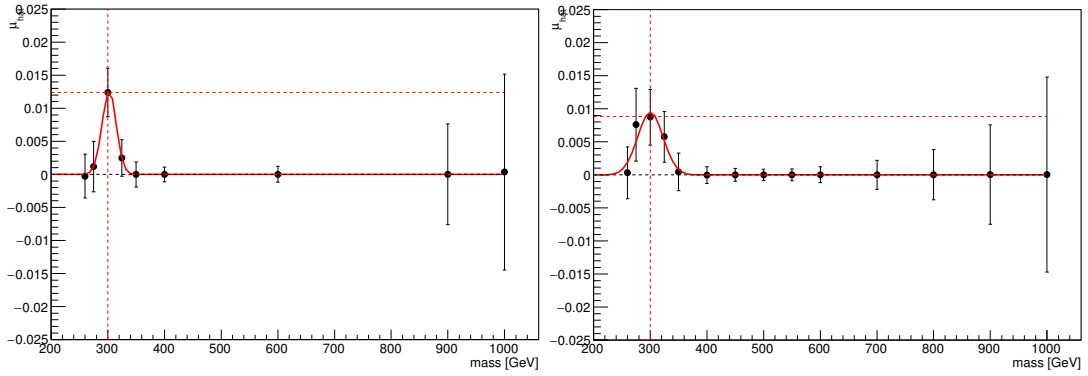


FIGURE 13.6: Distributions of m_{hh} for scalar signal at a number of different mass points in the range $260 \leqslant m_H \leqslant 1000$ GeV. The resolution of the m_{hh} peaks is higher at lower mass; this high resolution leads to problems when searching for a signal mass between the masses for which the BDTs are trained.

then performed in order to test whether the injected signal is observed. The value of the reconstructed signal mass, m_{hh}, and signal strength, μ, (also defined in Section 5.1) are calculated; these are expected to match the injected values. This test was performed for both resonant hypotheses using signal injections of $m_{H,G} = 300$ GeV, $m_{H,G} = 400$ GeV and $m_{H,G} = 500$ GeV. Figure 13.7 shows the reconstructed width of the injected signal mass; when training the BDTs with three signal samples, the mass width is greater and the signal strength is lower. This shows that the sensitivity of the search is more uniform across a wider mass range. Figure 13.8 shows the result of this test in the form of 95% confidence level upper limits for each of the three injected masses.

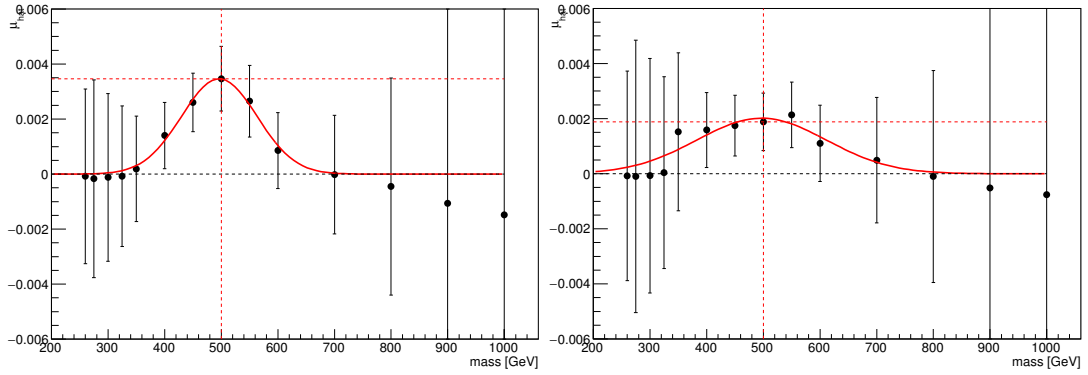
2. As shown in Figure 13.9, the BDT sensitivity to intermediate mass points can be tested by applying the BDT from the lower (higher) mass point as the final discriminant in the likelihood fit and propagating this to the limits. For example, the 300 GeV BDT is applied to the 400 GeV signal sample and the limits are calculated. When using the three-mass BDT training, the limits using the nominal BDT and those trained for the lower and higher mass points are more similar; therefore, we can expect a smaller drop in sensitivity than when using the single-mass BDT training.

From the results of both tests, it can be assumed that when using BDTs trained with three signal MC samples, the analysis is sensitive to a signal with an intermediate mass. However, the sensitivity of the search is negatively impacted across the mass range, resulting in higher expected limits.



(a) Single-mass 300 GeV. Reconstructed width = 12.5 GeV.

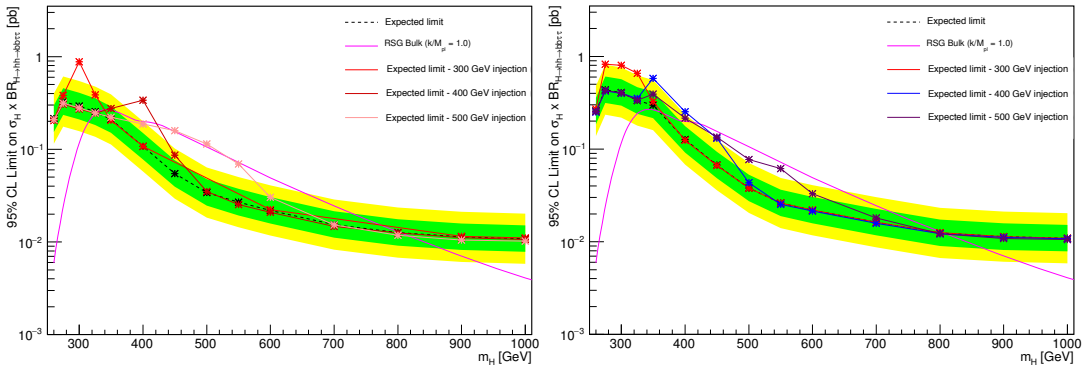
(b) Merged-mass 300 GeV. Reconstructed width = 22.4 GeV.



(c) Single-mass 500 GeV. Reconstructed width = 66.7 GeV.

(d) Merged-mass 500 GeV. Reconstructed width = 114 GeV.

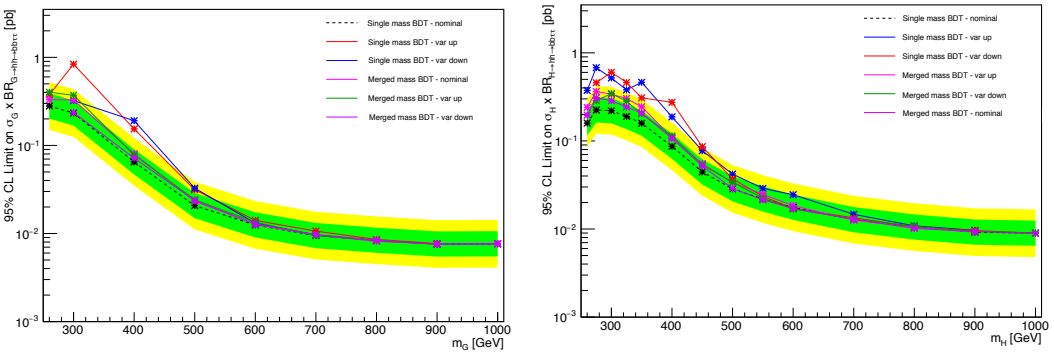
FIGURE 13.7: The reconstructed width of the injected mass in the case where (left) the BDTs are trained on only one signal sample and (right) the BDTs are trained using a merged sample of signal masses. These are shown for (a, b) $m_G = 300$ GeV and (c, d) $m_G = 500$ GeV.



(a) Single-mass BDT limits

(b) Merged-mass BDT limits

FIGURE 13.8: Expected limits for the scalar signal with injected masses of $m_H = 300, 400, 500$ GeV, obtained using (a) BDTs trained on only one signal sample and (b) BDTs trained using a merged sample of three signal masses. The $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the nominal expected limit are shown by green and yellow bands, respectively.



(a) RS graviton (b) Scalar

FIGURE 13.9: Expected 95% CL upper limits on (a) an RS graviton and (b) a narrow-width scalar using BDTs trained with a single mass and with a merged sample of three masses. Limits are also shown when using the BDT from the mass 'up' ('down'), e.g. for a signal with $m_G = 300$ GeV the BDT trained for $m_G = 260$ GeV ($m_G = 400$ GeV) is used. The $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the nominal expected limit are shown by green and yellow bands, respectively.

Chapter 14

Systematic Uncertainties

Systematic uncertainties arise from both theoretical and experimental sources and impact the analysis in two ways: *normalisation* uncertainties affect the expected yields of the signal and background and *shape* uncertainties affect the shapes of distributions. Each of the systematic uncertainties outlined in this chapter is propagated through the full analysis and is included in the final fit as a nuisance parameter (defined in Section 5.1), reducing the analysis sensitivity by introducing additional flexibility to the model. Whilst normalisation variations are always propagated to the fit, only significant shape variations are considered.

The systematic uncertainties are summarised in Section 14.4, which provides a ranking of the systematic uncertainties for the non-resonant fit. The dominant uncertainties are those related to the modelling of the $t\bar{t}$ and $Z \to \tau\tau$ + heavy-flavour backgrounds, the reconstruction and identification of hadronic τ-leptons, and the flavour-tagging efficiencies.

14.1 Experimental Uncertainties

14.1.1 Luminosity

The luminosity uncertainty arises from the calibration of the luminosity scale using x-y beam-separation scans (van der Meer scans) [189] performed in August 2015 and May 2016. The uncertainty on the combined 2015+2016 dataset is $\pm 2.1\%$.

14.1.2 Pile-Up

A pile-up reweighting is applied to MC simulations to correct distributions of the number of pile-up interactions, μ, to match the pile-up conditions for the relevant data-taking periods. This measurement has an associated uncertainty which is also applied [190].

14.1.3 Electrons

Uncertainties on the electron reconstruction, identification, trigger and isolation efficiency scale factors are derived by varying the *tag-and-probe* selection [154]. The tag-and-probe method is used to measure the electron identification efficiency using $Z \to ee$ and $J/\psi \to ee$ data events. Strict selection criteria are applied to one of the two decay electrons (*tag*) and the second electron (*probe*) is used for the efficiency measurements. The variations, provided as a function of the electron transverse energy, E_{T}, and pseudorapidity, η, are between 1% and 5%.

Systematic variations on the electron energy resolution are quantified by smearing the electron energies in MC simulations. The effect of varying the electron energy scale is also included as a systematic uncertainty. The systematic uncertainty in the energy scale calibration varies between 0.03% and 0.2% for electrons with p_{T} around 45 GeV. For electrons with $p_{\mathrm{T}} = 10$ GeV, the uncertainty ranges between 0.3% and 0.8% [191].

14.1.4 Muons

Measurement of the muon reconstruction, identification and trigger efficiency scale factors is performed using the tag-and-probe method with $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ MC events in the high-p_{T} ($p_{\mathrm{T}}^{\mu} > 20$ GeV) and low-p_{T} ($p_{\mathrm{T}}^{\mu} < 20$ GeV) regions respectively, resulting in variations in p_{T}^{μ} between 1% and 7% [156].

The momentum scale and energy resolution variations are also derived in $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ MC events; the shape-dependent systematic uncertainties range between 1.7% and 2.9%.

14.1.5 τ-Leptons

The energy calibration and efficiency scale factors of the τ_{had} reconstruction and identification are measured using $Z \to \tau\tau$ data, using the tag-and-probe method. The uncertainty on the offline τ_{had} identification efficiency is 5% (6%) for one-prong (three-prong) τ_{had} candidates. The τ_{had} energy scale uncertainty is 2% (3%) for one-prong

(three-prong) τ_{had} candidates. The probability of misidentifying an electron as a τ_{had} is associated with an uncertainty ranging from 3% to 14% [173]; this is measured in $Z \to ee$ data.

14.1.6 Jet Energy Scale and Resolution

The jet energy scale calibration (detailed in Section 9.3.2) corrects the reconstructed jet energy to match that of simulated truth jets. The associated systematic uncertainties, parametrised in p_{T} and η, are grouped into a reduced set of uncertainties with three nuisance parameters. This simplifies the correlations between the different sources of uncertainty. The JES uncertainty is largest for low-p_{T} jets – 4.5% for jets with $p_{\text{T}} = 20$ GeV – decreasing to 1% for jets with $200 < p_{\text{T}} < 1800$ GeV, and rising again to 2% for jets with higher p_{T} [160].

The jet energy resolution (JER) uncertainties are obtained by smearing the nominal JER, giving a total uncertainty between 10% and 20% [192].

The jet vertex tagger (JVT), described in Section 9.3.2, is a multivariate algorithm which vetoes jets which may arise from pile-up events by applying a threshold to the algorithm output. This threshold is varied to produce the associated uncertainty.

14.1.7 Flavour Tagging

Scale factors are used to correct the b-tagging efficiency (and mistag rate for c- and light-flavour jets) in simulation to match data (explained in detail in Section 9.4.2). These are calculated separately for b-, c- and light-jets and are parametrised in jet p_{T} and η. The three calibrations have multiple sources of uncertainty; these are simplified to form uncorrelated sets which are treated independently, with three separate uncertainties for b-jets, four for c-jets and five for light-jets [165].

The b-jet calibration and its associated uncertainties are taken from fully-leptonic $t\bar{t}$ events. For jets with $20 < p_{\text{T}} < 200$ GeV, the uncertainty is between 2% and 4%, rising to 12% for $p_{\text{T}} < 300$ GeV.

The mistag rates for c- and light-jets are derived using $t\bar{t}$ and multi-jet data events, respectively; the uncertainties on these measurements are derived from the variation of the scale factors. A further uncertainty on the c-jet scale factor originates from the difference observed between W + c-jet events and semi-leptonic $t\bar{t}$ decays. The total uncertainty on the c-jet mistag rate is 5-13% and the uncertainty on the light-jet mistag rate is 20-50%, depending on the b-tagging working point.

14.1.8 $E_{\mathrm{T}}^{\mathrm{miss}}$

Systematic uncertainties on the energy scale and resolution of the electrons, muons, jets and τ-leptons are propagated to the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$. Additionally, dedicated uncertainties are included to account for uncertainty in the measurements of the scale, resolution and reconstruction efficiency of tracks that are not associated to any reconstructed objects, as well as the modelling of the underlying event [174].

14.2 Background Modelling Uncertainties

The systematic uncertainties associated with the modelling of each background process are described in Chapter 12.

14.3 Theoretical Uncertainties on the Signal Models

The cross-section for non-resonant SM Higgs pair-production, taken from [28, 37], is given by

$$\sigma_{\mathrm{SM}} = 33.49^{+1.44}_{-2.00}(\mathrm{scale}) \pm 1.67(\mathrm{theory}) \pm 0.70(\mathrm{PDF}) \pm 0.77(\alpha_s)\ \mathrm{fb}. \tag{14.1}$$

The uncertainties are added in quadrature and are applied as a normalisation uncertainty of 8%.

Theoretical uncertainties on the signal acceptance for the non-resonant and resonant searches are applied to cover variations in the parton density functions, renormalisation and factorisation scales, and modelling of the parton shower and underlying event. The renormalization and factorization scales used in the signal generation are varied by factors of 0.5 and 2, with the combined uncertainty resulting from the envelope of all of the uncertainties. The uncertainty due to the choice of PDF set is calculated by comparing the nominal PDF set to the PDF4LHC set [185]. The parton shower uncertainties are calculated by comparing Herwig++ with Pythia 8.

The parton density function uncertainties are negligible for all signal hypotheses. For non-resonant SM Higgs pair-production, the renormalisation and factorisation scale uncertainties are negligible, whilst the uncertainty on the parton shower modelling is 5%. For both resonant signals, the renormalisation and factorisation scale uncertainty and the parton shower uncertainty add in quadrature to give 12%.

14.4 Summary of Systematic Uncertainties

The parameter of interest calculated by the fit described in Section 5.1, i.e. the signal strength, μ, can change with respect to each systematic uncertainty (or nuisance parameter). Figure 14.1 shows the *fractional impact*, $\Delta\mu/\Delta\mu_{\text{tot}}$, of the systematic uncertainties on the non-resonant SM Higgs pair-production signal strength, μ. The uncertainties are listed in decreasing order of their impact on μ, with only the fiften most highly-ranked nuisance parameters shown. The boxes, which refer to the top x-axis, show $\Delta\mu/\Delta\mu_{\text{tot}}$ when fixing the corresponding nuisance parameter, θ, to its fitted value, $\hat{\theta}$, modified up or down by its fitted uncertainty, and performing the fit again, allowing all other parameters to vary such that correlations between systematics can be taken into account. The hatched and open areas refer to the impact of the up and down variations, respectively. The two parameters with the greatest impact on the non-resonant fit are the background normalisations associated with the $t\bar{t}$ and $Z \to \tau\tau$ + heavy-flavour backgrounds, though this is not the case for all signal hypotheses. The systematic uncertainties that have a consistently high impact for all signals (resonant and non-resonant) are those related to the $t\bar{t}$ and $Z \to \tau\tau$ + heavy-flavour backgrounds, the reconstruction and identification of hadronic τ-leptons, and the flavour-tagging efficiencies.

The filled circles, to be read from the bottom x-axis, show the deviations (or *pulls*) of the fitted nuisance parameters, $\hat{\theta}$, from their nominal values, θ_0, as a fraction of the standard deviation, $\Delta\theta$. The associated error bars show the fitted uncertainties of the nuisance parameters, relative to their nominal uncertainties. The open circles and their error bars, also read from the bottom axis, show the fitted value and uncertainty of the $t\bar{t}$ and $Z \to \tau\tau$ + heavy-flavour normalisation parameters which are floated freely in the fit. For the non-resonant signal, the nuisance parameters showing the largest deviations from their nominal values are the $t\bar{t}$ shape uncertainties, whilst no nuisance parameters are pulled more than expected for any signal hypothesis.

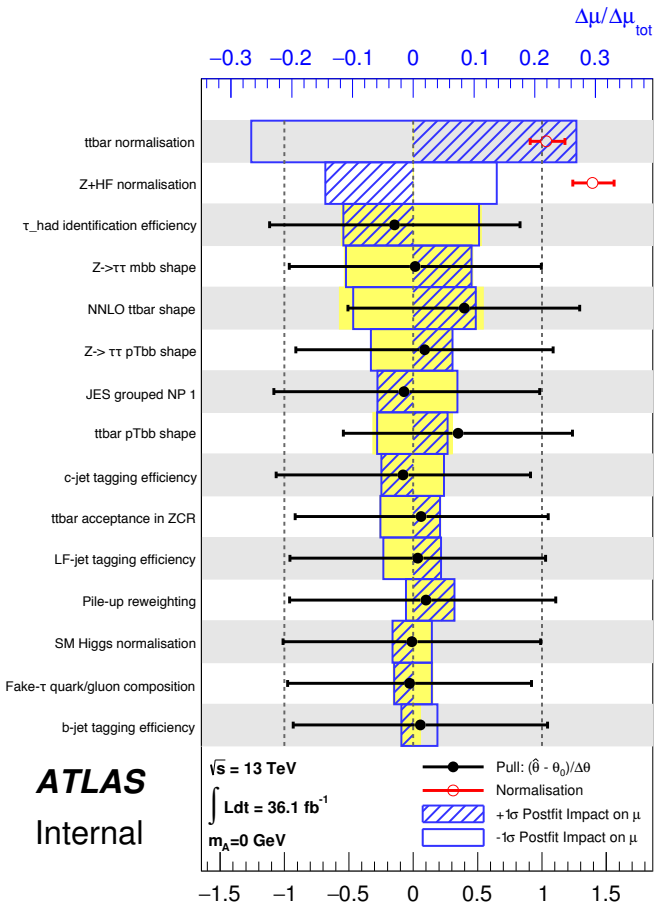


FIGURE 14.1: Nuisance parameter rankings for the non-resonant SM di-Higgs fit.

Chapter 15

Results

As explained in Section 5, the statistical interpretation of the results is performed using a likelihood fit. In the absence of a signal observation, results are presented as 95% confidence level upper limits.

Results are presented for the $\tau_{\text{lep}}\tau_{\text{had}}$ channel in Section 15.2, followed by the combined $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ results in Section 15.3. Section 15.4 presents the results of the ATLAS combination of the $hh \to b\bar{b}b\bar{b}$, $hh \to b\bar{b}\gamma\gamma$ and $hh \to b\bar{b}\tau^+\tau^-$ channels.

15.1 Fitting Procedure and Presentation of Results

The BDT output score is used as the discriminating variable in the fit for all signal models. The binning of the distribution used in the final fit can affect the results; finer binning increases the sensitivity to signal, whilst also increasing the statistical uncertainties associated with each bin. Therefore, the binning of the BDT score distributions is optimised for each signal scenario (separately for each mass hypothesis) by ensuring that the events in each bin (from right to left) obey the following:

$$\frac{\sigma_B^i}{N_B^i} < x\frac{N_S^i}{N_S^{\text{tot}}} \quad \text{and} \quad N_B^i < y, \tag{15.1}$$

where N_S^i and N_B^i are the number of signal and background events in the i^{th} bin, respectively, and N_S^{tot} is the total number of signal events. This ensures that the statistical uncertainty on the background in the i^{th} bin, σ_B^i, remains below a fraction, x, ($x = 0.2$ for the resonant signal and $x = 0.4$ for the non-resonant signal) of the signal fraction, N_S^i/N_S^{tot}, while keeping a minimum of y events per bin ($y = 10$ for all signal hypotheses). For bins with no signal, the statistical uncertainty is kept below 1%.

There are three parameters floated in the final fit:

- the signal strength, μ,
- the true-τ_{had} $t\bar{t}$ normalisation,
- the $Z \to \tau\tau$ + heavy-flavour normalisation.

As well as the BDT output score (the signal region), the $Z \to \mu\mu$ + heavy-flavour jets control region is included (as a single bin) in the fit in order to derive the normalisation factor for the $Z \to \tau\tau$ + heavy-flavour background processes (see Section 12.3). The $t\bar{t}$ normalisation is constrained by the low BDT score region of the BDT distribution (see Section 12.1).

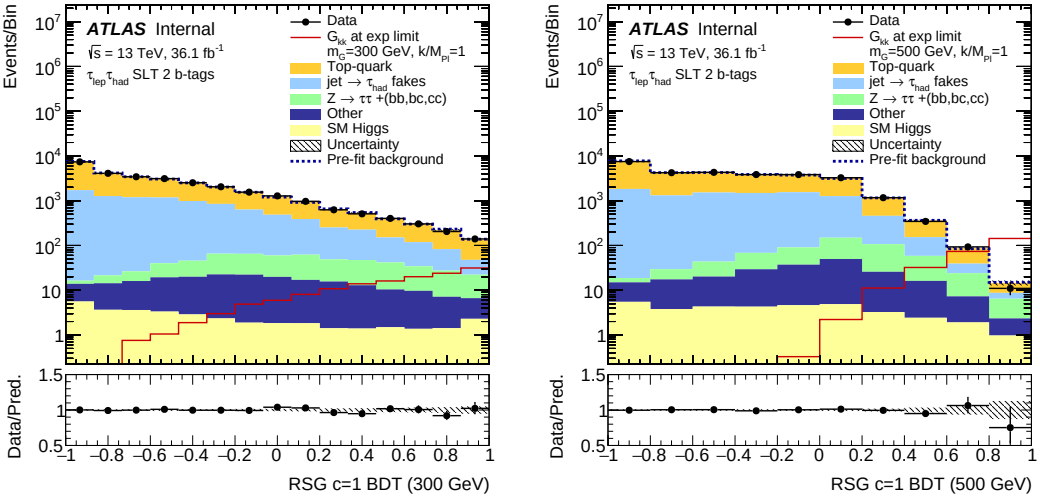
Using a profile likelihood test (as outlined in Section 5.1, the observed number of events is found to be compatible with the fitted number of background events (i.e. no significant excess over the SM background is observed) for all signal hypotheses. Therefore, the data are used to set upper limits at the 95% confidence level, following the CL_s prescription (see Section 5.2).

15.2 $\tau_{\text{lep}}\tau_{\text{had}}$ SLT Results

Figures 15.1, 15.2, 15.3 and 15.4 show the BDT output distributions for an RS graviton with $k/\bar{M}_{\text{Pl}} = 1.0$, an RS graviton with $k/\bar{M}_{\text{Pl}} = 2.0$, a generic narrow width scalar, and non-resonant signals respectively, after performing the fit assuming a background-only hypothesis.

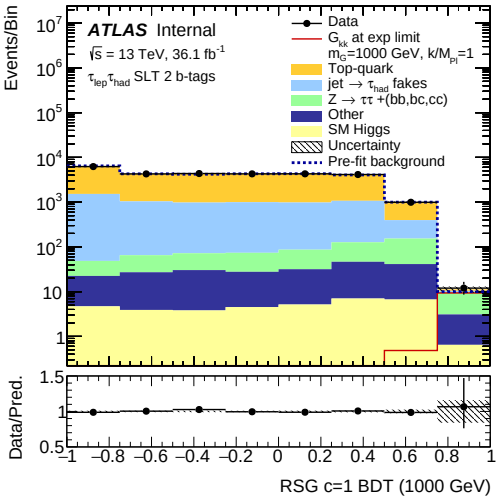
The expected number of signal and background events and the observed number of data events after applying the selection criteria in Section 11.2, requiring exactly two b-jets, and performing the fit assuming a background-only hypothesis are given in Table 15.1.

The 95% confidence level upper limits on $\sigma_{H,G} \times \text{BR}_{H,G\to hh\to bb\tau_{\text{lep}}\tau_{\text{had}}}$ as a function of the resonant signal mass are shown in Figure 15.5 for RS gravitons with $k/\bar{M}_{\text{Pl}} = 1.0$, RS gravitons with $k/\bar{M}_{\text{Pl}} = 2.0$, and generic narrow-width scalars with $\tan\beta = 2.0$. For scalar resonances, the results are interpreted in a simplified minimal supersymmetric model, the hMSSM [49, 52], which is summarised in Section 2.4. The observed (expected) range in which RS gravitons with $k/\bar{M}_{\text{Pl}} = 1.0$ are excluded is 335 (348) $< m_G <$ 868 (850) GeV. RS gravitons with $k/\bar{M}_{\text{Pl}} = 2.0$ are excluded between $260 < m_G < 1000$ GeV but heavy scalar Higgs bosons in the hMSSM are not excluded for any masses between $260 < m_H < 1000$ GeV.



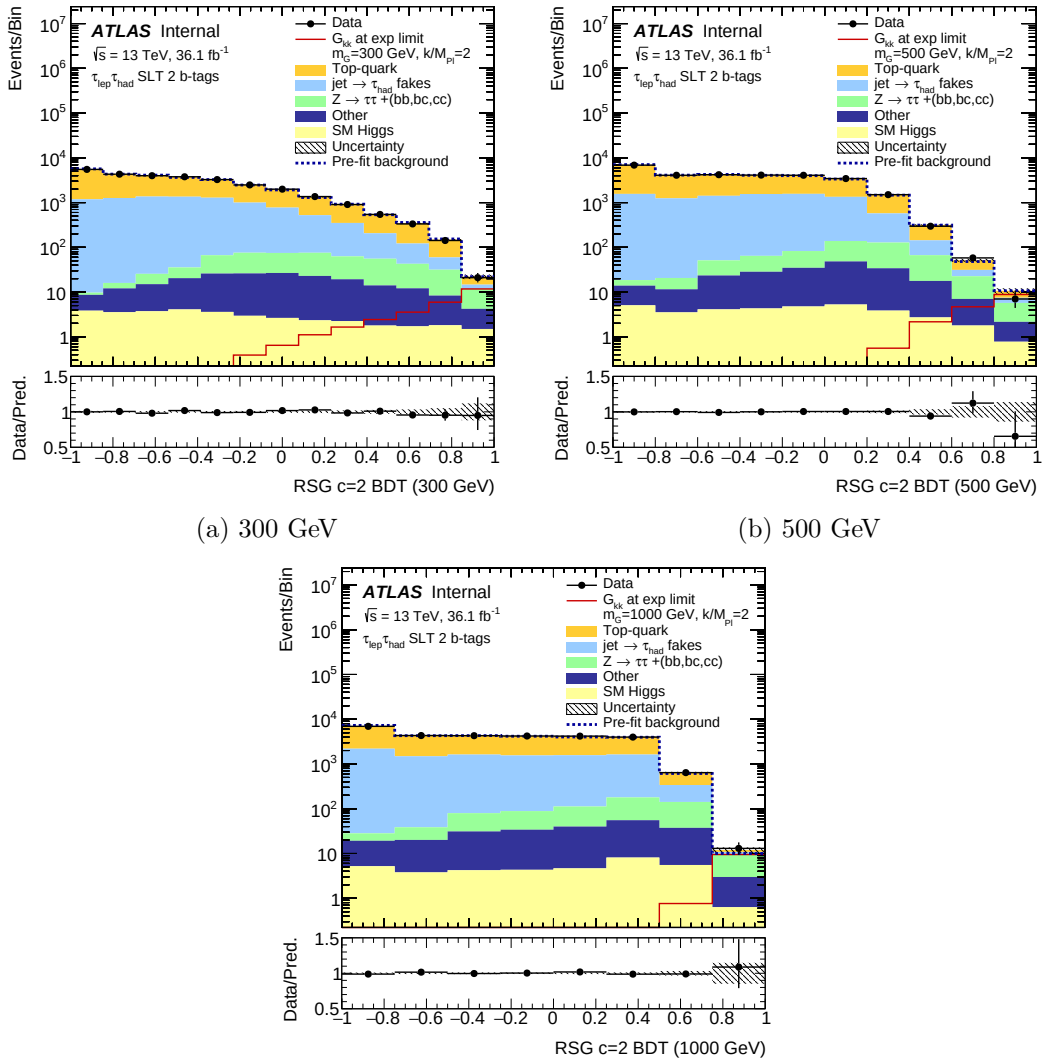
(a) 300 GeV

(b) 500 GeV



(c) 1000 GeV

FIGURE 15.1: Post-fit BDT output distributions for the RSG signal with $k/\bar{M}_{\mathrm{Pl}} = 1.0$ with (a) $m_G = 300$ GeV, (b) $m_G = 500$ GeV and (c) $m_G = 1000$ GeV. 'Top-quark' refers to $t\bar{t}$ processes involving a real τ_{had} and single-top processes; 'jet→τ_{had} fakes' refers to all processes where a jet fakes a τ_{had}, i.e. QCD multi-jet, $t\bar{t}$ and W + jets; the 'SM Higgs' background combines associated Vh production processes and tth processes; backgrounds included in 'Other' are $Z \to \ell\ell$, $Z \to \tau\tau$ + light-flavour jets, W boson decays involving a real τ_{had}, and di-boson processes involving pairs of Z and W bosons.



(a) 300 GeV

(b) 500 GeV

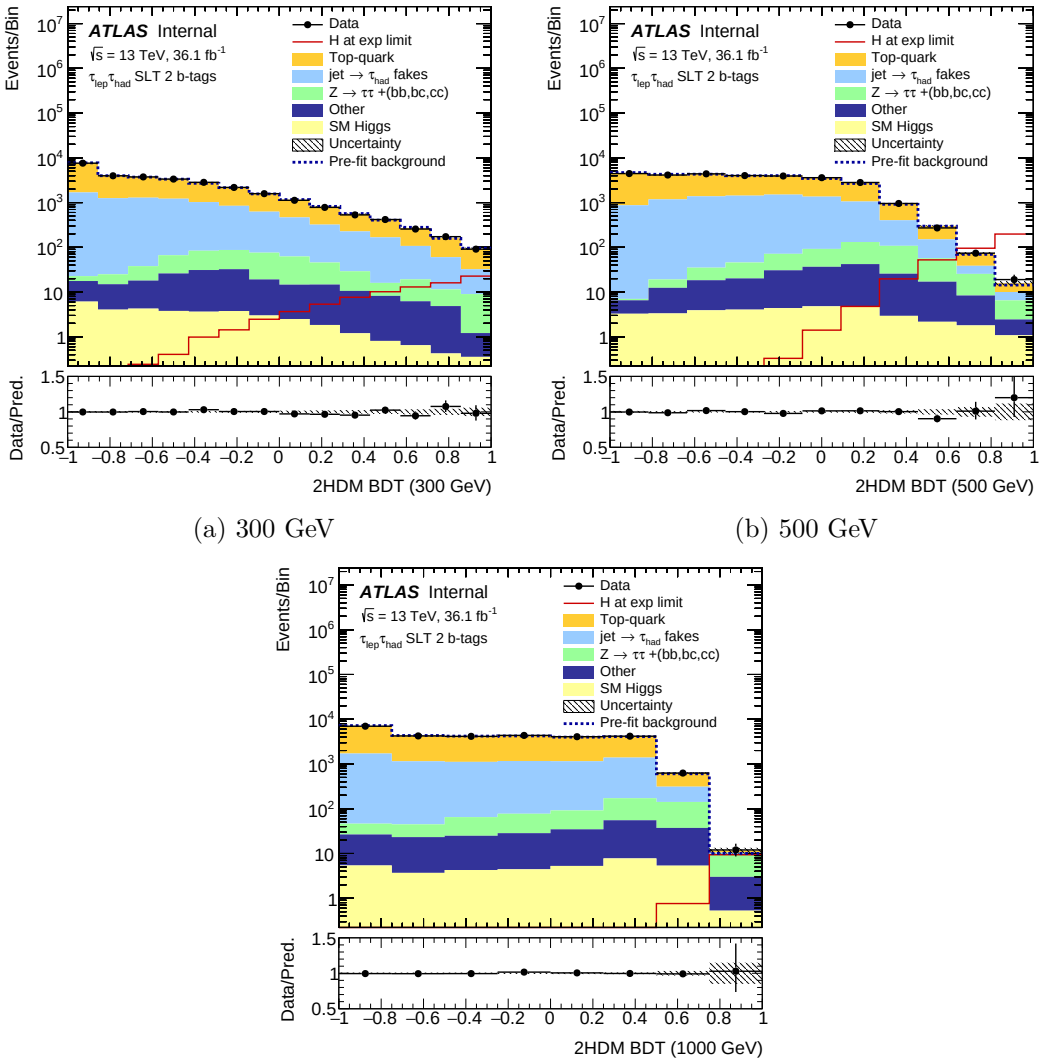
(c) 1000 GeV

FIGURE 15.2: Post-fit BDT output distributions for the RSG signal with $k/\bar{M}_{\mathrm{Pl}} = 2.0$ with (a) $m_G = 300$ GeV, (b) $m_G = 500$ GeV and (c) $m_G = 1000$ GeV.

The 95% confidence level upper limits on the non-resonant di-Higgs production cross-section are given in Table 15.4. The observed (expected) limit is 21.34 (29.33) times the SM cross-section.

15.3 Combined $b\bar{b}\tau^+\tau^-$ Results

The analysis presented in [193] provides a combination of the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ di-τ decay modes (the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ decay is selected using the single-τ and di-τ triggers). The $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel also includes events selected by the lepton-plus-τ_{had} trigger (LTT), which are analysed separately and combined with those selected by the SLTs. The LTT allows the inclusion of data with lower lepton transverse momentum, p_{T}^{ℓ}. The



(a) 300 GeV

(b) 500 GeV

(c) 1000 GeV

FIGURE 15.3: Post-fit BDT output distributions for the generic narrow-width scalar signal with (a) $m_H = 300$ GeV, (b) $m_H = 500$ GeV and (c) $m_H = 1000$ GeV.

three regions ($\tau_{\text{lep}}\tau_{\text{had}}$ SLT, $\tau_{\text{lep}}\tau_{\text{had}}$ LTT and $\tau_{\text{had}}\tau_{\text{had}}$) are all included in the combined fit, alongside the single-bin $Z \to \mu\mu$ + heavy-flavour jets control region. Table 15.3 outlines the extrapolation uncertainties applied to the $t\bar{t}$ and $Z \to \tau\tau$ + heavy-flavour normalisations in all regions of the combined fit: the $t\bar{t}$ normalisation is derived in the $\tau_{\text{lep}}\tau_{\text{had}}$ (SLT and LTT) signal region and extrapolated to the $\tau_{\text{had}}\tau_{\text{had}}$ signal region, where it is allowed to vary from the value derived in the $\tau_{\text{lep}}\tau_{\text{had}}$ signal region within an uncertainty of $^{+30\%}_{-32\%}$.[1] The $Z \to \tau\tau$ + heavy-flavour normalisation, which is derived in the $Z \to \mu\mu$ + heavy-flavour control region, is extrapolated to the $\tau_{\text{had}}\tau_{\text{had}}$ signal region with an uncertainty of $\pm 35\%$, which is derived in a similar manner to the extrapolation to the $\tau_{\text{lep}}\tau_{\text{had}}$ channel described in Section 12.3.2.

[1] This is estimated following a similar procedure to that used to calculate the extrapolation uncertainty applied to the $t\bar{t}$ normalisation in the $Z \to \mu\mu$ + heavy-flavour CR; see Section 12.3.2.

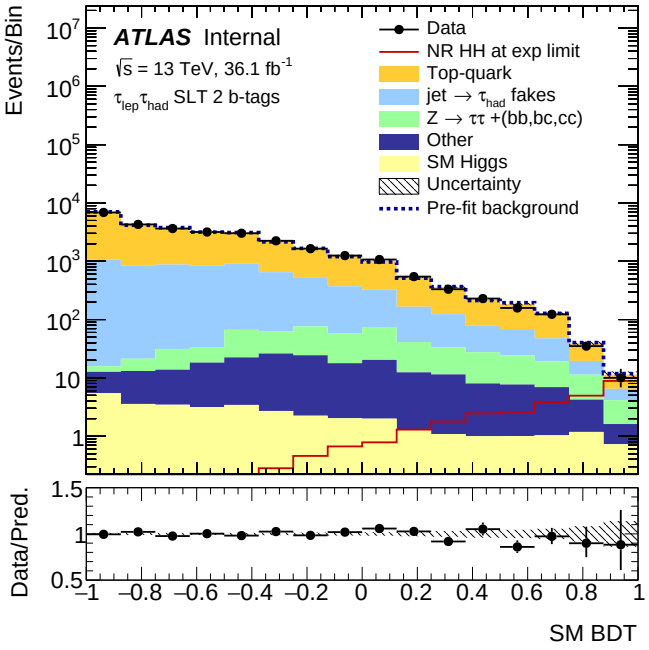
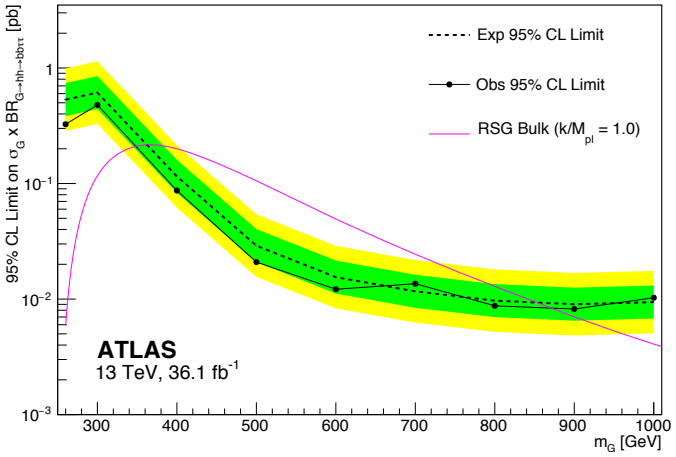


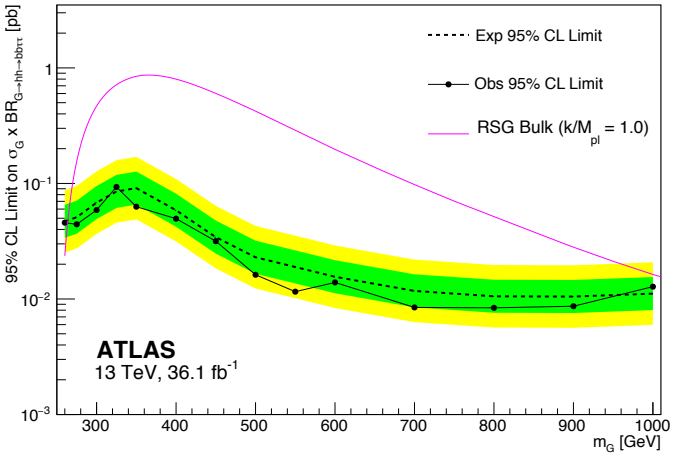
FIGURE 15.4: The post-fit BDT output distribution for the non-resonant SM di-Higgs signal.

Sample	Number of events
$t\bar{t}$	17800 ± 1100
$Single-top$	1130 ± 110
Fake τ_{had}	9000 ± 1100
$Z \to \tau\tau + (bb, bc, cc)$	416 ± 97
Other	197 ± 32
SM Higgs	38 ± 10
Total background	28610 ± 180
Data	28612
$G(m_G = 300$ GeV, $k/\bar{M}_{\text{Pl}} = 1.0)$	23.6 ± 3.7
$G(m_G = 500$ GeV, $k/\bar{M}_{\text{Pl}} = 1.0)$	42.4 ± 6.4
$G(m_G = 1000$ GeV, $k/\bar{M}_{\text{Pl}} = 1.0)$	2.6 ± 0.4
$G(m_G = 300$ GeV, $k/\bar{M}_{\text{Pl}} = 2.0)$	327 ± 50
$G(m_G = 500$ GeV, $k/\bar{M}_{\text{Pl}} = 2.0)$	193 ± 29
$G(m_G = 1000$ GeV, $k/\bar{M}_{\text{Pl}} = 2.0)$	8.6 ± 1.3
$H(m_H = 300$ GeV)	39.1 ± 6.3
$H(m_H = 300$ GeV)	3.41 ± 0.52
$H(m_H = 300$ GeV)	0.0267 ± 0.0041
Non-resonant hh	0.99 ± 0.13

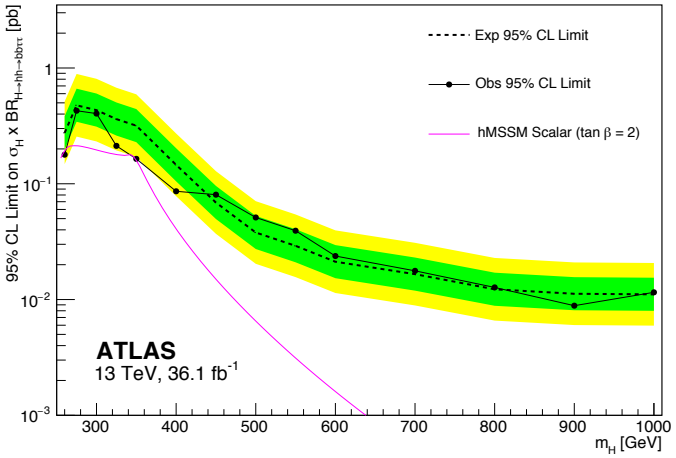
TABLE 15.1: The expected number of signal and background events and the observed number of data events after applying the selection criteria in Section 11.2, requiring exactly two b-jets, and performing the background-only fit. The background labelled 'Fake τ_{had}' includes all processes ($t\bar{t}$, QCD multi-jet and W + jets) in which a jet is misidentified as a τ_{had}. The total background yield is not equal to the sum of the individual backgrounds due to rounding. Individual uncertainties can be larger than the total uncertainty due to the large correlations.



(a) Graviton G, $k/\bar{M}_{\mathrm{Pl}} = 1.0$



(b) Graviton G, $k/\bar{M}_{\mathrm{Pl}} = 2.0$



(c) Scalar H

FIGURE 15.5: 95% confidence level upper limits on $\sigma_{H,G} \times \mathrm{BR}_{H,G\to hh\to bb\tau_{\mathrm{lep}}\tau_{\mathrm{had}}}$ as a function of the resonant signal mass for (a) an RS graviton with $k/\bar{M}_{\mathrm{Pl}} = 1.0$, (b) an RS graviton with $k/\bar{M}_{\mathrm{Pl}} = 2.0$, or (c) a generic narrow-width scalar particle. The expected limit is shown as a dotted line, the observed limit is shown as a solid black line, and the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the expected limit are shown by green and yellow bands, respectively. The theoretical prediction is shown as a solid pink line.

	Observed	-2σ	-1σ	Expected	$+1\sigma$	$+2\sigma$
σ [pb]	0.71	0.53	0.71	0.98	1.36	1.83
$\sigma/\sigma_{\text{SM}}$	21.34	15.74	21.14	29.33	40.82	54.73

TABLE 15.2: The 95% CL upper limits on the non-resonant di-Higgs production cross-section, σ, for the $\tau_{\text{lep}}\tau_{\text{had}}$ channel.

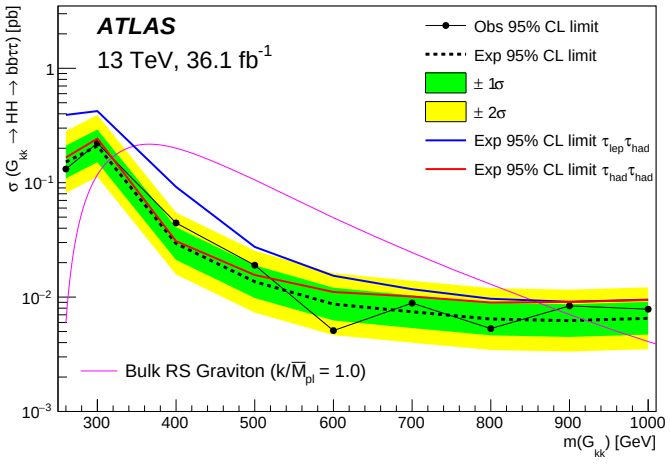
Region	Input	Extrapolation uncertainty	
		$t\bar{t}$	$Z \to \tau\tau$ + heavy-flavour
$\tau_{\text{lep}}\tau_{\text{had}}$ SLT (signal)	BDT	Derived	$\pm 29\%$
$\tau_{\text{lep}}\tau_{\text{had}}$ LTT (signal)	BDT	Derived	$\pm 29\%$
$\tau_{\text{had}}\tau_{\text{had}}$	BDT	$^{+30}_{-32}\%$	$\pm 35\%$
$Z \to \mu\mu$ + heavy-flavour (control)	Yield	$^{+8}_{-9}\%$	Derived

TABLE 15.3: The uncertainties applied to the normalisation values to account for the extrapolation of the normalisation from the region in which it was derived to the other regions included in the combined fit.

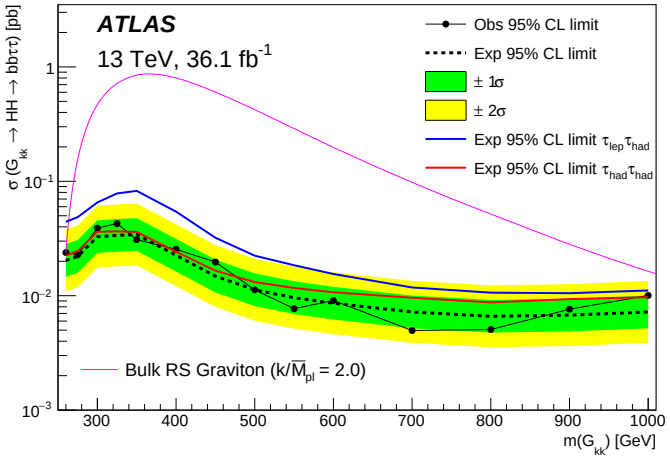
After combining the channels, no excess is seen and the results are presented as 95% confidence level upper limits on the Higgs pair-production cross-section times the $hh \to b\bar{b}\tau^+\tau^-$ branching ratio. The limits for a narrow-width scalar resonance, H, and RS gravitons, G, with $k/\bar{M}_{\text{Pl}} = 1.0, 2.0$ are given as a function of resonance mass in Figure 15.6. For a heavy scalar Higgs boson in the hMSSM, the mass range $305 < m_H < 402$ GeV is excluded at the 95% confidence level for $\tan\beta = 2.0$. RS gravitons are excluded in the range $325 < m_G < 885$ GeV for $k/\bar{M}_{\text{Pl}} = 1.0$; for $k/\bar{M}_{\text{Pl}} = 2.0$, the entire mass range of this search ($260 < m_G < 1000$ GeV) can be excluded. By comparing Figure 15.6a with 15.6b, it can be seen that the limits are largely insensitive to the value of $k/\bar{M}_{\text{Pl}}$ above ~ 600 GeV.

The individual limits for the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels are shown alongside the combined limits. At lower resonant masses, the $\tau_{\text{had}}\tau_{\text{had}}$ channel is significantly more sensitive to a resonant signal; at higher masses, the limits are much closer. This is, in part, due to the larger $t\bar{t}$ background in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel.

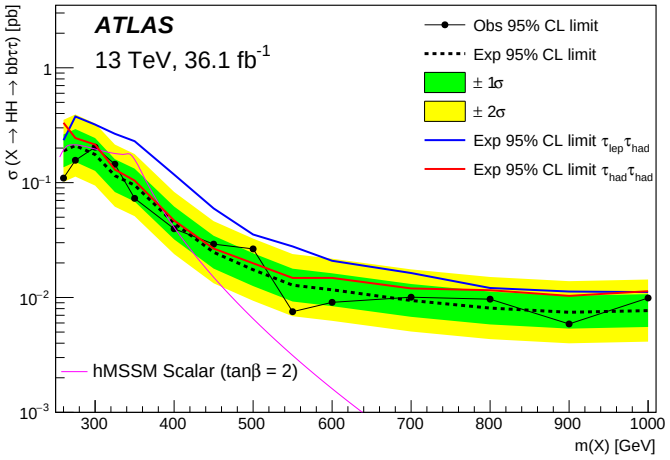
Observed and expected 95% confidence level upper limits on non-resonant SM Higgs pair-production cross-section for events decaying to a $b\bar{b}\tau^+\tau^-$ final state are presented in Table 15.4. These are presented as a cross-section limit and as a ratio to the cross-section predicted by the SM. The observed (expected) limit is 12.7 (14.8) times the SM cross-section.



(a) Graviton G, $k/\bar{M}_{\mathrm{Pl}} = 1.0$



(b) Graviton G, $k/\bar{M}_{\mathrm{Pl}} = 2.0$



(c) Scalar H

FIGURE 15.6: Observed (solid black) and expected (dotted black) limits at 95% confidence level on the cross-sections of (a) an RS graviton with $k/\bar{M}_{\mathrm{Pl}} = 1.0$, (b) an RS graviton with $k/\bar{M}_{\mathrm{Pl}} = 2.0$ and (c) a generic narrow-width scalar interpreted in the hMSSM model. The $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the expected limit are shown by green and yellow bands, respectively. The scalar, labelled H throughout this thesis, is labelled X in this figure.

Channel	Units	Observed	-2σ	-1σ	Expected	$+1\sigma$	$+2\sigma$
$\tau_{\text{lep}}\tau_{\text{had}}$ SLT	σ [pb]	0.71	0.53	0.71	0.98	1.36	1.83
	$\sigma/\sigma_{\text{SM}}$	21.3	15.7	21.1	29.3	40.8	54.7
$\tau_{\text{lep}}\tau_{\text{had}}$ LTT	σ [pb]	4.47	1.68	2.26	3.14	4.37	5.85
	$\sigma/\sigma_{\text{SM}}$	134	50.4	67.7	93.9	131	175
$\tau_{\text{lep}}\tau_{\text{had}}$ combined	σ [pb]	0.79	0.51	0.68	0.95	1.32	1.77
	$\sigma/\sigma_{\text{SM}}$	23.5	15.2	20.5	28.4	39.5	53.0
$\tau_{\text{had}}\tau_{\text{had}}$	σ [pb]	0.52	0.31	0.41	0.57	0.80	1.07
	$\sigma/\sigma_{\text{SM}}$	15.5	9.21	12.4	17.2	23.9	32.0
Combined	σ [pb]	0.42	0.26	0.36	0.49	0.69	0.92
	$\sigma/\sigma_{\text{SM}}$	12.7	7.93	10.7	14.8	20.6	27.6

TABLE 15.4: Observed and expected 95% CL upper limits on the non-resonant di-Higgs production cross-section, σ, and their ratios to the value predicted by the SM, for the $b\bar{b}\tau^+\tau^-$ final state (combining the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels). Additionally, $\pm1\sigma$ and $\pm2\sigma$ variations about the expected limit are shown.

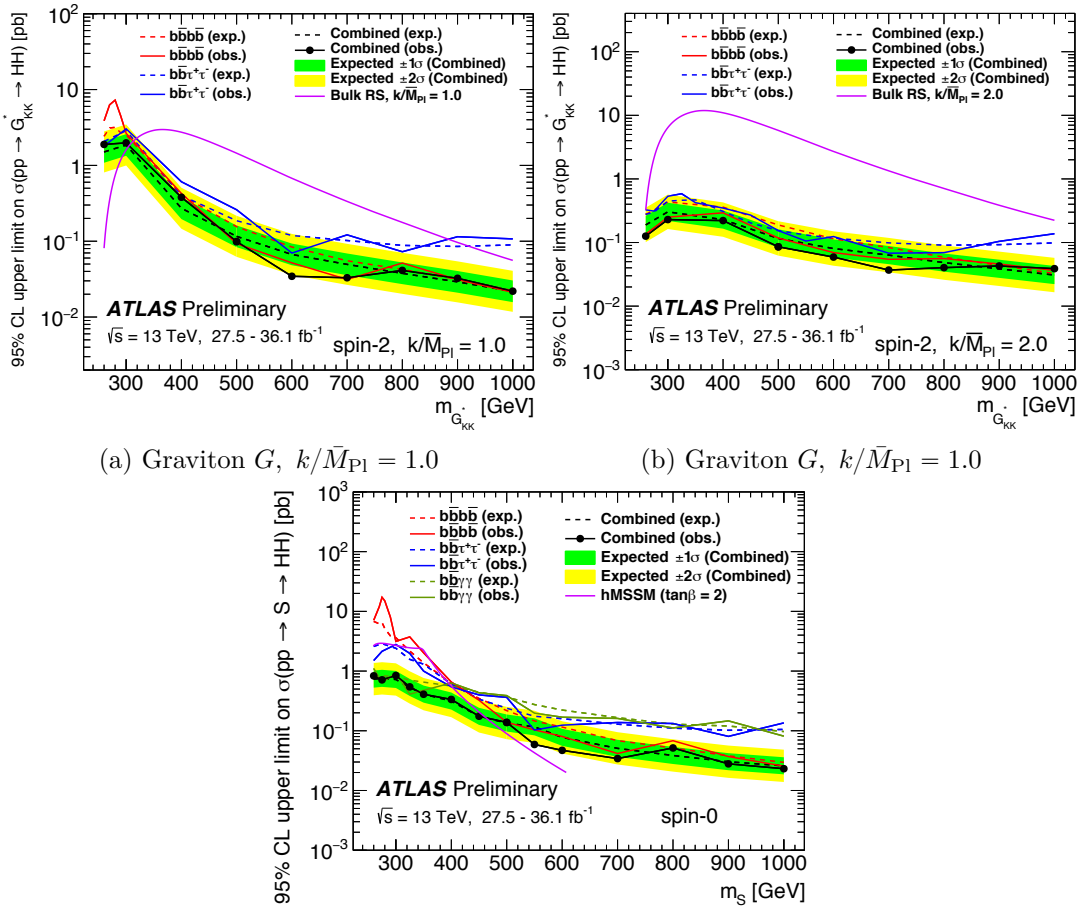
15.4 Combined ATLAS Results

As outlined in Section 2.4.1, the other main decay channels utilised by the ATLAS experiment in searching for Higgs pairs are $hh \to b\bar{b}b\bar{b}$ and $hh \to b\bar{b}\gamma\gamma$. A combination of these channels alongside $hh \to b\bar{b}\tau^+\tau^-$ is presented in [194].

No statistically significant excess is observed across the probed resonance mass range. The results are presented as 95% confidence level upper limits on the Higgs pair-production cross-section, σ, for a narrow width scalar resonance, H (referred to as S in the relevant plots), and an RS graviton, G, with $k/\bar{M}_{\text{Pl}} = 1.0, 2.0$. These are shown in Figure 15.7; the RS graviton results are a combination of only the $hh \to b\bar{b}b\bar{b}$ and $hh \to b\bar{b}\tau^+\tau^-$ searches.

RS gravitons with $k/\bar{M}_{\text{Pl}} = 1.0$ are excluded for $m_G > 307$ GeV, with an upper bound set by the $b\bar{b}b\bar{b}$ analysis of $m_G < 1362$ GeV. RS gravitons with $k/\bar{M}_{\text{Pl}} = 2.0$ are excluded for the entire mass range for which the combination is performed, $260 < m_G < 1000$ GeV. Again, the upper bound is provided by the $b\bar{b}b\bar{b}$ search where no combination is performed, excluding $m_G < 1744$ GeV [194]. In the hMSSM, for $\tan\beta = 2.0$ the scalar resonance is excluded in the mass range $260 < m_H < 462$ GeV.

In each case, the $b\bar{b}\tau^+\tau^-$ search provides a greater contribution to the combined sensitivity at lower resonant masses, $m_{H,G}$. In the cases of the RS graviton signal with $k/\bar{M}_{\text{Pl}} = 1.0$ and the narrow-width scalar resonance, the $b\bar{b}\tau^+\tau^-$ sensitivity at low mass is greater than that of the $b\bar{b}b\bar{b}$ search, which provides the greatest sensitivity at high mass.



(a) Graviton G, $k/\bar{M}_{\rm Pl} = 1.0$

(b) Graviton G, $k/\bar{M}_{\rm Pl} = 1.0$

(c) Scalar H

FIGURE 15.7: Observed (solid black) and expected (dotted black) limits at 95% confidence level on the cross-sections of (a) an RS graviton with $k/\bar{M}_{\rm Pl} = 1.0$, (b) an RS graviton with $k/\bar{M}_{\rm Pl} = 2.0$ and (c) a generic narrow-width scalar interpreted in the hMSSM. The $\pm1\sigma$ and $\pm2\sigma$ uncertainties on the expected limit are shown by green and yellow bands, respectively. The scalar, labelled H throughout this thesis, is labelled S in this figure. The RS graviton limits are the result of the combination of the $hh \to b\bar{b}\tau^+\tau^-$ and $hh \to b\bar{b}b\bar{b}$ channels; the scalar limits also include the $hh \to b\bar{b}\gamma\gamma$ channel. Figures taken from [194].

Figure 15.8 displays the expected and observed 95% CL upper limits on the non-resonant Higgs boson pair-production signal strength in units of the SM $gg \to hh$ cross-section. Limits from the $hh \to b\bar{b}\tau^+\tau^-$, $hh \to b\bar{b}b\bar{b}$ and $hh \to b\bar{b}\gamma\gamma$ searches and their statistical combination are shown with $\pm1\sigma$ and $\pm2\sigma$ uncertainty bands. The strongest limit is that from the $b\bar{b}\tau^+\tau^-$ channel.

All three channels report a deficit of data with respect to the background-only prediction, particularly the $b\bar{b}b\bar{b}$ channel where the observed (expected) limit is 12.9 (20.7) $\times\, \sigma_{\rm SM}$. Therefore, the channels combine to give an observed upper limit which is stronger than expected, but within the 2σ uncertainty band; the combined observed (expected) limit is 6.7 (10.4) $\times\, \sigma_{\rm SM}$, which corresponds to 0.22 pb (0.35 pb).

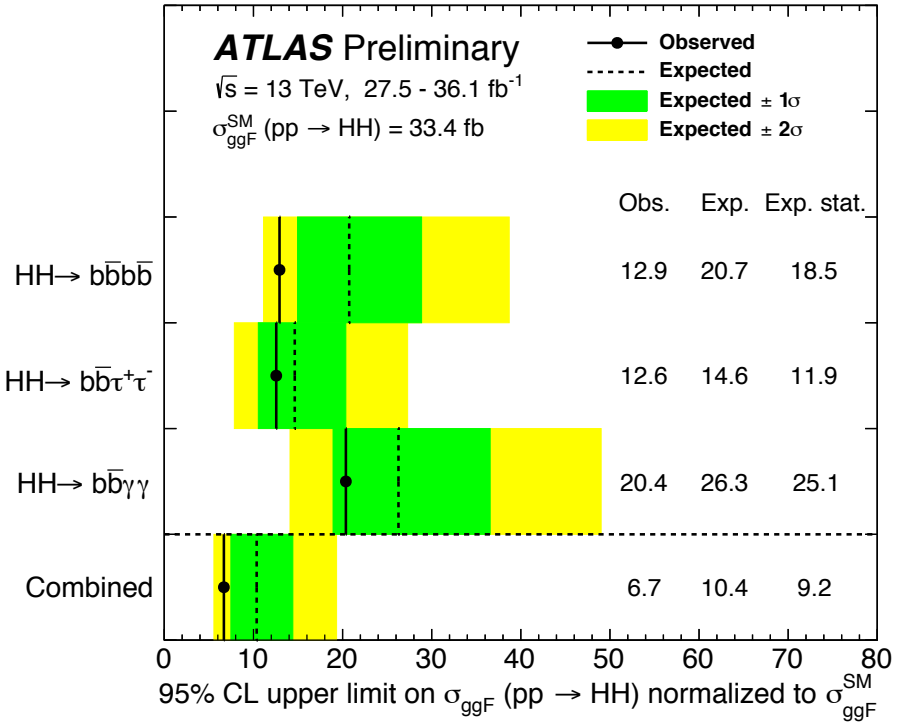


FIGURE 15.8: Upper limits at 95% CL on the non-resonant Higgs boson pair-production cross-section from the $hh \to b\bar{b}\tau^+\tau^-$, $hh \to b\bar{b}b\bar{b}$ and $hh \to b\bar{b}\gamma\gamma$ searches and their statistical combination. The expected limits are shown by dotted lines, with $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties shown by green and yellow bands, respectively. The solid black lines are the observed limits. The column labelled 'obs.' represents the observed limits, 'exp.' the expected limits with all statistical and systematic uncertainties, and 'exp. stat.' the expected limits obtained with statistical uncertainties only. Figure taken from [194].

The combined non-resonant limits are also presented as a scan over the value of $\kappa_\lambda = \lambda/\lambda_{\text{SM}}$, where λ is the trilinear Higgs self-coupling defined in Section 2.2.3.

All couplings except λ are set to their SM values, with kinematic distributions computed at the generator level for each value of κ_λ in the range $-20 < \kappa_\lambda < 20$. These are produced by linearly combining three LO samples to produce samples for any value of κ_λ, as described in [194]. All samples are analysed using a new BDT training with $\kappa_\lambda = 20$.

The results of the scan are presented as expected and observed 95% CL upper limits on the non-resonant Higgs boson pair-production cross-section as a function of κ_λ in the range $-20 < \kappa_\lambda < 20$. A combination of the $hh \to b\bar{b}\tau^+\tau^-$, $hh \to b\bar{b}b\bar{b}$ and $hh \to b\bar{b}\gamma\gamma$ channels is presented in Figure 15.9. The allowed values of κ_λ based on the combined observed limits at 95% CL are in the interval $-5.0 < \kappa_\lambda < 12.1$.

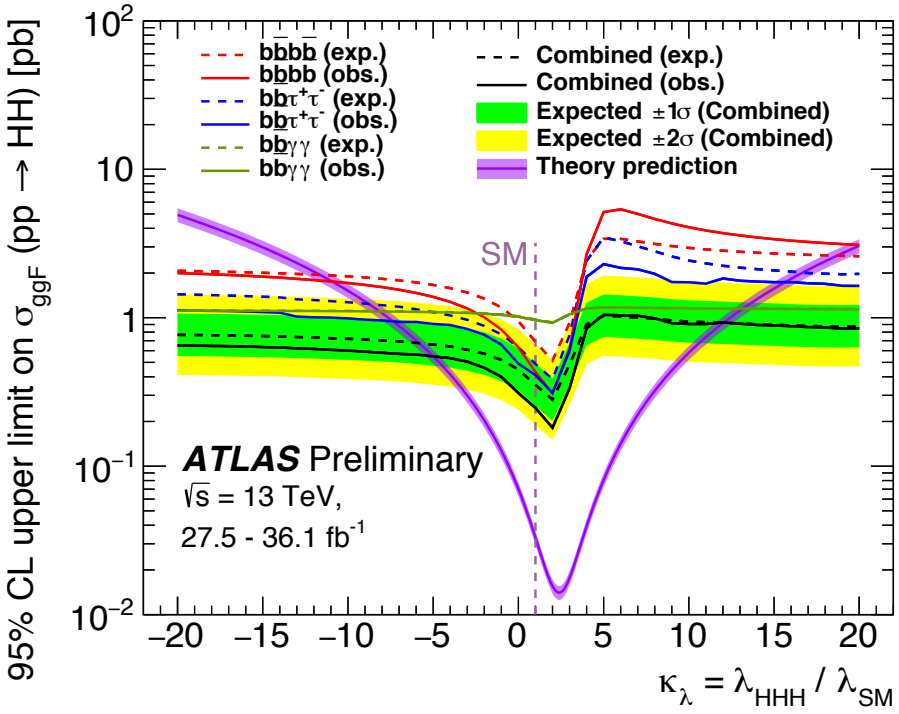


FIGURE 15.9: 95% CL upper limits on the $gg \to hh$ cross-section as a function of κ_λ. Expected (observed) limits are shown as dashed (solid) lines. The $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the expected limits are shown by green and yellow bands, respectively. The $b\bar{b}b\bar{b}$ limits are shown in red, $b\bar{b}\tau^+\tau^-$ in blue, $b\bar{b}\gamma\gamma$ in green, and the combined limit in black. The $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands are shown only for the combined limit. The theory prediction is shown as a purple line with purple uncertainty band. Figure taken from [194].

15.5 Conclusion

In this section, results are presented for searches for resonant and non-resonant Higgs pair production for the $hh \to b\bar{b}\tau_{\text{lep}}\tau_{\text{had}}$ decay channel, which are also combined with the $hh \to b\bar{b}\tau_{\text{lep}}\tau_{\text{had}}$ channel. The $hh \to b\bar{b}\tau^+\tau^-$ results are followed by a summary of the combination of all di-Higgs decay channels studied by the ATLAS collaboration.

Whilst no significant excess is observed over the SM background in any channel, the 95% confidence level upper limits set by the $hh \to b\bar{b}\tau^+\tau^-$ analysis are shown to be competitive with the $hh \to b\bar{b}b\bar{b}$ search for all signal hypotheses, with the current world's most stringent upper limit on Higgs pair production in a single decay channel being set by the $hh \to b\bar{b}\tau^+\tau^-$ search. The observed (expected) upper limit on the di-Higgs production cross section is 12.7 (14.8) times the SM prediction.

The $hh \to b\bar{b}\tau^+\tau^-$ search in ATLAS outperforms the results achieved by the CMS experiment from a combination of all channels. The BDT introduced to improve the discrimination between signal and background was imperative in achieving these results.

Whilst the CMS $hh \to b\bar{b}\tau^+\tau^-$ search does utilise a BDT to separate signal and background, the BDT distribution is used to perform a cut on the data before using kinematic distributions as the final discriminant. This is in contrast to the method presented in this thesis, in which the BDT distribution is used in the statistical fit.

The $hh \to b\bar{b}\tau^+\tau^-$ has proven to be a promising channel to search for di-Higgs production and measure the trilinear self-coupling; this channel will be of great interest as the amount of ATLAS data increases. Prospects for di-Higgs searches at the High Luminosity LHC (HL-LHC) are outlined in [**?**].

Finally, for a heavy scalar Higgs boson in the hMSSM, the mass range $305 < m_H < 402$ GeV is excluded at the 95% confidence level for $\tan\beta = 2$. RS gravitons are excluded in the range $325 < m_G < 885$ GeV for $k/\bar{M}_{\mathrm{Pl}} = 1.0$; for $k/\bar{M}_{\mathrm{Pl}} = 2.0$, the entire mass range of the search ($260 < m_G < 1000$ GeV) can be excluded. These results suggest that future graviton searches might concentrate on low values of $k/\bar{M}_{\mathrm{Pl}}$ for lower m_G, or higher masses for a larger range of $k/\bar{M}_{\mathrm{Pl}}$.

Part IV

Search for Pair-Production of Third-Generation Scalar Leptoquarks Decaying to a $b\tau^{-}\bar{b}\tau^{+}$ Final State

Chapter 16

Analysis Strategy

Leptoquarks, as explained in Section 3.5, are encountered in a number of extensions to the Standard Model. This search focuses on pair-production of third-generation scalar leptoquarks, i.e. those that decay to third-generation SM particles.

The search for pair-produced leptoquarks decaying to $b\tau$ pairs is based on the search for pair-production of Higgs bosons, where one Higgs decays to a pair of b-quarks ($h \to b\bar{b}$) and the other to a pair of τ-leptons ($h \to \tau^+\tau^-$), which is the focus of Part III. Again, the search is performed for di-τ final states where one τ-lepton decays leptonically and the other hadronically ($\tau_{\text{lep}}\tau_{\text{had}}$) and where both τ-leptons decay hadronically ($\tau_{\text{had}}\tau_{\text{had}}$). The work performed by the author and described in this thesis is focused on the $\tau_{\text{lep}}\tau_{\text{had}}$ decay channel.

As explained in Section 3.5, third-generation leptoquarks can be either up-type ($\text{LQ}_3^u \to t\nu/b\tau$) or down-type ($\text{LQ}_3^d \to b\nu/t\tau$). The search is optimised for up-type leptoquarks with a branching ratio into charged leptons of $B = 1$, though it is sensitive to the down-type leptoquark decay channel due to the decay $\text{LQ}_3^d\overline{\text{LQ}_3^d} \to t\tau t\tau \to Wb\tau Wb\tau$, where the W bosons decay into jets. This decay chain is allowed because the analysis does not veto additional jets, but no optimisation has been performed to adapt the search for down-type leptoquarks. The search for up-type leptoquarks is performed for the mass range $200 < m(\text{LQ}_3^u) < 1500$ GeV, whilst the search for down-type leptoquarks is performed for the range $400 < m(\text{LQ}_3^d) < 1100$ GeV.

The analysis is presented in [195], alongside four reinterpretations of ATLAS searches for supersymmetric particles. The following chapters refer regularly to Part III, highlighting the differences between this search and the di-Higgs search.

16.1 Trigger and Event Selection

For the di-Higgs search, the triggers used in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel are the single-lepton triggers (SLTs) and the lepton-plus-τ trigger (LTT). The LTT is used in the di-Higgs search to provide increased sensitivity at low resonance mass, whereas for the leptoquark search, the low-mass range is less interesting as it has been excluded by previous searches. The SLTs and data cleaning are described in Section 11.1.

The event selection is similar to that in Section 11.2. Events are subject to the following requirements (using the object definitions in Section 9):

- exactly one electron passing the *Tight* identification criteria or exactly one muon passing the *Medium* identification criteria (with the additional requirement $|\eta^{\mu}| < 2.5$), with p_{T} required to be 1 GeV higher than the trigger threshold corresponding to the data-taking period of the event, such that the trigger efficiency reaches the plateau where the MC modelling improves;
- exactly one hadronic τ-lepton with $p_{\text{T}} > 25$ GeV and $|\eta| < 2.3$ (the ECAL crack region at $1.37 < |\eta| < 1.52$ is also vetoed);
- at least two jets, where the leading (subleading) jet fulfils $p_{\text{T}} > 60(20)$ GeV;
- no other electrons or muons (as defined in Sections 9.1 and 9.2) in the event;
- the light lepton and hadronic τ-lepton must have opposite sign charges;
- the invariant mass of the di-τ system, calculated using the MMC (as detailed in Section 9.7), $m_{\tau\tau}^{\text{MMC}} > 0$ GeV.

The requirement on $p_{\text{T}}^{\tau_{\text{had}}}$ is increased relative to the di-Higgs selection, as is the p_{T} requirement on the leading jet. This is because leptoquarks produce harder decays; the di-Higgs analysis was optimised to search for non-resonant production (in particular the triangle diagram) which produces low-p_{T} decay products. The MMC requirement of $m_{\tau\tau}^{\text{MMC}} > 0$ ensures that the MMC calculation has not failed. Whilst the di-τ mass is irrelevant in $b\tau$ decays, the MMC is used to produce the τ-lepton 4-vectors to calculate the variables used in the $b\tau$ pairing. The variables used for the BDT (outlined in Section 16.3.1), on the other hand, use the visible mass of the τ-leptons.

16.1.1 Pairing of b-quarks and τ-leptons

The decay of leptoquarks to $b\tau$ pairs, as opposed to $b\bar{b}$ and $\tau^+\tau^-$ pairs as in the di-Higgs decay, presents a combinatorial dilemma. Many of the kinematic variables utilised by the BDTs require knowledge of which b-jet and τ-lepton decayed from the same leptoquark.

A number of pairing strategies were explored in order to correctly reconstruct the leptoquarks:

- minimise $|\Delta m(\mathrm{LQ}_0, \mathrm{LQ}_1)|$, i.e. pair the b-jets and τ-leptons such that the mass difference between the two reconstructed leptoquarks, LQ_0 and LQ_1, is minimised;
- minimise $|\Delta p_{\mathrm{T}}(\mathrm{LQ}_0, \mathrm{LQ}_1)|$, i.e. pair the b-jets and τ-leptons such that the difference in p_{T} of the two reconstructed leptoquarks is minimised;
- minimise $|\pi - \Delta R(\ell, \mathrm{jet})| + |\pi - \Delta R(\tau_{\mathrm{had}}, \mathrm{jet})|$, i.e. pair b-jets and τ-leptons such that the sum of the values of $|\Delta R|$ between the paired particles is maximised.

The efficiencies with which the final state particles were paired correctly to reconstruct the parent leptoquarks were calculated using MC simulation and are shown in Figure 16.1. The pairing efficiency is greatest when minimising $|\Delta p_{\mathrm{T}}|$ between the two reconstructed leptoquarks. The full analysis chain was performed for each pairing strategy, including training separate BDTs and producing limits, which are shown in Figure 16.2. For $m(\mathrm{LQ}_3^u) < 900$ GeV, the analysis sensitivity is greatest when minimising $|\Delta m|$ between the two reconstructed leptoquarks. Above this threshold, the limits are not influenced by the pairing strategy. Although the $|\Delta p_{\mathrm{T}}|$ method has the best efficiency, the pairing strategy also affects the background distributions, which affects the signal sensitivity. As such, the $|\Delta m|$ strategy is used.

16.1.2 Inclusion of one b-tag events in the signal region

Since the leptoquark search extends to higher signal masses than the di-Higgs search, the inclusion of one b-tag events in the signal region brings a significant improvement in sensitivity. Events with one b-tag consider the b-tagged jet and the highest-p_{T} non-tagged jet as the leptoquark decay products.

Figure 16.3 shows the expected limits on the up-type leptoquark cross-section, $\sigma(pp \to \mathrm{LQ}_3^u\overline{\mathrm{LQ}_3^u})$, as a function of the leptoquark mass. The expected limit produced using only two b-tag events in the signal region is shown as a black dotted line with $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands shown in green and yellow, respectively. For $m(\mathrm{LQ}_3^u) > 800$ GeV,

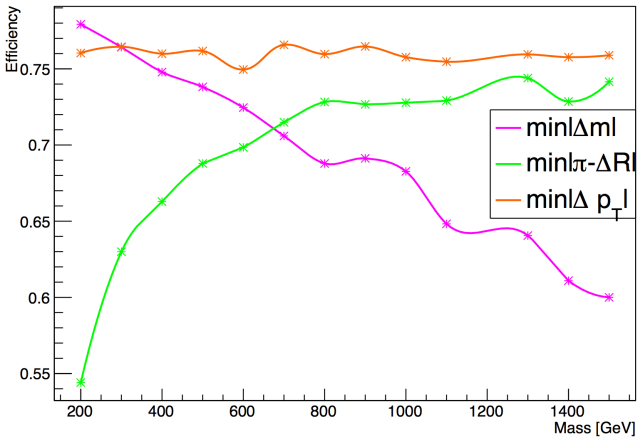


FIGURE 16.1: The efficiencies with which the final state particles were paired correctly. The efficiency achieved when minimising the mass difference between the two reconstructed leptoquarks is shown in pink; the efficiency achieved when minimising $|\Delta p_{\mathrm{T}}|$ between the paired particles is shown in red; and the efficiency achieved when minimising the sum of $|\pi - \Delta R|$ between the paired particles is shown in green.

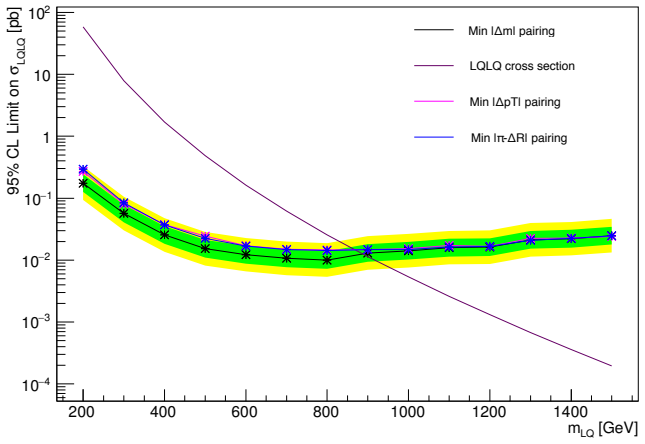


FIGURE 16.2: Expected 95% confidence level upper limits on the leptoquark pair-production cross-section using three different pairing methods (the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands are shown in green and yellow, respectively). The strongest limits are achieved when pairing the b-jets and τ-leptons such that the mass difference between the two reconstructed leptoquarks is minimised. These limits are calculated using the two b-tag signal region only.

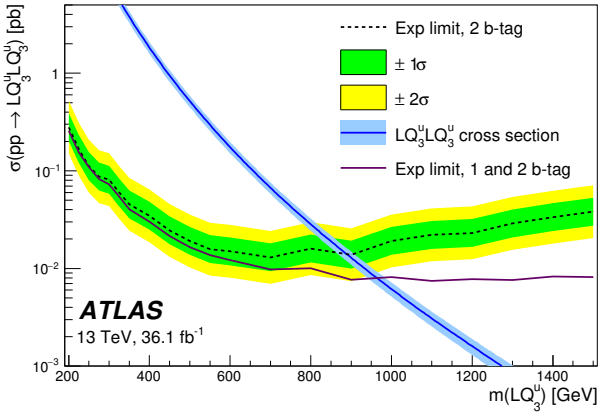


FIGURE 16.3: 95% confidence level upper limit on the up-type leptoquark cross-section, $\sigma(pp \to \mathrm{LQ}_3^u\mathrm{LQ}_3^u)$, as a function of the leptoquark mass, $m(\mathrm{LQ}_3^u)$. The expected limit produced using only two b-tag events in the signal region is shown as a black dotted line, with $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands shown in green and yellow, respectively. Limits produced using one and two b-tag events are shown as a solid purple line.

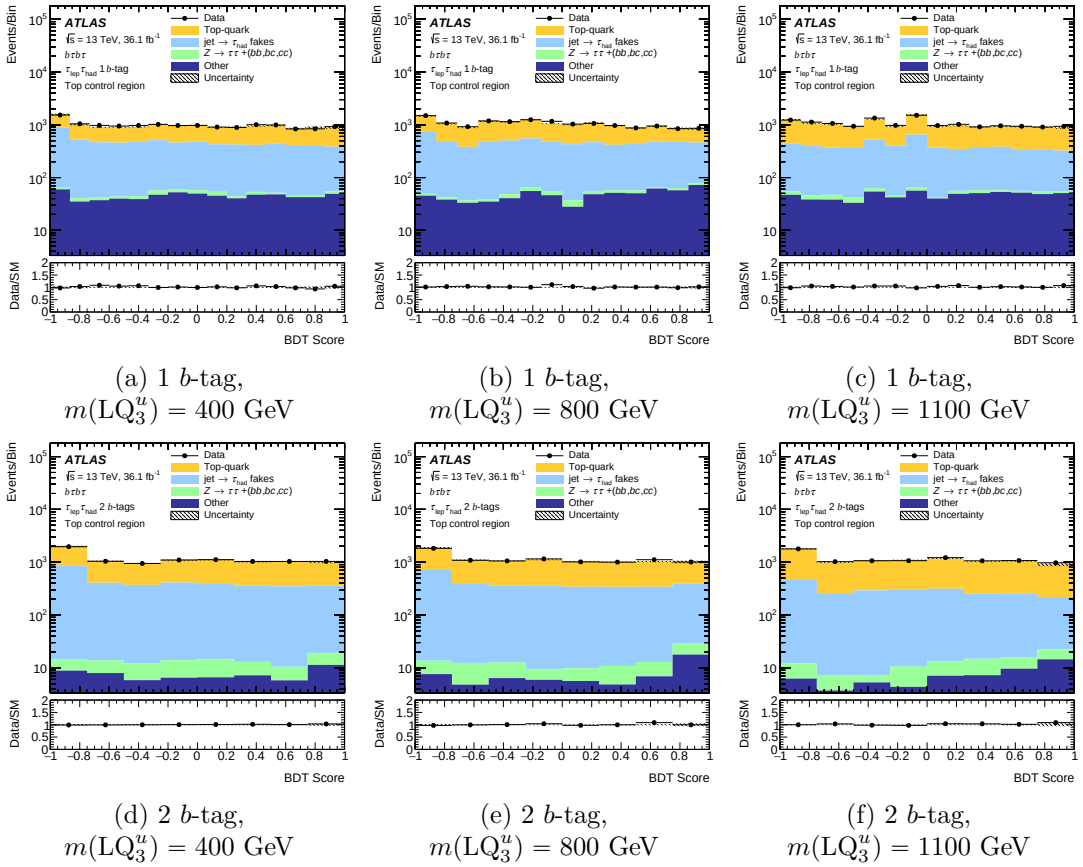
the analysis sensitivity decreases with increasing mass because the b-tagging efficiency decreases with increasing transverse momentum (at high-p_{T}, the separation between the tracks inside the b-jet is of the same order as the ID resolution). Therefore, the inclusion of the b-tag events in the signal region results in improved limits at high mass; these limits are shown as a solid purple line.

16.2 Background Estimation

The background estimation closely follows that in Chapter 12. Any deviations from the di-Higgs strategy are outlined in Appendix A. The main change for the leptoquark search is in the estimation of the fake-τ_{had} background processes. An additional selection criteria of $s_{\mathrm{T}} < 350$ GeV (s_{T} is defined in Section 16.3.1) is required for the control regions used in the data-driven fake factor method for $t\bar{t}$ and W + jets processes.

The background modelling is validated by comparing the estimated background with data in specially constructed validation regions. The modelling is studied in the signal/background BDT distributions; the BDTs are described in Chapter 16.3. The validation regions are studied separately for events with one and two b-tagged jets and are defined as follows:

- The real-τ_{had} $t\bar{t}$ background is validated in the high-m_{T}^W region, which is constructed by applying the signal region selection outlined in Section 16.1 and further requiring $m_{\mathrm{T}}^W > 40$ GeV and $s_{\mathrm{T}} < 350$ GeV.



(a) 1 b-tag, $m(\mathrm{LQ}_3^u) = 400$ GeV

(b) 1 b-tag, $m(\mathrm{LQ}_3^u) = 800$ GeV

(c) 1 b-tag, $m(\mathrm{LQ}_3^u) = 1100$ GeV

(d) 2 b-tag, $m(\mathrm{LQ}_3^u) = 400$ GeV

(e) 2 b-tag, $m(\mathrm{LQ}_3^u) = 800$ GeV

(f) 2 b-tag, $m(\mathrm{LQ}_3^u) = 1100$ GeV

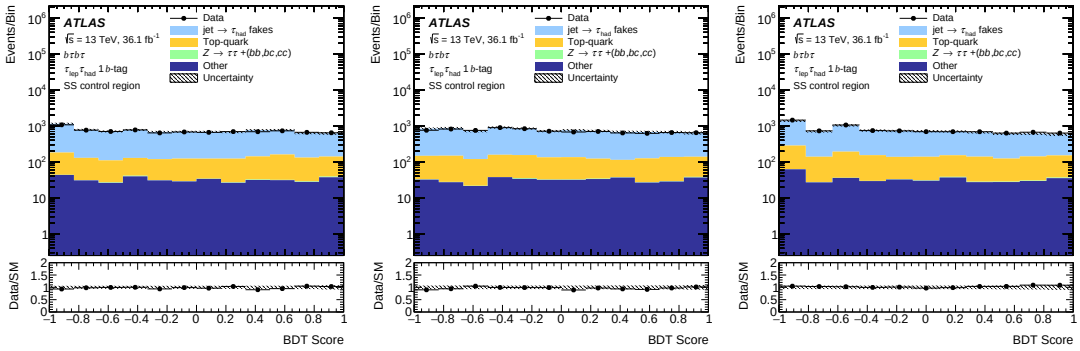
FIGURE 16.4: Post-fit BDT distributions in the high-m_{T}^W validation region for up-type leptoquarks with $m(\mathrm{LQ}_3^u) = 400,\ 800,\ 1100$ GeV (left, middle, right respectively) for events with one and two b-tagged jets (top and bottom respectively).

- The fake-τ_{had} backgrounds, estimated using the data-driven fake-factor method, are validated in the same-sign region, defined in Section 12.6.

Post-fit BDT distributions in the high-m_{T}^W validation region are shown in Figure 16.4 for up-type leptoquarks. BDT distributions in the same-sign validation region are shown in Figure 16.5. All regions show good agreement between data and MC simulation over the full BDT score distribution.

16.3 BDT Training

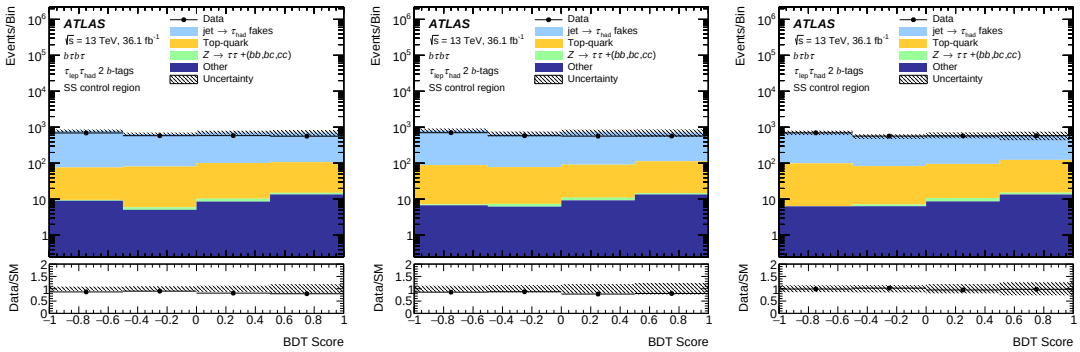
As in the di-Higgs search, BDTs are trained to separate signal and background, with their output distributions used as the final discriminant to test for the presence of a signal. The BDTs are trained separately for one and two b-tag events using the same variables. The BDTs trained using up-type leptoquark signal MC samples are applied to both up- and down-type signals.



(a) 1 b-tag, $m(\mathrm{LQ}_3^u) = 400$ GeV

(b) 1 b-tag, $m(\mathrm{LQ}_3^u) = 800$ GeV

(c) 1 b-tag, $m(\mathrm{LQ}_3^u) = 1100$ GeV



(d) 2 b-tag, $m(\mathrm{LQ}_3^u) = 400$ GeV

(e) 2 b-tag, $m(\mathrm{LQ}_3^u) = 800$ GeV

(f) 2 b-tag, $m(\mathrm{LQ}_3^u) = 1100$ GeV

FIGURE 16.5: Post-fit BDT distributions in the same-sign validation region for up-type leptoquarks with $m(\mathrm{LQ}_3^u) = 400,\ 800,\ 1100$ GeV (left, middle, right respectively) for events with one and two b-tagged jets (top and bottom respectively).

16.3.1 Variables

The variables used to train the BDTs used in the leptoquark search were selected in the same way as those for the di-Higgs search (outlined in Section 13.1). The variables are listed below in order of their importance in the two b-tag BDTs trained for leptoquarks with $m(\mathrm{LQ}_3^u) = 400$ GeV:

- s_T is the scalar sum of the missing transverse momentum ($E_\mathrm{T}^\mathrm{miss}$), the transverse momentum of the hadronic τ-lepton, the transverse momentum of the light lepton, and the transverse momentum of the two selected jets.
- $m_{\tau,\mathrm{jet}}$ is the invariant mass of the hadronic τ-lepton and the b-jet with which it is paired when minimising the mass difference between the two reconstructed leptoquarks (the pairing is described in Section 16.1.1).
- $m_{\ell,\mathrm{jet}}$ is the invariant mass of the light lepton and the b-jet with which it is paired when minimising the mass difference between the two reconstructed leptoquarks (the pairing is described in Section 16.1.1).

- $\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$ is the $\Delta\phi$ separation between the light lepton and the missing transverse momentum.
- $\Delta R(\ell, \tau_{\mathrm{had}})$ is the ΔR separation between the light lepton and the hadronic τ-lepton.
- $E_{\mathrm{T}}^{\mathrm{miss}}\phi$ centrality, defined in Section 13.1, quantifies the position in ϕ of the $E_{\mathrm{T}}^{\mathrm{miss}}$ with respect to the two τ-leptons.
- p_{T}^{τ} is the transverse momentum of the hadronic τ-lepton.

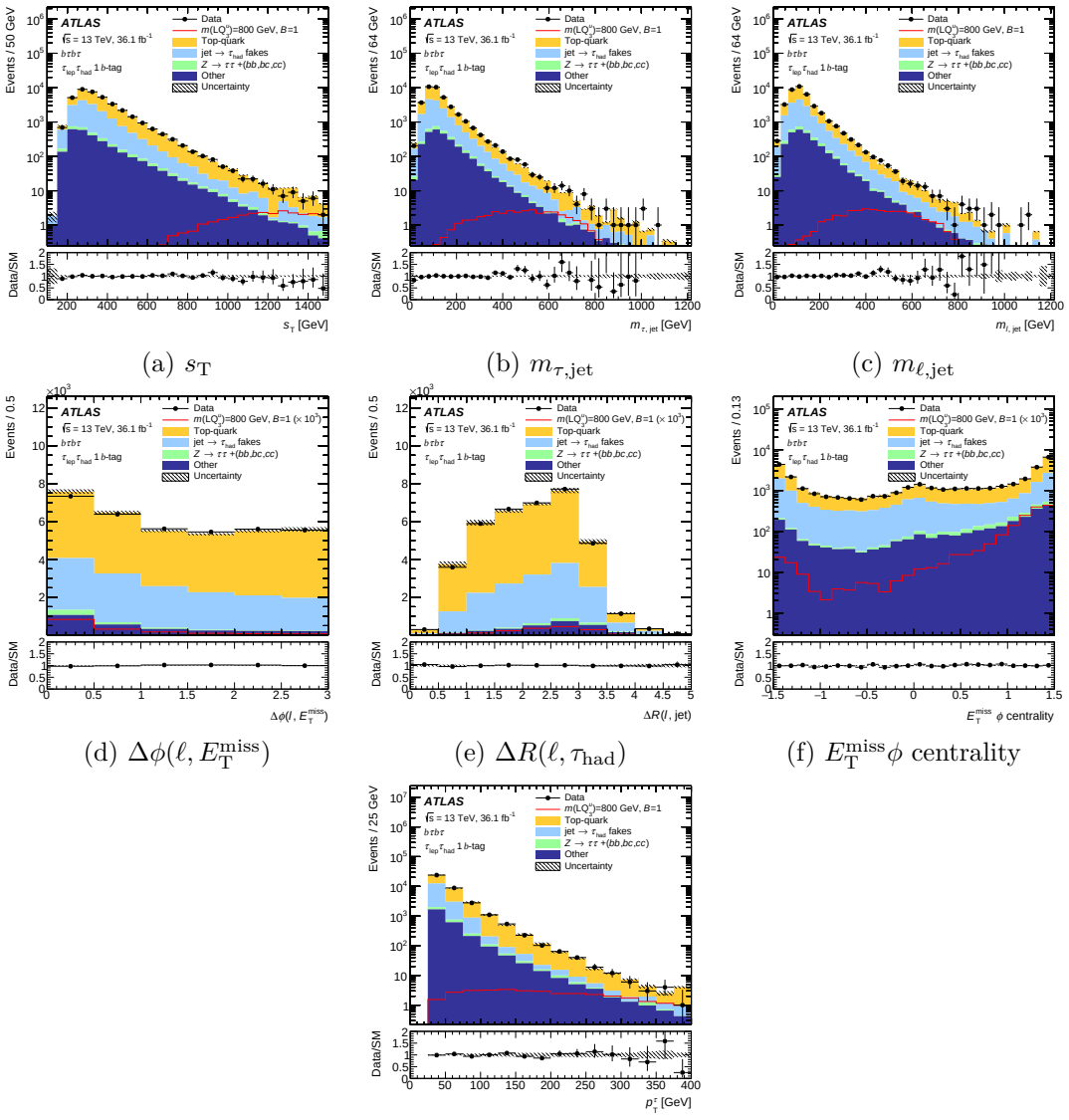
Variables involving a τ-lepton are calculated using the visible decay products. Figure 16.6 shows distributions of the variables used to train the one b-tag BDTs, after performing the full fitting procedure with a background-only hypothesis. The input variable distributions in the two b-tag region are shown in Figure 16.7. In both cases, the up-type leptoquark signal with $m(\mathrm{LQ}_3^u) = 800$ GeV is plotted as a red line to show how each variable discriminates between signal and background.

16.3.2 Training

Separate BDTs are trained for one and two b-tag events for each up-type signal sample. In both the one and two b-tag regions, BDTs are trained using $t\bar{t}$ processes as the only background. The inclusion of the $Z \to \tau\tau+$ heavy-flavour background as well as $t\bar{t}$ for one b-tag events was found to result in poorer sensitivity to the signal process.

As described in Section 13.2.1, each BDT used in the di-Higgs search is trained using three signal MC samples to ensure that the analysis is sensitive to a signal with a mass between those for which BDTs are trained. For this analysis, the BDTs are also trained using three merged signal samples. For example, for a leptoquark with $m(\mathrm{LQ}_3^u) = 500$ GeV, the BDT training is performed using a merged signal sample comprising the samples generated with $m(\mathrm{LQ}_3^u) = 450$ GeV, $m(\mathrm{LQ}_3^u) = 500$ GeV and $m(\mathrm{LQ}_3^u) = 550$ GeV.

Figure 16.8 shows the distribution of BDT scores from TMVA for one training and test set for an up-type leptoquark signal with (a) $m(\mathrm{LQ}_3^u) = 300$ GeV, (b) $m(\mathrm{LQ}_3^u) = 600$ GeV, and (c) $m(\mathrm{LQ}_3^u) = 1300$ GeV. The training and test distributions are consistent within errors and exhibit no overtraining. Similarly to the di-Higgs analysis, signals with higher mass are more easily distinguishable from the $t\bar{t}$ background.



(a) s_{T} (b) $m_{\tau,\mathrm{jet}}$ (c) $m_{\ell,\mathrm{jet}}$

(d) $\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$ (e) $\Delta R(\ell, \tau_{\mathrm{had}})$ (f) $E_{\mathrm{T}}^{\mathrm{miss}}\phi$ centrality

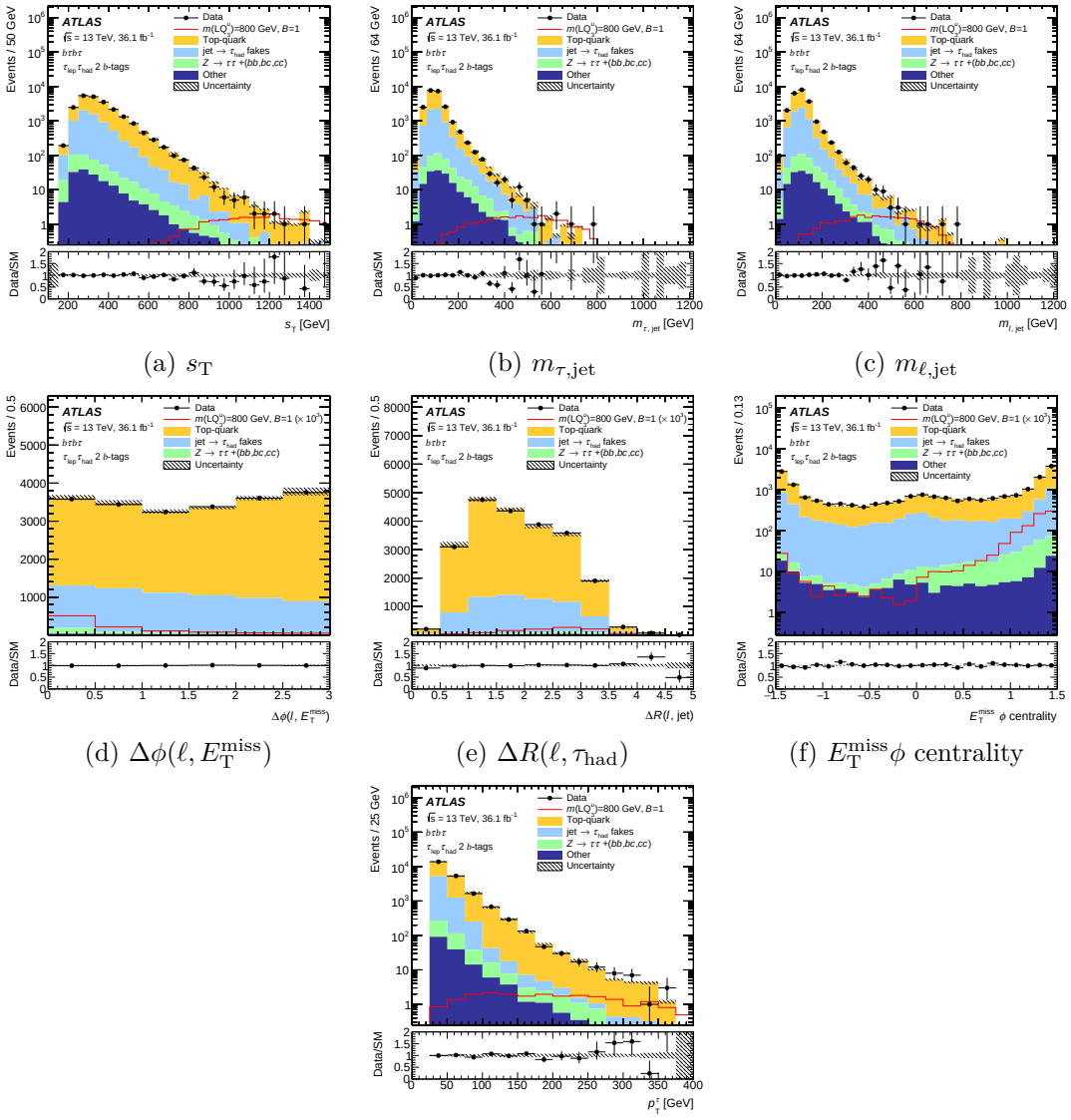
(g) p_{T}^{τ}

FIGURE 16.6: Post-fit distributions of the variables used to train the one b-tag BDTs, shown after performing the background-only fit (see Section 5.1). The up-type leptoquark signal with $m(\mathrm{LQ}_3^u) = 800$ GeV is shown as a red line.

16.4 Systematics

Systematics associated with the background modelling, where different from the di-Higgs analysis, are described in Appendix A. All other systematics are defined in Chapter 14, with no difference between the two analyses.

Figure 16.9 shows the fractional impact of the systematic uncertainties on the signal strength for an up-type leptoquark with $m(\mathrm{LQ}_3^u) = 500$ GeV, $\Delta\mu/\Delta\mu_{\mathrm{tot}}$. The uncertainties are listed in decreasing order of their impact on μ.

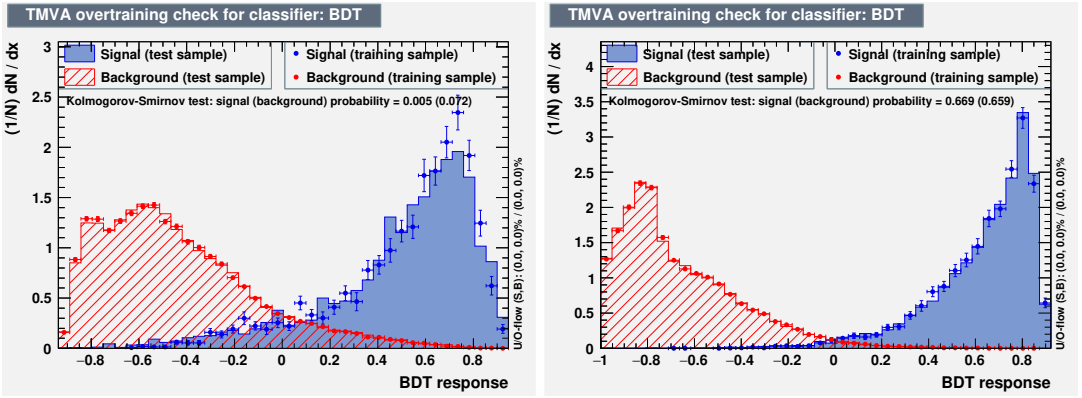


(a) s_{T} (b) $m_{\tau,\mathrm{jet}}$ (c) $m_{\ell,\mathrm{jet}}$

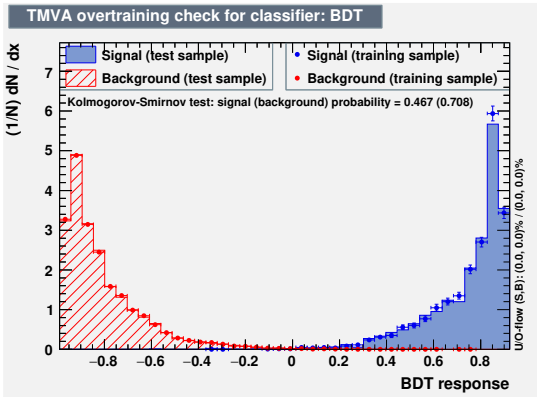
(d) $\Delta\phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$ (e) $\Delta R(\ell, \tau_{\mathrm{had}})$ (f) $E_{\mathrm{T}}^{\mathrm{miss}}\phi$ centrality

(g) p_{T}^{τ}

FIGURE 16.7: Post-fit distributions of the variables used to train the two b-tag BDTs, shown after performing the background-only fit (see Section 5.1). The up-type leptoquark signal with $m(\mathrm{LQ}_3^u) = 800$ GeV is shown as a red line.



(a) $m(\mathrm{LQ}_3^u) = 300$ GeV (b) $m(\mathrm{LQ}_3^u) = 600$ GeV



(c) $m(\mathrm{LQ}_3^u) = 1300$ GeV

FIGURE 16.8: The distributions of BDT scores (for one training and test set) for an up-type leptoquark signal with (a) $m(\mathrm{LQ}_3^u) = 300$ GeV, (b) $m(\mathrm{LQ}_3^u) = 600$ GeV, and (c) $m(\mathrm{LQ}_3^u) = 1300$ GeV. The lower-mass signals suffer from lower statistics than the higher mass; therefore, they exhibit a slight discrepancy between the training and test distributions and a smaller KS probability. Because the signal kinematics for lower-mass signals are more similar to the background, the BDTs are less able to discriminate between signal and background.

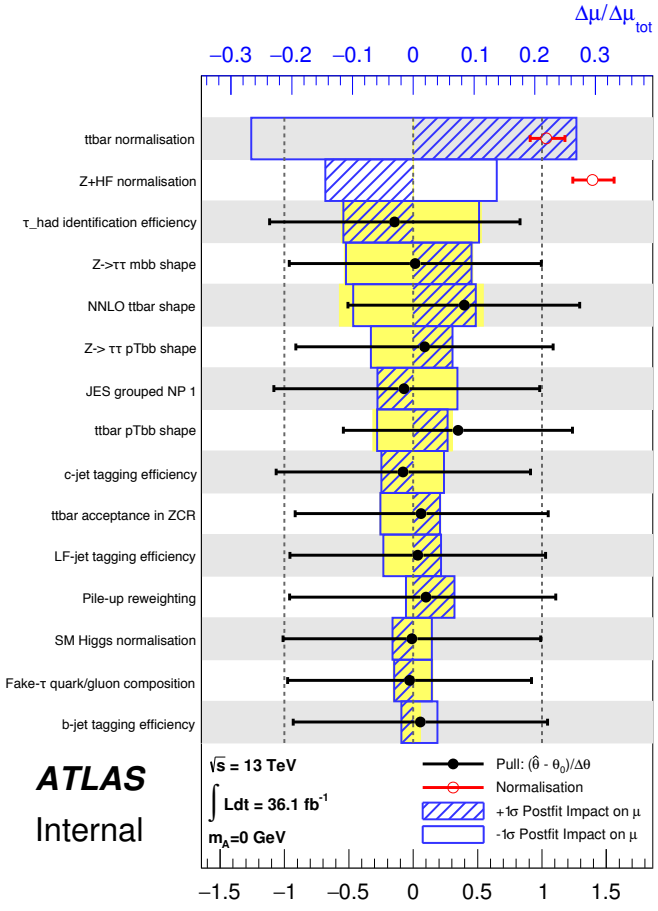


FIGURE 16.9: Nuisance parameter rankings for the up-type leptoquark with $m(\mathrm{LQ}_3^u) = 400$ GeV.

Chapter 17

Results

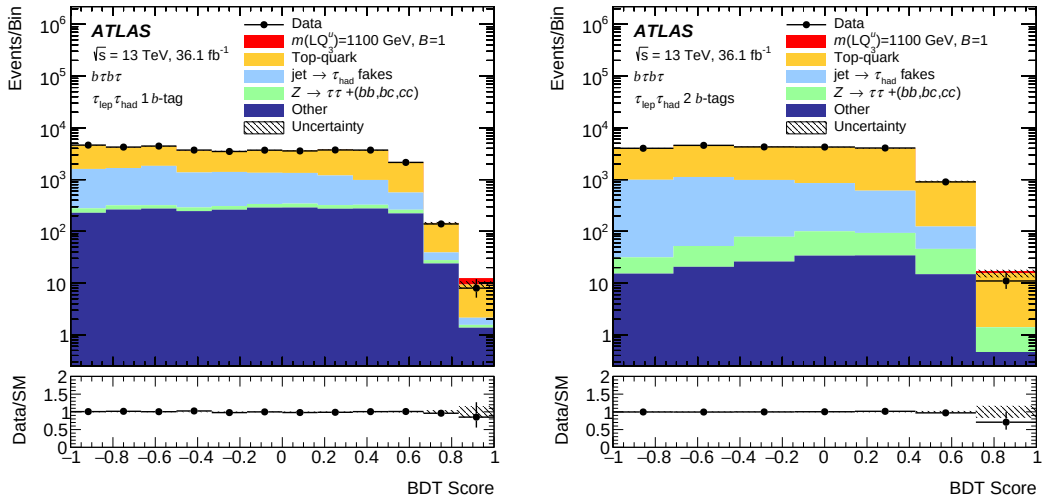
As explained in Section 5, the statistical interpretation of the results is performed using a likelihood fit. In the absence of a signal observation, results are presented as 95% confidence level upper limits.

Results are presented for the $\tau_{\text{lep}}\tau_{\text{had}}$ channel in Section 17.2 and the $\tau_{\text{had}}\tau_{\text{had}}$ channel in Section 17.3, followed by the combined $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ results in Section 17.4. Limits on the leptoquark mass, $m(\text{LQ}_3^u)$ ($m(\text{LQ}_3^d)$), as a function of the leptoquark branching ratio, $B(\text{LQ}_3^u \to b\tau)$ ($B(\text{LQ}_3^u \to t\tau)$), are presented in Section 17.5.

17.1 Fitting Procedure and Presentation of Results

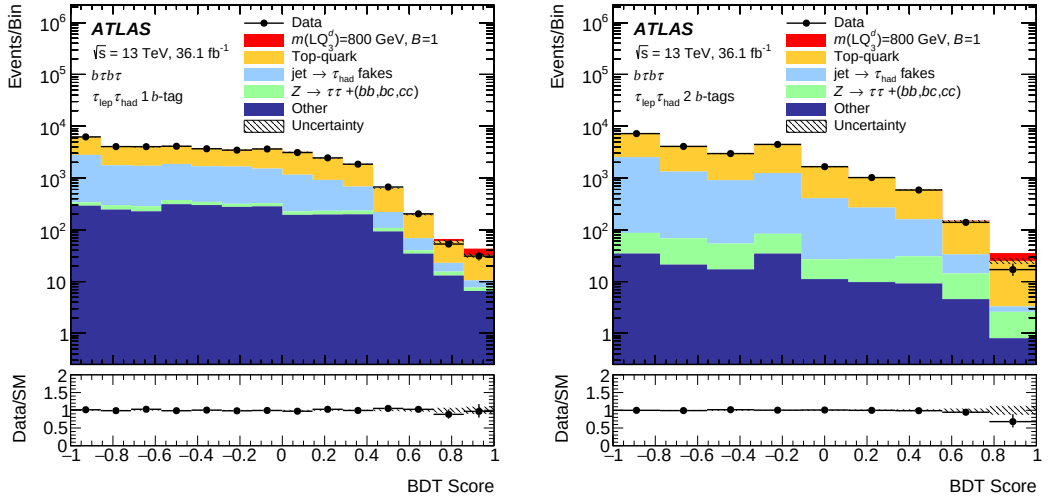
The BDT output score is used as the discriminating variable in the fit with the binning transformed according to Equation 15.1 with $x = 0.2$ and $y = 10$. The $\tau_{\text{lep}}\tau_{\text{had}}$ one and two b-tag signal regions and the $Z \to \mu\mu$ + heavy-flavour jets control region are all included in the fit. The extrapolation uncertainties applied in the di-Higgs analysis and described in Sections 12.3.2 and 15.3 are kept the same for the leptoquark search.

Using a profile likelihood test, defined in Chapter 5, the observed number of events is found to be compatible with the fitted number of background events (i.e. no significant excess over the SM background is observed) for all signal hypotheses. Therefore, the data are used to set upper limits at the 95% confidence level, following the CL_s prescription [84].



(a) 1 b-tag, $m(\mathrm{LQ}_3^u) = 1100$ GeV

(b) 2 b-tag, $m(\mathrm{LQ}_3^u) = 1100$ GeV



(c) 1 b-tag, $m(\mathrm{LQ}_3^d) = 800$ GeV

(d) 2 b-tag, $m(\mathrm{LQ}_3^d) = 800$ GeV

FIGURE 17.1: BDT output distributions in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel, shown after performing the combined likelihood fit. Plots are shown for the up- and down-type leptoquark signals with $m(\mathrm{LQ}_3^u) = 1100$ GeV (top) and $m(\mathrm{LQ}_3^d) = 800$ GeV (bottom) respectively. These are shown in the one and two b-tag regions on the left and right respectively.

17.2 $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ Results

Figure 17.1 shows the BDT output distributions in the one and two b-tag signal regions for up-type and down-type leptoquarks with $m(\mathrm{LQ}_3^u) = 1100$ GeV and $m(\mathrm{LQ}_3^d) = 800$ GeV respectively. Plots are shown after performing the combined $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ fit (which is detailed in Section 17.4) assuming a background-only hypothesis, for $m(\mathrm{LQ}_3^u) = 1100$ GeV.

The expected number of signal and background events and the observed number of data events after applying the selection criteria in Section 16.1, requiring exactly one or two

Sample	Number of events	
	1 b-tag	2 b-tags
$t\bar{t}$	18400 ± 1500	14460 ± 1000
Single-top	2490 ± 180	851 ± 73
Fake τ_{had}	13300 ± 1600	6200 ± 1100
$Z \to \tau\tau + (bb, bc, cc)$	540 ± 160	285 ± 85
Other	2790 ± 280	157 ± 28
Total background	37520 ± 200	22120 ± 160
Data	37527	22117
$m(\text{LQ}_3^u) = 400$ GeV	2210 ± 160	1970 ± 160
$m(\text{LQ}_3^d) = 400$ GeV	1420 ± 170	1096 ± 82
$m(\text{LQ}_3^u) = 800$ GeV	39.8 ± 3.1	25.2 ± 2.4
$m(\text{LQ}_3^d) = 800$ GeV	23.8 ± 2.4	16.7 ± 1.5
$m(\text{LQ}_3^u) = 1500$ GeV	0.25 ± 0.02	0.08 ± 0.01

TABLE 17.1: The expected number of signal and background events and the observed number of data events in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel with one and two b-tagged jets after performing the background-only fit. The background labelled 'Fake τ_{had}' includes all processes ($t\bar{t}$, QCD multi-jet and W + jets) in which a jet is misidentified as a τ_{had}. The total background yield is not equal to the sum of the individual backgrounds due to rounding. Individual uncertainties can be larger than the total uncertainty due to the large correlations.

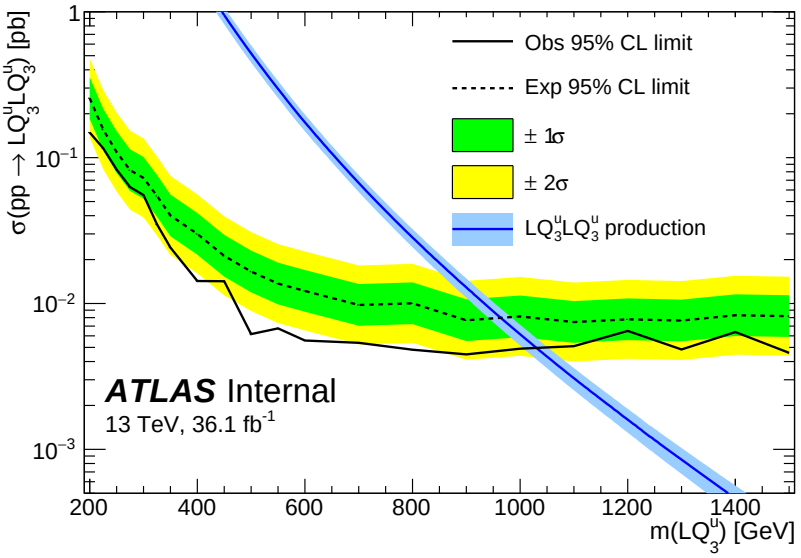
b-jets, and performing the $\tau_{\text{lep}}\tau_{\text{had}}$-only fit assuming a background-only hypothesis are given in Table 17.1.

Observed and expected 95% confidence level upper limits on the third-generation leptoquark pair-production, $\sigma(pp \to \text{LQ}_3^{u,d}\overline{\text{LQ}_3^{u,d}})$, are shown in Figure 17.2 for up- and down-type leptoquarks as a function of the leptoquark mass, $m(\text{LQ}_3^{u,d})$. The observed limits are between one and two standard deviations lower than the expected limits for the entire mass range for both up- and down-type leptoquarks. This is discussed in Section 17.2.1.

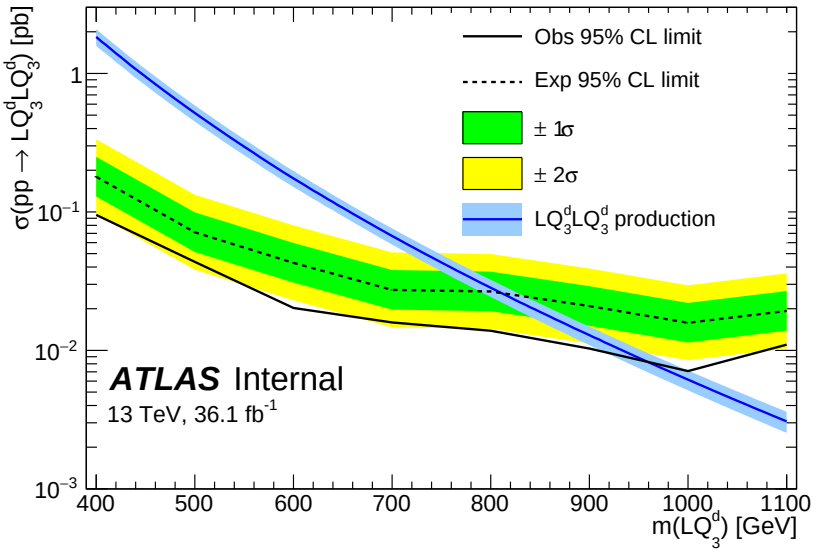
17.2.1 Deficit Check

The observed limits, shown in Figure 17.2, are between one and two standard deviations lower than the expected limits for both up- and down-type leptoquarks. This is due to a deficit in the number of observed data events compared to the predicted number of events, primarily in the final bin of the two b-tag BDT distributions. This can be seen in Figure 17.1 and in Appendix C.

Figure 17.3 shows the ratio of data to prediction, as well as the number of standard deviations to which the discrepancy in each bin of the BDT distributions corresponds, for up-type leptoquarks with $m(\text{LQ}_3^u) = 300$ GeV and $m(\text{LQ}_3^u) = 1000$ GeV in the one

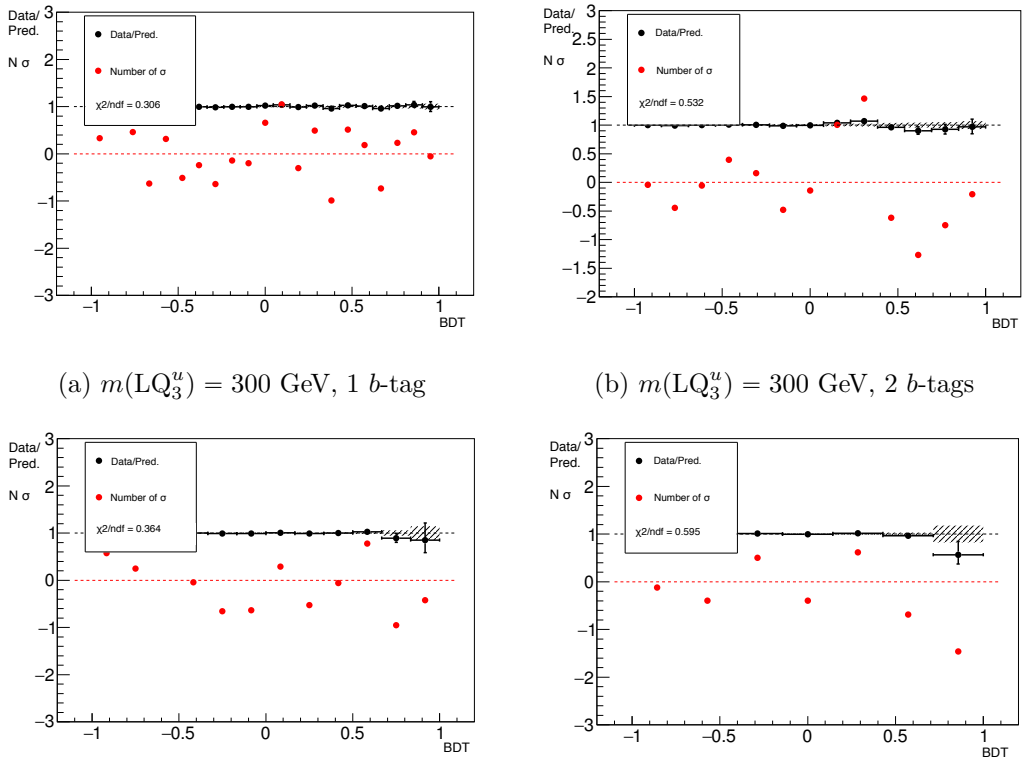


(a) Up-type leptoquarks



(b) Down-type leptoquarks

FIGURE 17.2: Observed (solid black) and expected (dotted black) limits at 95% confidence level on the leptoquark pair-production cross-section for (a) up-type leptoquarks, and (b) down-type leptoquarks, for the $\tau_{\text{lep}}\tau_{\text{had}}$ decay channel. The $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the expected limit are shown by green and yellow bands, respectively.



(a) $m(\mathrm{LQ}_3^u) = 300$ GeV, 1 b-tag

(b) $m(\mathrm{LQ}_3^u) = 300$ GeV, 2 b-tags

(c) $m(\mathrm{LQ}_3^u) = 1000$ GeV, 1 b-tag

(d) $m(\mathrm{LQ}_3^u) = 1000$ GeV, 2 b-tags

FIGURE 17.3: The ratio of the number of data events to the predicted number of events for each bin of the BDT distribution (black) and the number of standard deviations to which the discrepancy in each bin corresponds (red), shown for up-type leptoquarks with $m(\mathrm{LQ}_3^u) = 300$ GeV (top) and $m(\mathrm{LQ}_3^u) = 1000$ GeV (bottom) for events with one (left) and two (right) b-tags.

and two b-tag regions. A χ^2 fit is also performed, giving the compatibility of the data to prediction ratio with unity.

Table 17.2 shows the fraction of two b-tag events with the highest BDT scores for each mass that overlap with the highest-scoring events for the $m(\mathrm{LQ}_3^u) = 1500$ GeV BDT. The $m(\mathrm{LQ}_3^u) = 1500$ GeV BDT is chosen because the deficit at low mass is smaller. Similar fractions are seen for one b-tag events. The observed deficit is highly correlated across the range of masses; the background events that receive high BDT scores for one mass overlap with those that receive high BDT scores for the other masses. Therefore, a deficit observed at high BDT score for one mass point can be expected to be observed across the mass range.

$m(\mathrm{LQ}_3^u)$ (GeV)	Overlap (%)			
	$N = 10$	$N = 50$	$N = 100$	$N = 200$
300	60.0	76.0	82.0	99.3
600	70.0	78.0	80.0	99.3
900	80.0	76.0	83.0	99.3
1200	80.0	70.0	81.0	99.3

TABLE 17.2: Overlap of N highest-scoring $t\bar{t}$ events with $m(\mathrm{LQ}_3^u) = 1500$ GeV BDT (%).

17.3 $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ Results

Figure 17.4 shows the BDT distributions for the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel after performing the combined likelihood fit, for up-type and down-type leptoquarks in the one and two b-tag channels with $m(\mathrm{LQ}_3^u) = 1100$ GeV and $m(\mathrm{LQ}_3^d) = 800$ GeV. More distributions are shown in Appendix C.

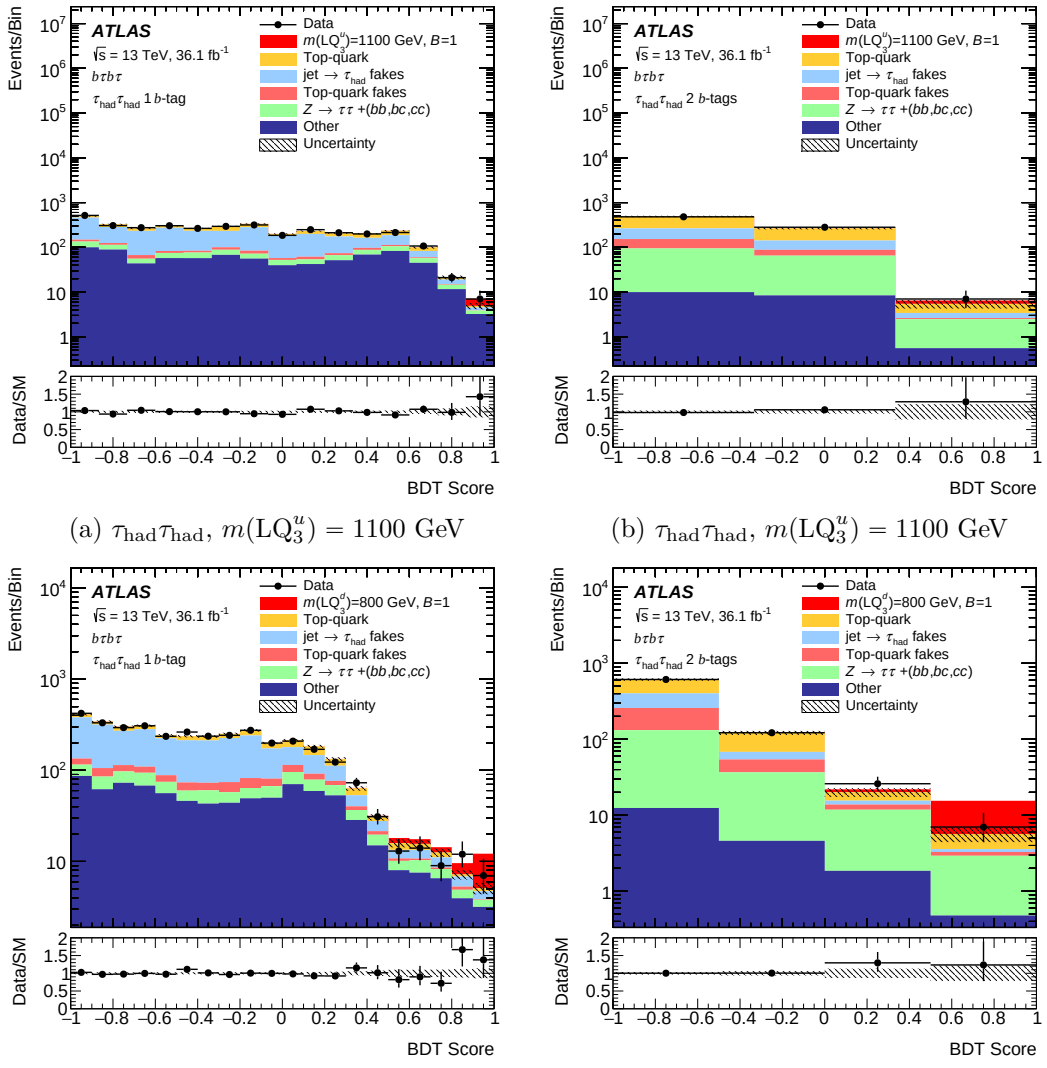
The $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$-only limits are shown in Figure 17.5. An excess of one to two standard deviations is observed for almost the entire mass range for both up- and down-type leptoquarks. Similar studies to those conducted for the deficit observed in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel concluded that the excess is correlated for all masses. As will be seen in Section 17.4, when the two channels are combined, the observed limit is generally consistent with the expected within one standard deviation.

17.4 Combined $b\tau^+\bar{b}\tau^-$ Results

Results are presented here for the combination of the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channels (the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ decay is selected using the single-τ and di-τ triggers). The combination is performed by including the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ one and two b-tag signal regions, again with the $Z \to \mu\mu$ + heavy-flavour jets control region, in the likelihood fit.

The expected number of signal and background events and the observed number of data events after performing the background-only fit are given in Table 17.3. They are given separately for the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channels with one and two b-tagged events. The $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ yields differ slightly from those shown in Table 17.1, whilst remaining consistent within errors, due to the inclusion of the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ regions in the likelihood fits.

No excess is seen after combining the channels and the results are presented as 95% confidence level upper limits on the third-generation leptoquark pair-production cross-section, $\sigma(pp \to \mathrm{LQ}_3^{u,d}\overline{\mathrm{LQ}_3^{u,d}})$. These are shown in Figure 17.6 for up- and down-type leptoquarks as a function of the leptoquark mass, $m(\mathrm{LQ}_3^{u,d})$. From the data, up-type



(a) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 1100$ GeV

(b) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 1100$ GeV

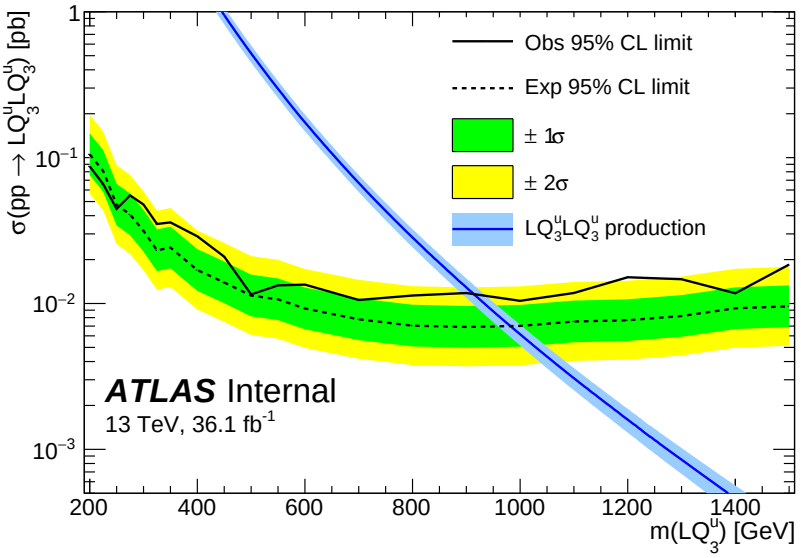
(c) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 800$ GeV

(d) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 800$ GeV

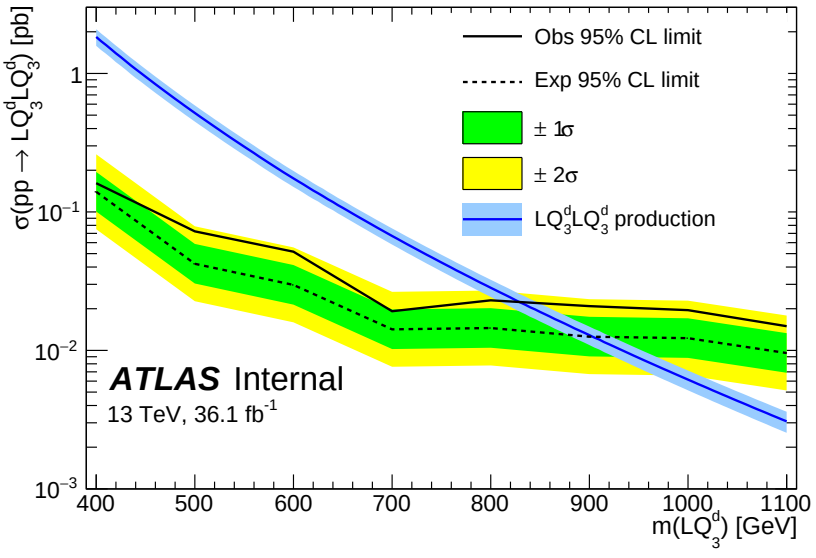
FIGURE 17.4: BDT output distributions in the $\tau_{\text{had}}\tau_{\text{had}}$ channel, shown after performing the combined likelihood fit. Plots are shown for the up- and down-type leptoquark signals with $m(\text{LQ}_3^u) = 1100$ GeV (top) and $m(\text{LQ}_3^d) = 800$ GeV (bottom) respectively. These are shown in the one and two b-tag regions on the left and right respectively.

(down-type) leptoquarks with $m(\text{LQ}_3^u) < 1030$ GeV ($m(\text{LQ}_3^d) < 930$ GeV) are excluded. The expected exclusion ranges are $m(\text{LQ}_3^u) < 1020$ GeV and $m(\text{LQ}_3^d) < 928$ GeV for up- and down-type leptoquarks, respectively.

For down-type leptoquarks, the $\tau_{\text{had}}\tau_{\text{had}}$ channel is more sensitive than the $\tau_{\text{lep}}\tau_{\text{had}}$ channel for the full range over which the search was performed. However, for up-type leptoquarks, the two channels become closer in sensitivity as $m(\text{LQ}_3^u)$ increases. Therefore, for up-type leptoquarks, the two channels contribute approximately equally to give the observed exclusion range of $m(\text{LQ}_3^u) < 1030$ GeV.

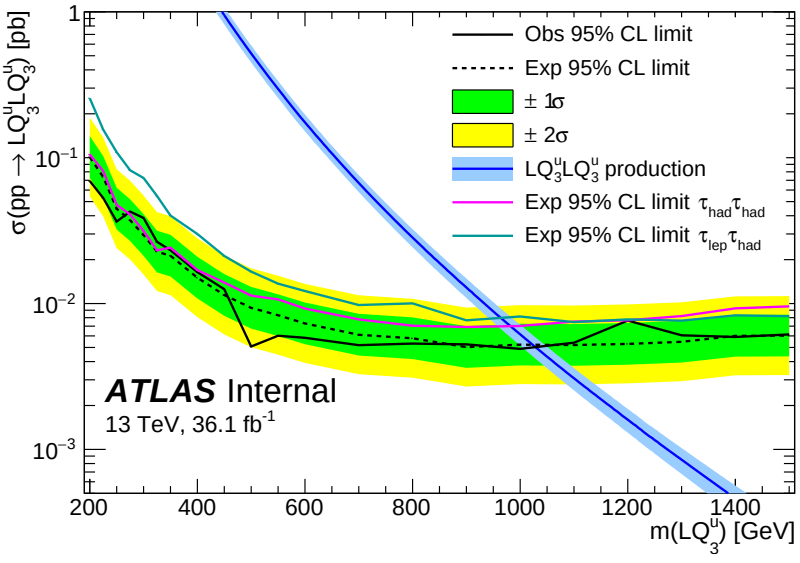


(a) Up-type leptoquarks

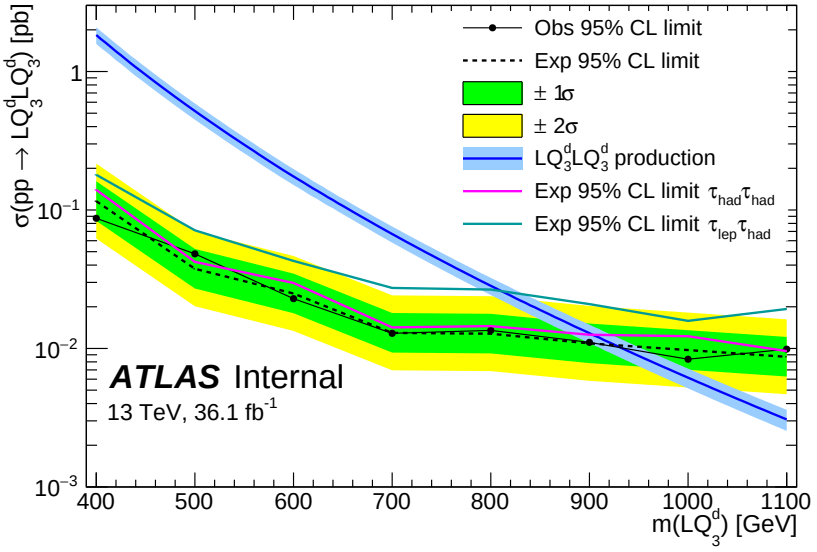


(b) Down-type leptoquarks

FIGURE 17.5: Observed (solid black) and expected (dotted black) limits at 95% confidence level on the leptoquark pair-production cross-section in the $\tau_{\text{had}}\tau_{\text{had}}$ channel for (a) up-type leptoquarks, and (b) down-type leptoquarks. The $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the expected limit are shown by green and yellow bands, respectively.



(a) Up-type leptoquarks



(b) Down-type leptoquarks

FIGURE 17.6: Observed (solid black) and expected (dotted black) limits at 95% confidence level on the leptoquark pair-production cross-section for (a) up-type leptoquarks, and (b) down-type leptoquarks. The individual $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ limits are shown in cyan and pink, respectively. The $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the expected limit are shown by green and yellow bands, respectively.

Sample	Number of events			
	$\tau_{\text{lep}}\tau_{\text{had}}$		$\tau_{\text{had}}\tau_{\text{had}}$	
	1 b-tag	2 b-tags	1 b-tag	2 b-tags
$t\bar{t}$	17800 ± 1500	14460 ± 980	285 ± 83	238 ± 69
$Single-top$	2500 ± 180	863 ± 73	63 ± 8	27 ± 3
Fake τ_{had}	13900 ± 1700	6400 ± 1000	-	-
QCD fake τ_{had}	-	-	1860 ± 110	173 ± 34
$t\bar{t}$ fake τ_{had}	-	-	200 ± 110	142 ± 79
$Z \to \tau\tau + (bb, bc, cc)$	520 ± 160	285 ± 83	258 ± 64	156 ± 36
Other	2785 ± 270	158 ± 26	817 ± 95	21 ± 4
Total background	37510 ± 220	22120 ± 160	3482 ± 59	756 ± 27
Data	37527	22117	3469	768
$m(\text{LQ}_3^u) = 400$ GeV	2140 ± 140	1950 ± 160	1430 ± 190	1430 ± 200
$m(\text{LQ}_3^d) = 400$ GeV	1420 ± 170	1096 ± 82	850 ± 110	67288
$m(\text{LQ}_3^u) = 800$ GeV	39.1 ± 2.8	25.2 ± 2.3	25.6 ± 3.9	16.8 ± 2.7
$m(\text{LQ}_3^d) = 800$ GeV	23.0 ± 2.3	16.6 ± 1.4	17.8 ± 2.8	12.4 ± 2.2
$m(\text{LQ}_3^u) = 1500$ GeV	0.25 ± 0.02	0.08 ± 0.01	0.16 ± 0.03	0.05 ± 0.01

TABLE 17.3: The expected number of signal and background events and the observed number of data events after performing the background-only fit in the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels with one and two b-tagged events. Fake τ_{had} events in the $\tau_{\text{had}}\tau_{\text{had}}$ channel are estimated separately for $t\bar{t}$ and QCD multi-jet processes. The total background yield is not equal to the sum of the individual backgrounds due to rounding. Individual uncertainties can be larger than the total uncertainty due to the large correlations. Yields for the $\tau_{\text{lep}}\tau_{\text{had}}$ channel are slightly different to those in Table 17.1 (but consistent within errors) due to the inclusion of the $\tau_{\text{had}}\tau_{\text{had}}$ regions in the fit.

17.5 Limits on the leptoquark mass as a function of the leptoquark branching ratio

The leptoquark search is presented in [195], alongside four reinterpretations of ATLAS searches for supersymmetric particles:

- searches for top squarks in the $t\bar{t} + E_{\text{T}}^{\text{miss}}$ channel in final states with one or zero leptons are reinterpreted to search for $LQ_3^u\overline{LQ_3^u}$ with $B = 0$;
- a search for bottom squarks in the $b\bar{b} + E_{\text{T}}^{\text{miss}}$ channel is reinterpreted to search for $LQ_3^d\overline{LQ_3^d}$ with $B = 0$;
- a search for top squarks in the $\tau + E_{\text{T}}^{\text{miss}}$ channel is sensitive to $LQ_3^u\overline{LQ_3^u}$ production with medium and high branching ratios into charged leptons.

Limits on the leptoquark mass, $m(\text{LQ}_3^u)$ ($m(\text{LQ}_3^d)$), as a function of the leptoquark branching ratio, $B(\text{LQ}_3^u \to b\tau)$ ($B(\text{LQ}_3^u \to t\tau)$), were derived using signal MC samples with model parameter $\beta = 0.5$, reweighted to achieve a range of branching ratios (following the method outlined in Section 8). These are presented in Figure 17.7 for all

search channels described in [195]. The combined limits for the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels are shown in green.

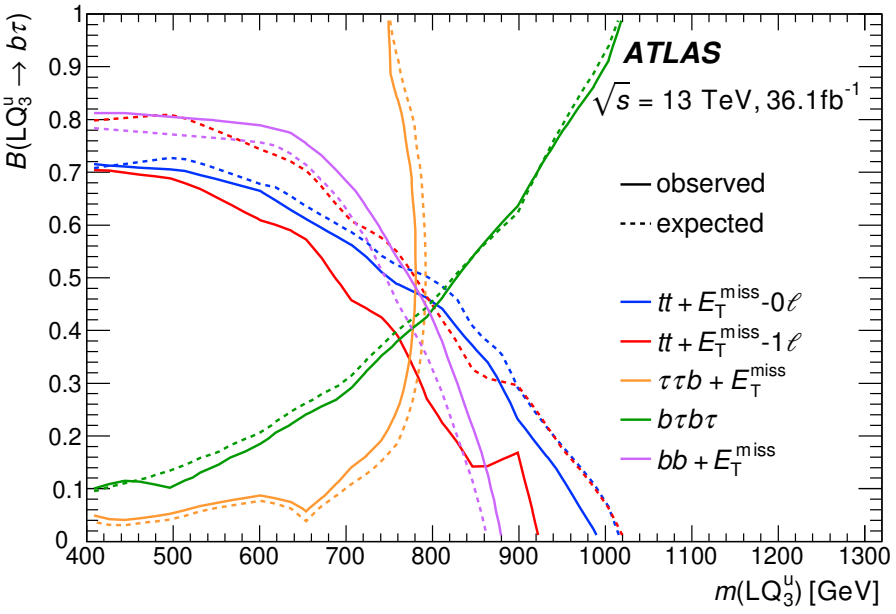
The region to the left of the contour lines is excluded at the 95% confidence level. For up-type leptoquarks, the $b\tau$ search provides the strongest limits in terms of mass exclusion for $B(\text{LQ}_3^u \to b\tau) > 0.5$ (the search is optimised for $B(\text{LQ}_3^u \to b\tau) = 1.0$). The strongest limit for $B(\text{LQ}_3^u \to b\tau) < 0.5$ is for the $t\bar{t} + E_{\text{T}}^{\text{miss}}$ channel, which is optimal for $B = 0$. For down-type leptoquarks, the $b\tau$ search provides the strongest limits for $B(\text{LQ}_3^u \to b\tau) > 0.7$. For low $B(\text{LQ}_3^u \to b\tau)$, the strongest limits are from the $b\bar{b} + E_{\text{T}}^{\text{miss}}$ channel. Up- and down-type leptoquarks with $m(\text{LQ}_3^{u,d}) < 800$ GeV are excluded independently of $B(\text{LQ}_3^u \to b\tau)$.

17.6 Conclusion

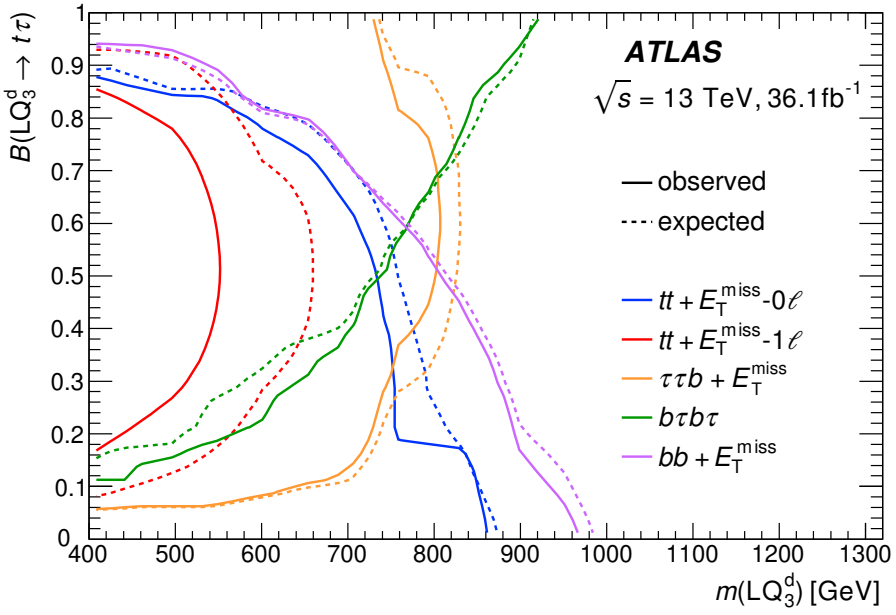
This chapter presents a search for pair-produced third-generation scalar leptoquarks, where each leptoquark decays to a b-quark and a τ-lepton. No excess is seen after combining the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels and the results are presented as 95% confidence level upper limits on the third-generation leptoquark pair-production cross-section.

Up-type (down-type) leptoquarks with $m(\text{LQ}_3^u) < 1030$ GeV ($m(\text{LQ}_3^d) < 930$ GeV) are excluded. The analysis is not optimised to search for down-type leptoquarks.

Limits on the leptoquark mass, $m(\text{LQ}_3^u)$ ($m(\text{LQ}_3^d)$), are also presented as a function of the leptoquark branching ratio, $B(\text{LQ}_3^u \to b\tau)$ ($B(\text{LQ}_3^u \to t\tau)$). Up- and down-type leptoquarks with $m(\text{LQ}_3^{u,d}) < 800$ GeV are excluded independently of $B(\text{LQ}_3^u \to b\tau)$.



(a) Up-type leptoquarks



(b) Down-type leptoquarks

FIGURE 17.7: Observed (solid) and expected (dotted) limits at 95% confidence level on the leptoquark mass, $m(\mathrm{LQ}_3)$, as a function of the leptoquark branching ratio, B, for (a) up-type and (b) down-type leptoquark pair-production. The region to the left of the contour lines is excluded at the 95% confidence level. The $b\tau$ channel is shown in green.

Part V

Conclusion and Appendices

Chapter 18

Conclusion

This thesis presents searches for resonant and non-resonant Higgs pair-production and leptoquark pair-production in final states with two b-quarks and two τ-leptons. Both searches used 36.1 fb^{-1} of proton-proton collision data collected by the ATLAS experiment in 2015 and 2016. The work of the author was concentrated on the $\tau_{\text{lep}}\tau_{\text{had}}$ channel and, for the leptoquark search, the combination with the $\tau_{\text{had}}\tau_{\text{had}}$ channel.

When searching for signal processes with small cross-sections, it is imperative that the background processes are well modelled and are assigned appropriate uncertainties. The background modelling and the estimation of the associated systematic uncertainty is documented in Section 12, alongside the validation of the modelling, for the di-Higgs analysis (the leptoquark analysis follows the same general strategy). Processes where a jet is misidentified as a hadronically-decaying τ-lepton are modelled using a data driven method; all other processes are modelled using MC simulations. For both analyses, BDTs are employed to boost the sensitivity to the signal process, producing the final discriminants which are used in the statistical fit and to set limits.

In the $hh \to bb\tau\tau$ channel, no excess is observed in either the resonant or the non-resonant search for Higgs pair-production. The observed (expected) 95% confidence level upper limit on the non-resonant Higgs pair-production cross-section is 30.9 fb (36.0 fb), i.e. 12.7 (14.8) times the SM prediction. This is, to date, the world's most stringent limit on non-resonant Higgs pair-production in a single decay channel; the ATLAS $hh \to b\bar{b}b\bar{b}$ observed (expected) limit is 12.9 (20.7) times the SM prediction. The ATLAS $hh \to bb\tau\tau$ channel alone outperforms the results achieved by the CMS experiment from a combination of all channels, which sets a 95% confidence level observed (expected) limit of 21.8 (12.4).

The search for resonant SM production is performed for two benchmark models: a narrow-width scalar Higgs in the hMSSM and a spin-2 Kaluza-Klein graviton in the bulk Randall-Sundrum model. For a heavy scalar Higgs boson in the hMSSM, the mass range $305 < m_H < 402$ GeV is excluded at the 95% confidence level for $\tan\beta = 2$. RS gravitons are excluded in the range $325 < m_G < 885$ GeV for $k/\bar{M}_{\mathrm{Pl}} = 1.0$; for $k/\bar{M}_{\mathrm{Pl}} = 2.0$, the entire mass range of the search ($260 < m_G < 1000$ GeV) can be excluded.

A search for pair-produced third-generation scalar leptoquarks is also presented, where each leptoquark decays to a b-quark and a τ-lepton. No excess is seen after combining the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ and $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channels and the results are presented as 95% confidence level upper limits on the third-generation leptoquark pair-production cross-section. Up-type (down-type) leptoquarks with $m(\mathrm{LQ}_3^u) < 1030$ GeV ($m(\mathrm{LQ}_3^d) < 930$ GeV) are excluded.

When the leptoquark limits are presented as a function of the leptoquark branching ratio alongside the other channels included in [195], the $b\tau$ search provides the strongest limits in terms of mass exclusion for $B(\mathrm{LQ}_3^u \to b\tau) > 0.5$ (the search is optimised for $B(\mathrm{LQ}_3^u \to b\tau) = 1.0$).

Appendix A

Background Estimation for Leptoquark Search

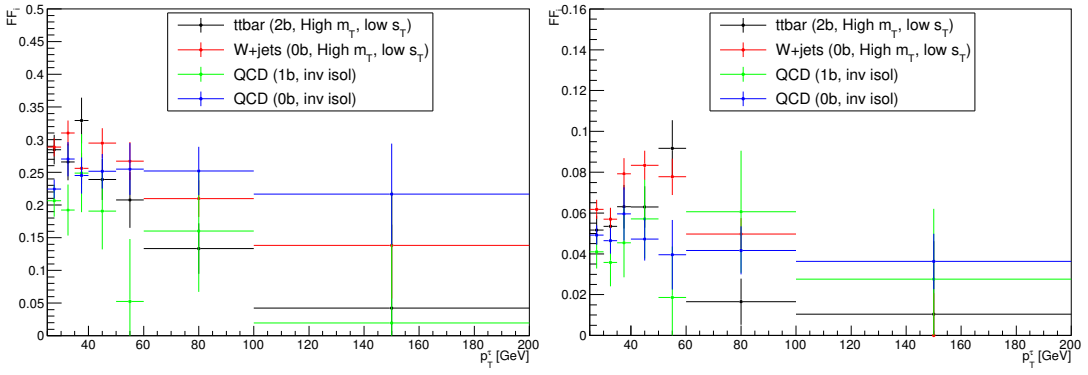
The background estimation for the leptoquark search in Part IV closely follows that in Chapter 12. Any deviations from the di-Higgs strategy are outlined in this appendix.

A.1 Backgrounds where a jet is mis-identified as a hadronic τ-lepton

The data-driven combined fake factor method used to estimate the fake-τ_{had} background closely follows that in Section 12.2. The control regions in which the fake factors are derived are redefined for $t\bar{t}$ and W + jets in order to eliminate leptoquark signal contamination, by adding a cut on s_{T} and increasing the requirement on m_{T}^{W} for W + jets. The updated control regions are defined as follows:

- **QCD multi-jet:** The isolation criteria for electrons and muons are inverted, i.e. electrons and muons are required to fail the *Loose* isolation working points. The fake factor is calculated separately for events with zero and one b-tagged jet(s), though these are similar. The fake factor calculated in the one b-tag region is used for events with two b-tagged jets, due to the lack of events in the two b-tag region.
- **$t\bar{t}$:** Events are required to have $m_{\text{T}}^{W} > 40$ GeV, $s_{\text{T}} < 350$ GeV and exactly two b-tagged jets.
- **W + jets:** Events are required to have $m_{\text{T}}^{W} > 60$ GeV, $s_{\text{T}} < 350$ GeV and exactly zero b-tagged jets.

The individual fake factors, which are parametrised in $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$, are shown in Figure A.1 for one- and three-prong jets. The $t\bar{t}$ and W + jets fake factors are generally consistent within their statistical errors and the QCD multi-jet fake factors are similar for events with zero and one b-tagged jets. As observed in the di-Higgs analysis, the $t\bar{t}$ and W + jets fake factors are larger than those for QCD multi-jet processes due to the higher fraction of quark-initiated jets in $t\bar{t}$ and W + jets processes than in QCD multi-jet processes. Furthermore, the fake factors are larger (for all three background sources) for one-prong events because a one-prong fake τ-lepton candidate is more likely to pass the τ-lepton identification than a three-prong candidate.



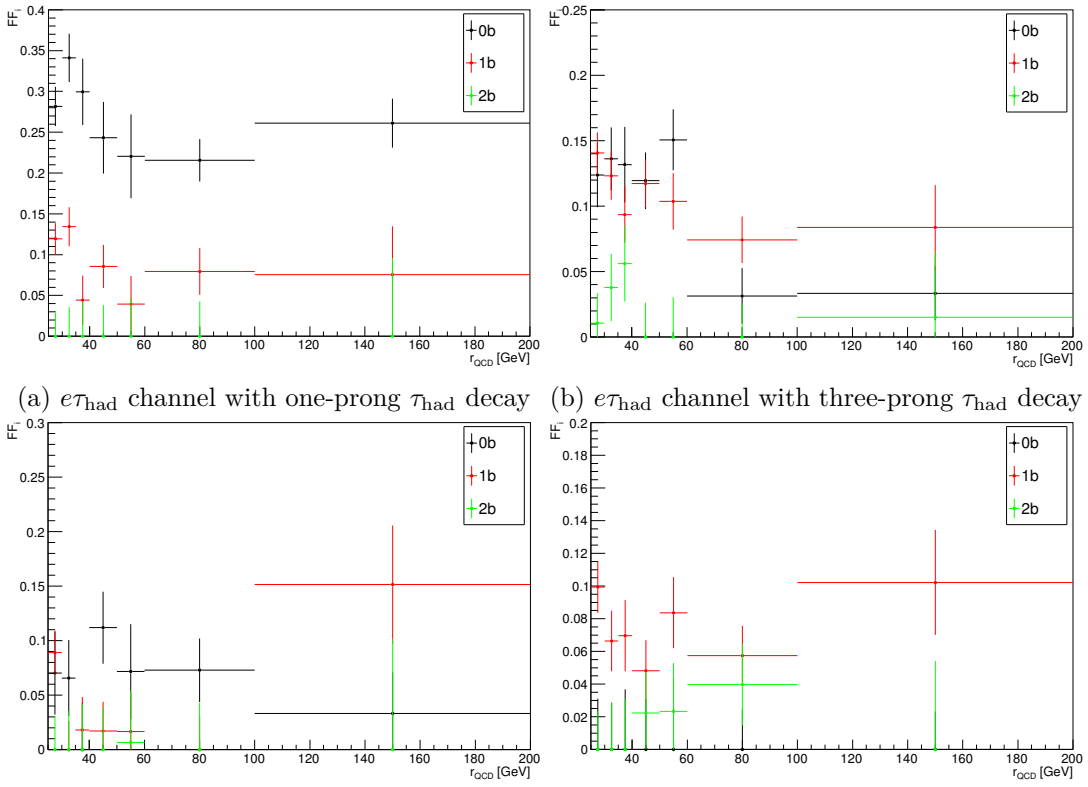
(a) One-prong (b) Three-prong

FIGURE A.1: The fake factors, FF_i, for the QCD multi-jet, $t\bar{t}$ and W + jets processes, as a function of $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ for (a) one-prong τ_{had} decays and (b) three-prong τ_{had} decays. The fake factor for QCD multi-jet is calculated separately for events with zero and one b-tagged jets, shown in blue and green respectively. The $t\bar{t}$ and W + jets fake factors are shown in black and red respectively.

The proportion of the total fake background for which QCD multi-jet processes are responsible, r_{QCD}, is parametrised in $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ and is provided separately according to the number of b-tagged jets, the number of prongs in the τ_{had} decay, and the type of light lepton to which the τ_{lep} decays. These are shown in Figure A.2. As expected, for one-prong τ_{had} decays, r_{QCD} decreases as the number of b-tagged jets in the event increases. This trend is not observed as clearly in the three-prong r_{QCD} distribution due to the lower numbers of three-prong decays. For all four decay channels, the two b-tag r_{QCD} values are consistent with zero. This is because the $t\bar{t}$ background is dominant.

A.1.1 Uncertainties on backgrounds where a jet fakes a τ_{had}

There are four uncertainties applied to backgrounds where a jet fakes a hadronic τ-lepton:



(a) $e\tau_{\text{had}}$ channel with one-prong τ_{had} decay (b) $e\tau_{\text{had}}$ channel with three-prong τ_{had} decay

(c) $\mu\tau_{\text{had}}$ channel with one-prong τ_{had} decay (d) $\mu\tau_{\text{had}}$ channel with three-prong τ_{had} decay

FIGURE A.2: The values of r_{QCD} as a function of $p_{\text{T}}^{\tau_{\text{had}}}$ for (a) the $e\tau_{\text{had}}$ channel with a one-prong τ_{had} decay, (b) the $e\tau_{\text{had}}$ channel with a three-prong τ_{had} decay, (c) the $\mu\tau_{\text{had}}$ channel with a one-prong τ_{had} decay and (d) the $\mu\tau_{\text{had}}$ channel with a three-prong τ_{had} decay. In each plot, r_{QCD} is shown for the regions with zero, one and two b-tagged jets; these are labelled '0b', '1b' and '2b' respectively.

- The uncertainty associated with the true-τ_{had} component of the template region is estimated by varying the true-τ_{had} $t\bar{t}$ MC sample which is subtracted from data within its uncertainties. Smaller backgrounds which are subtracted from data are subject to a conservative variation of $\pm 50\%$. This uncertainty is derived using the same procedure as the di-Higgs analysis, which is described in Section 12.2.1.2.

- An uncertainty is applied to account for the extrapolation of the fake factors from the high-m_{T} control regions to the signal region. This uncertainty is described in Section 12.2.1.3.

- The statistical uncertainty on the fake factor is taken into account by varying the fake factors up and down by their statistical uncertainty.

- The uncertainty applied to account for the difference in the quark and gluon flavour composition of jets in the signal region and the fake-τ_{had} enriched region is re-derived for the leptoquark search and is covered in Section A.1.1.1.

A.1.1.1 Quark and gluon composition variation

A systematic uncertainty is required to account for the difference in the quark and gluon flavour composition of jets misidentified as hadronically decaying τ-leptons in the signal region and the fake-τ_{had} enriched region used in the fake factor calculation. This uncertainty is estimated following the same procedure as in Section 12.2.1.1 but is performed separately for both one and two b-tag events as these are both included as signal regions.

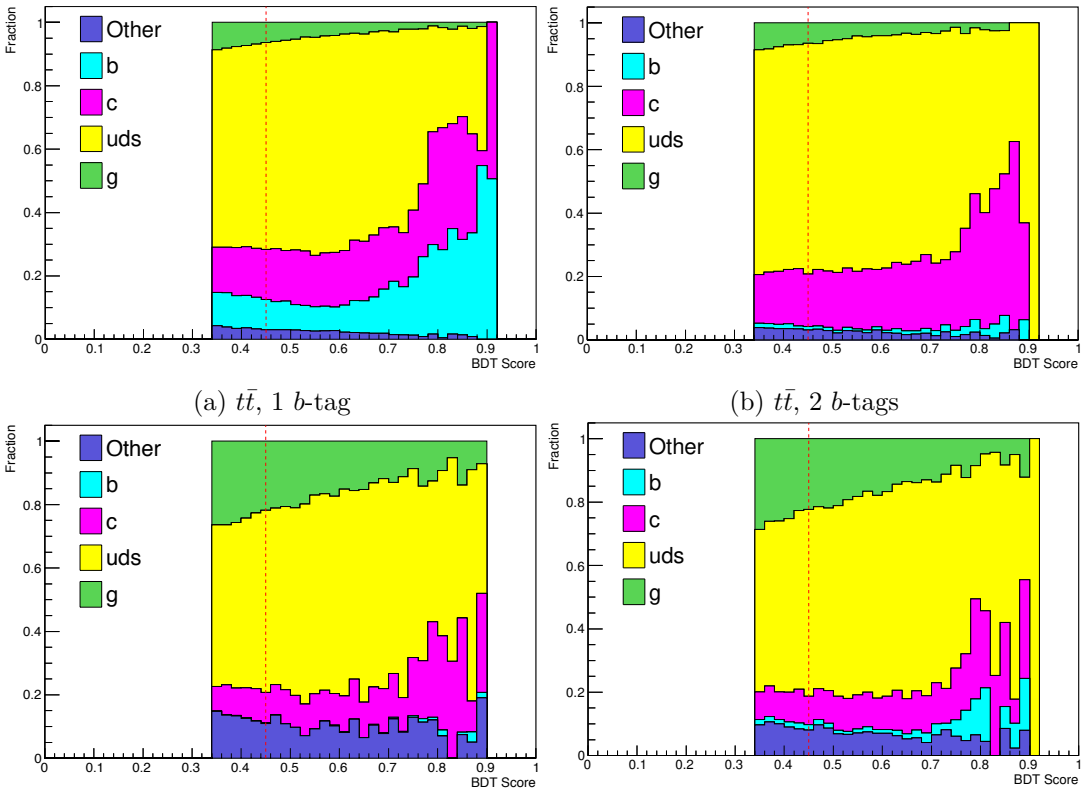
Figure A.3 shows the fraction of MC-simulated fake-τ_{had} events containing jets that originate from gluons, light-flavour quarks, c-quarks and b-quarks as a function of the τ-identification BDT score. The composition of these jets varies significantly with the τ-identification BDT score.

The estimation of the uncertainty is performed in the same-sign validation region, which requires the event to contain a hadronic τ-lepton and light lepton with same-sign electric charge. The fake-τ_{had} composition in the one and two b-tag signal regions is studied as a function of all input variables used in the BDT for signal/background discrimination (listed in Section 16.3). The greatest variation is observed as a function of $\Delta R(\ell, \tau)$; this is shown in Figure A.4.

The non-closure in the number of fake-τ_{had} events predicted by the fake factor method in the one and two b-tag same-sign validation regions is most evident as a function of $\Delta R(\ell, \tau)$. This is demonstrated in Figure A.5, which shows the ratio (calculated using Equation 12.7) of the number of fake-τ_{had} events in data to the number of fake-τ_{had} events estimated. The uncertainty is parametrised in $\Delta R(\ell, \tau)$ because it shows the greatest fake-τ_{had} composition and the greatest non-closure between data and estimation. A linear fit is performed and symmetrised to provide the up and down systematic uncertainties.

A.2 Top-quark pair-production decaying to a real τ_{had}

The $t\bar{t}$ background is modelled following the same procedure as 12.1. For an up-type leptoquark signal with $m(\text{LQ}_3^u) = 500$ GeV, the normalisation factor is 0.94 ± 0.11. A different normalisation factor is obtained for each fit that is performed (i.e. for each signal mass point), though they all agree within uncertainties.



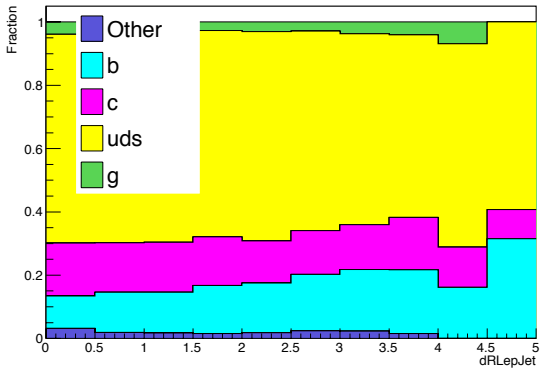
(a) $t\bar{t}$, 1 b-tag (b) $t\bar{t}$, 2 b-tags

(c) W+jets, 0 b-tags (d) W+jets 1 b-tag

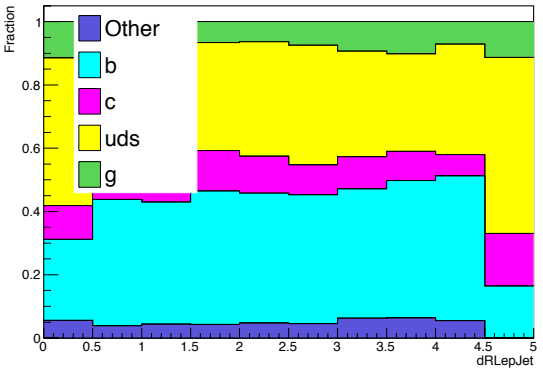
FIGURE A.3: The composition of jets mis-identified as hadronically-decaying τ-leptons as a function of the τ-identification BDT score for (a) $t\bar{t}$ events with one b-tagged jet, (b) $t\bar{t}$ events with two b-tagged jets, (c) W + jets events with no b-tagged jets and (d) W + jets events with one b-tagged jet. The plots show the fraction of events in which the fake τ_{had} originates from: a gluon ('g'), a light quark ('uds'), a c-quark ('c'), a b-quark ('b') or otherwise ('other', mainly pile-up jets). The red line marks the composition of jets at a τ-identification BDT score of 0.45, which is approximately (it is p_{T}-dependent) the minimum value required for a jet to be positively identified as a τ_{had}.

A.2.1 Uncertainties on top-quark processes with a real τ_{had}

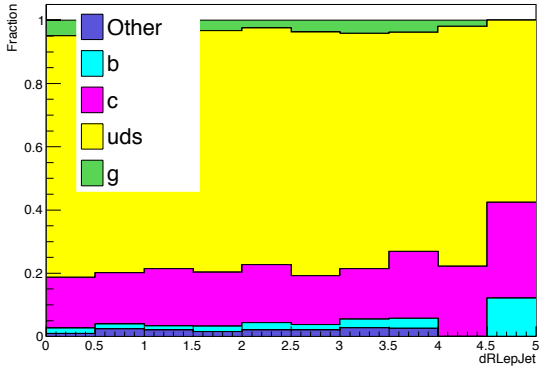
Generator-level uncertainties, listed in Section 12.1.1, are estimated by producing $t\bar{t}$ samples for each source of uncertainty and comparing these with the nominal sample. The variations with respect to the nominal sample are studied as functions of the variables used to train the signal/background BDT discriminant. The uncertainty is parametrised in s_{T}, as the greatest variation from the nominal is observed in the s_{T} distribution in both the one and two b-tag regions. This is shown in Figure A.6. The total uncertainty is taken as a symmetrised envelope of the individual effects in the two b-tag region, as the two regions exhibit similar variations in both size and shape. The largest variation comes from the HERWIG generator uncertainty.



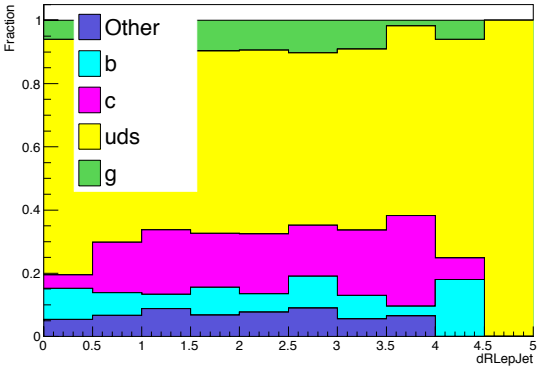
(a) Fake-τ_{had} composition in one b-tag signal region



(b) Fake-τ_{had} composition in one b-tag same-sign region

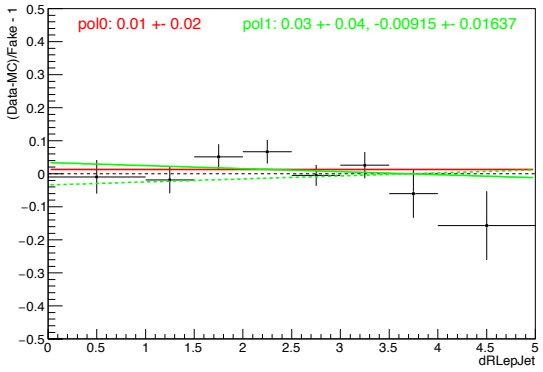


(c) Fake-τ_{had} composition in two b-tag signal region

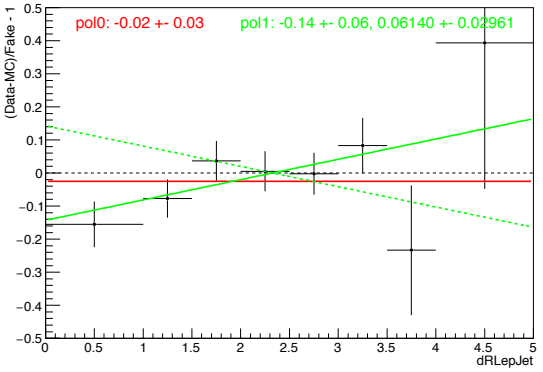


(d) Fake-τ_{had} composition in two b-tag same-sign region

FIGURE A.4: Composition of jets midentified as hadronic τ-leptons as a function of $\Delta R(\ell, \text{jet})$ in (a) the one b-tag opposite-sign signal region, (b) the one b-tag same-sign validation region, (c) the two b-tag opposite-sign signal region and (d) the two b-tag same-sign validation region.

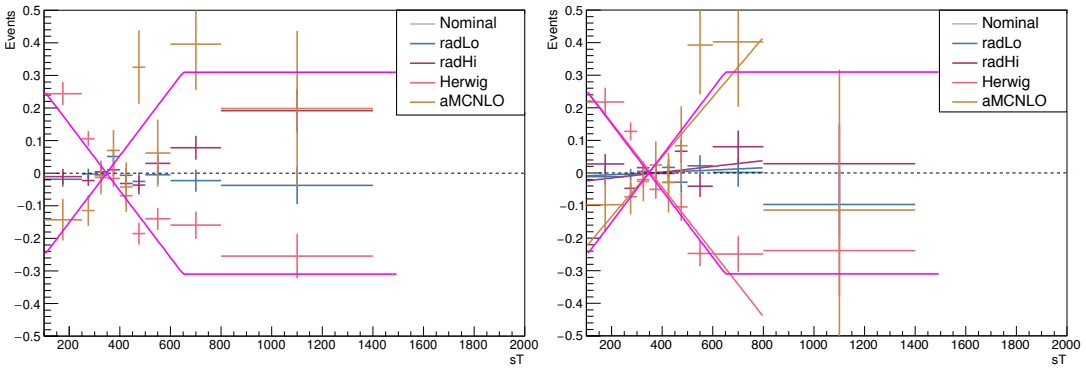


(a) $\Delta R(\ell, \tau)$, one b-tag



(b) $\Delta R(\ell, \tau)$, two b-tag

FIGURE A.5: The ratio of data to simulation in the same-sign region with (a) one and (b) two b-tags as a function of $\Delta R(\ell, \text{jet})$. The ratio is fitted and symmetrised to provide up and down systematic uncertainties (as illustrated by the green line).



(a) 1 b-tag (b) 2 b-tag

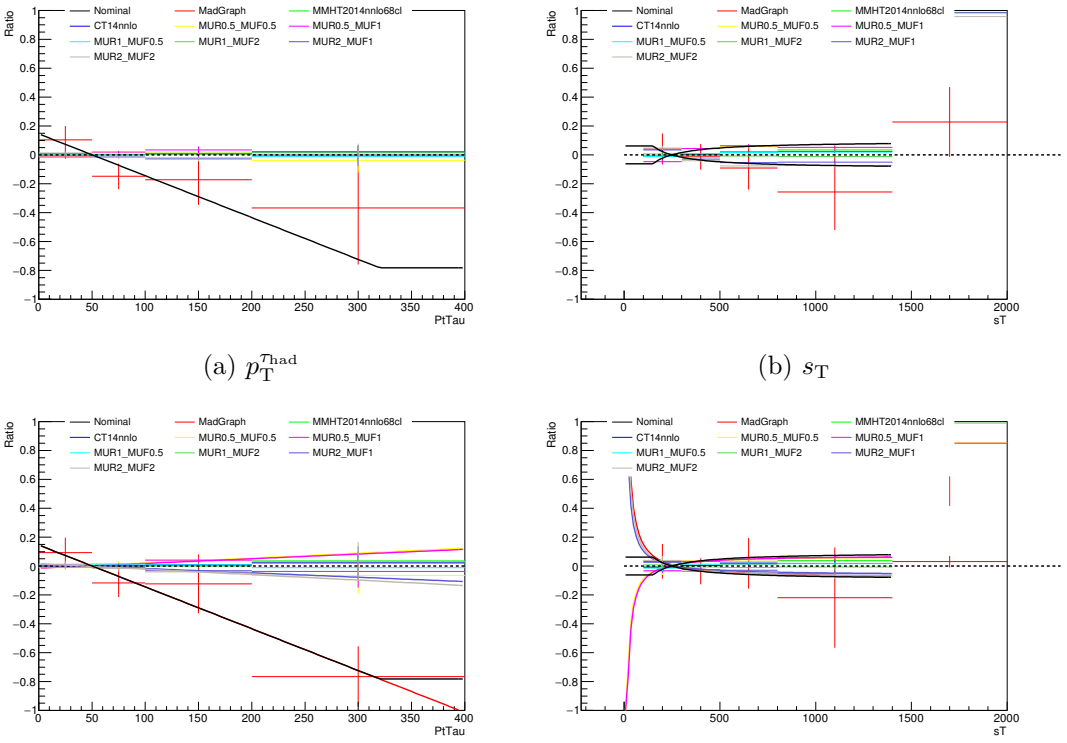
FIGURE A.6: Shape variations of the $t\bar{t}$ background as a function of s_{T} for (a) one b-tag events and (b) two b-tag events. The 'radHi' and 'radLo' variations refer to high and low radiation of additional jets, respectively; 'Herwig' refers to the variation of the fragmentation model from POWHEG+PYTHIA6 to POWHEG+HERWIG; and 'aMC-NLO' is the variation of the hard scatter simulation from that generated with POWHEG and showered using HERWIG++ to a sample generated with AMC@NLO. The total uncertainty, shown by the purple line in both plots, is a symmetrised envelope of the individual effects in the two b-tag region.

A.3 $Z \to \tau\tau+$ heavy-flavour jets

The cross-section for Z-boson production in association with a pair of heavy-flavour jets (combinations of b- and c-jets, i.e. $Z \to \tau\tau + (bb, bc, cc)$) is known to be poorly predicted by MC generators. Therefore, the MC simulations for these processes are normalised to data in a control region included in the likelihood fit, following the method outlined in Section 12.3. A different normalisation is obtained for every mass hypothesis; for an up-type leptoquark with $m(\mathrm{LQ}_3) = 500$ GeV, the normalisation is 1.42 ± 0.13.

A.3.1 Uncertainties on $Z \to \tau\tau+$ heavy-flavour jets processes

Generator-level uncertainties, listed in Section 12.3.1, are estimated by producing MC samples with varied normalisation and factorisation scales, PDF set and MC generator. These variations, relative to the nominal sample, are studied as a function of the variables used to train the signal/background BDT discriminant in both the one and two b-tag signal regions, as shown in Figure A.7. The uncertainty is parametrised by taking a symmetrised envelope of all variations as a function of $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ and s_{T} in the two b-tag signal region, which are the two variables in which the uncertainty is most evident. Because the variations in the one and two b-tag regions are similar, the same uncertainty is applied to both.



(a) $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ (b) s_{T}

(c) $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ (d) s_{T}

FIGURE A.7: Variation of $Z \rightarrow \tau\tau+$ heavy-flavour jets background as a function of $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ (left) and s_{T} (right) when varying the renormalisation and factorisation scales, the PDF set and the MC generator. These are shown in the one and two b-tag signal regions on the top and bottom rows, respectively. The total uncertainty applied to the $Z \rightarrow \tau\tau+$ heavy-flavour background is shown as a thick black line, taken as a symmetrised envelope of all uncertainties in the two b-tag signal region.

Appendix B

Sensitivity of Leptoquark Search to Intermediate Masses

As described in Section 13.2.1, each BDT used in the di-Higgs search is trained using three signal MC samples to ensure that the analysis is sensitive to a signal with a mass between those for which BDTs are trained. For the leptoquark analysis, the BDTs are also trained using three merged signal samples. For example, for a leptoquark with $m(\mathrm{LQ}_3^u) = 500$ GeV, the BDT training is performed using a merged signal sample comprising the samples generated with $m(\mathrm{LQ}_3^u) = 450$ GeV, $m(\mathrm{LQ}_3^u) = 500$ GeV and $m(\mathrm{LQ}_3^u) = 550$ GeV.

The sensitivity to intermediate mass points is tested by applying the BDT from the lower or higher mass point, using this as the final discriminant in the likelihood fit and propagating it to the limits. For example, for up-type leptoquarks, the 275 GeV BDT is applied to the 300 GeV signal sample and the 300 GeV BDT is applied to the 325 GeV signal. The results of this test are shown in Figure B.1. Very little difference is observed between the nominal limits and those using the neighbouring masses. Therefore, it is concluded that the analysis would be sensitive to a signal which fell between the mass points.

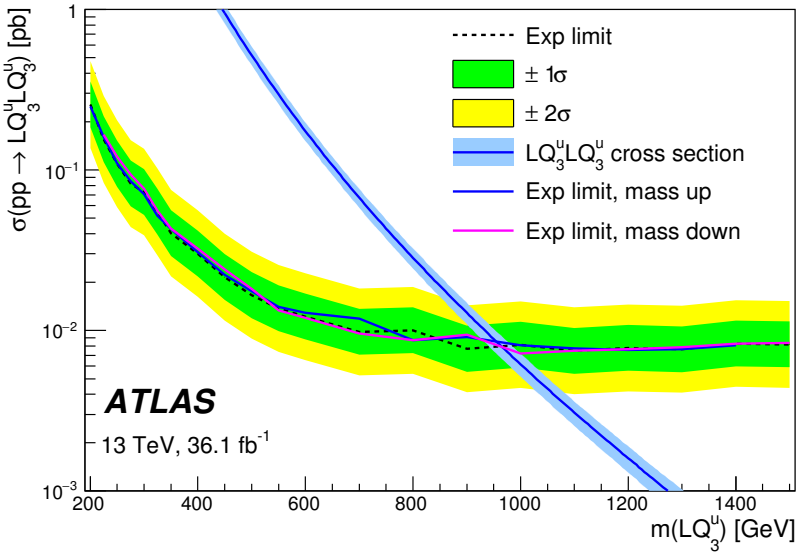
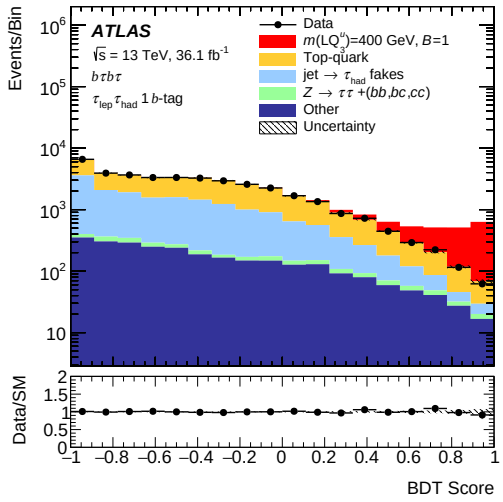


FIGURE B.1: Expected 95% CL upper limits on the up-type leptoquark pair-production cross section are shown as a black dotted line with uncertainty bands. The blue (magenta) line shows the limits produced when using the BDT trained for the signal mass 'up' ('down'), e.g. for a signal with $m(\mathrm{LQ}_3^u) = 300$ GeV the BDT trained for $m(\mathrm{LQ}_3^u) = 275$ GeV ($m(\mathrm{LQ}_3^u) = 325$ GeV is used.

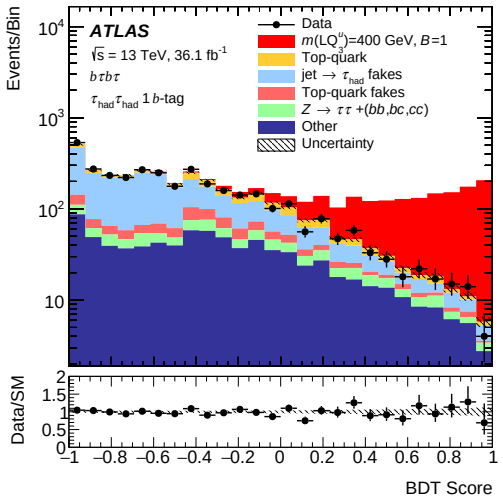
Appendix C

Post-fit BDT Distributions for Leptoquark Search

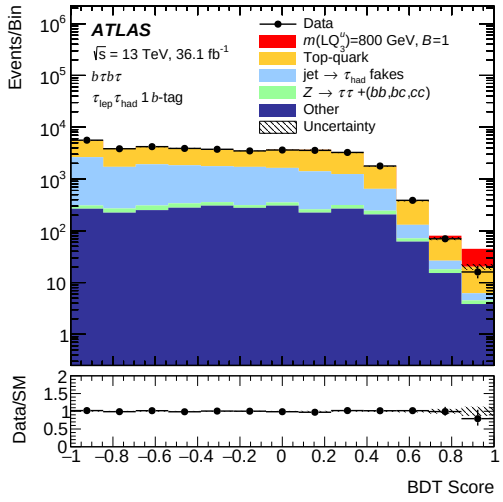
Figures C.1, C.2, C.3 and C.4 show postfit BDT distributions for the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ combined background-only fit, in the one and two b-tag signal regions for up- and down-type leptoquarks.



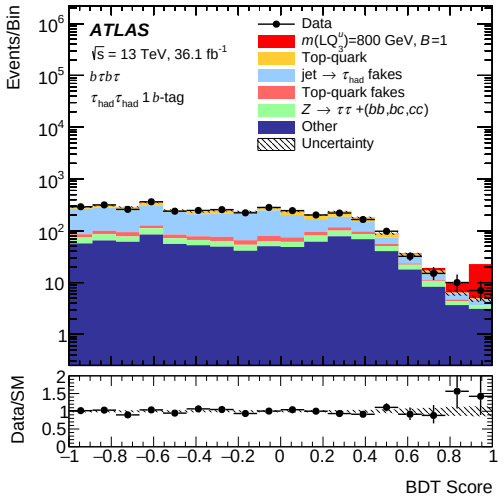
(a) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 400$ GeV



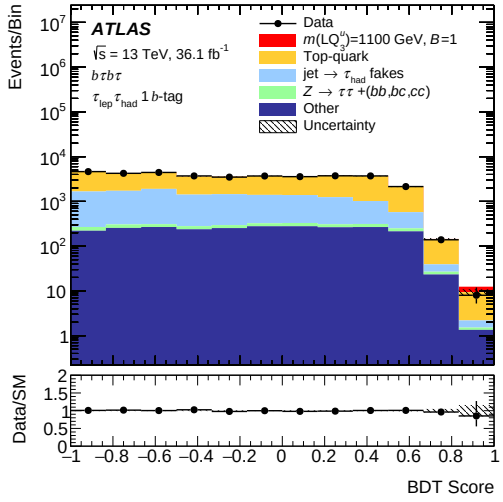
(b) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 400$ GeV



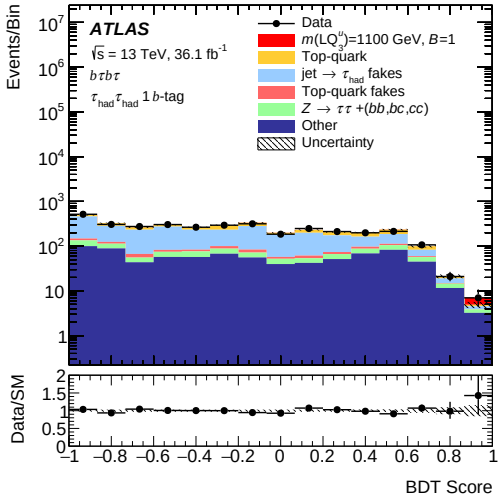
(c) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 800$ GeV



(d) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 800$ GeV

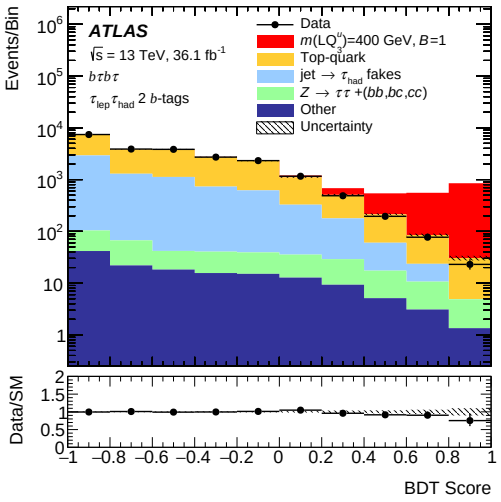


(e) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 1100$ GeV

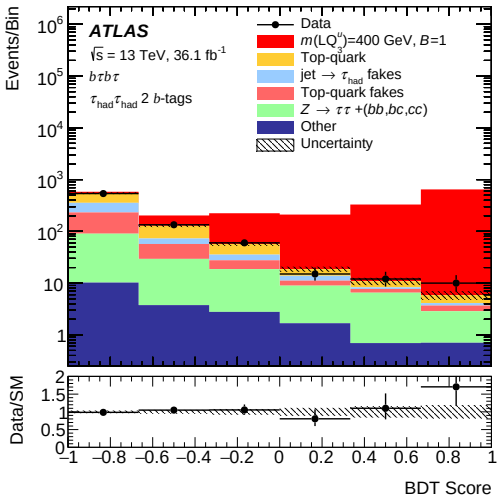


(f) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 1100$ GeV

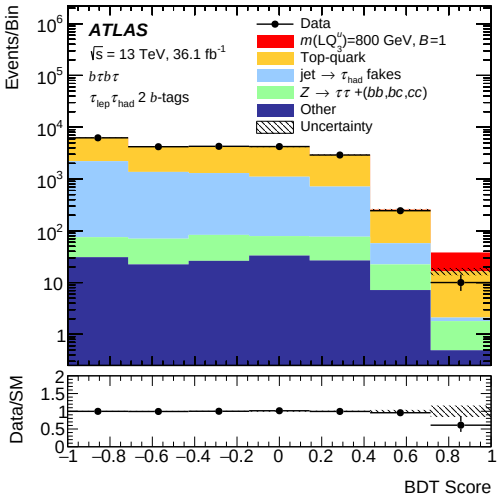
FIGURE C.1: BDT output distributions for the up-type leptoquark signal with $m(\text{LQ}_3^u) = 400$ GeV (top), $m(\text{LQ}_3^u) = 800$ GeV (middle) and $m(\text{LQ}_3^u) = 1100$ GeV (bottom), after performing the combined fit. Distributions are shown in the one b-tag signal region for the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels on the left and right respectively.



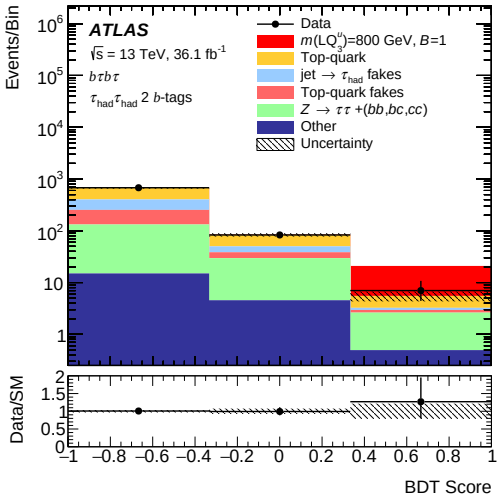
(a) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 400$ GeV



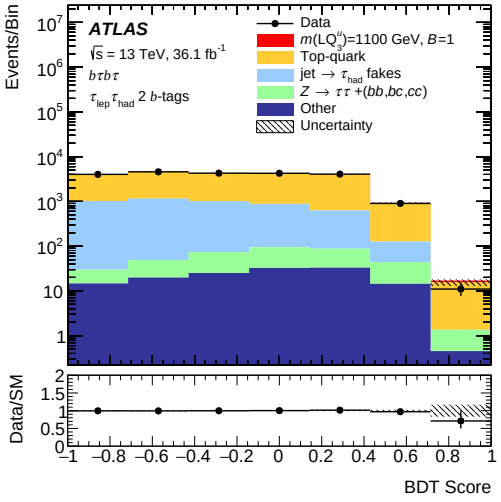
(b) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 400$ GeV



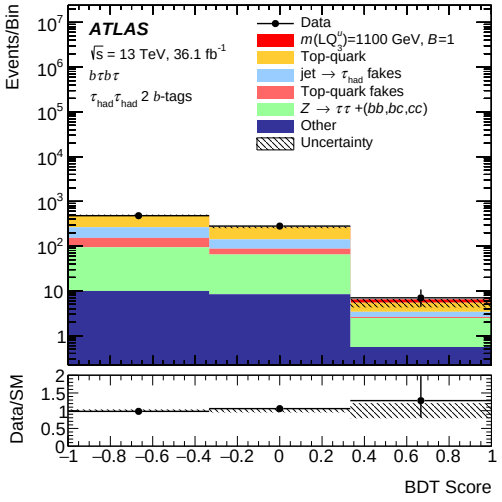
(c) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 800$ GeV



(d) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 800$ GeV

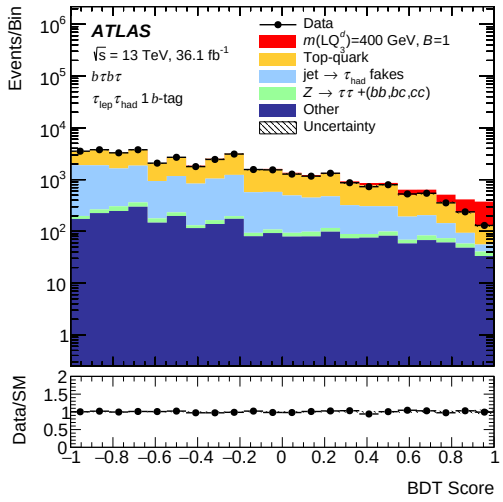


(e) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 1100$ GeV

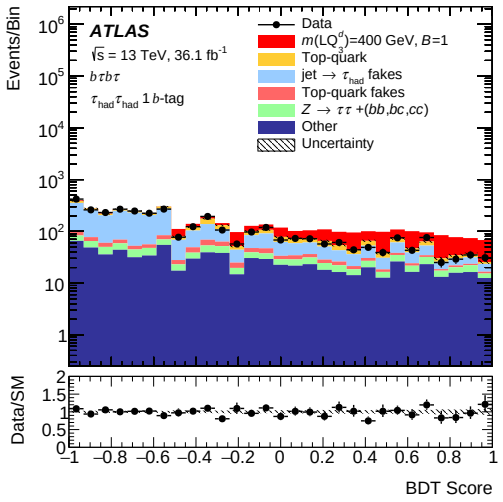


(f) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^u) = 1100$ GeV

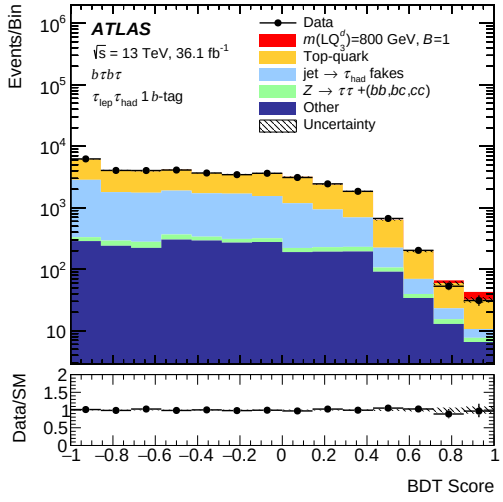
FIGURE C.2: BDT output distributions for the up-type leptoquark signal with $m(\text{LQ}_3^u) = 400$ GeV (top), $m(\text{LQ}_3^u) = 800$ GeV (middle) and $m(\text{LQ}_3^u) = 1100$ GeV (bottom), after performing the combined fit. Distributions are shown in the two b-tag signal region for the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels on the left and right respectively.



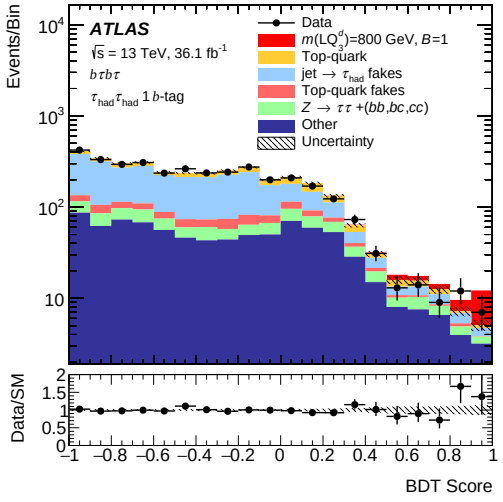
(a) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 400$ GeV



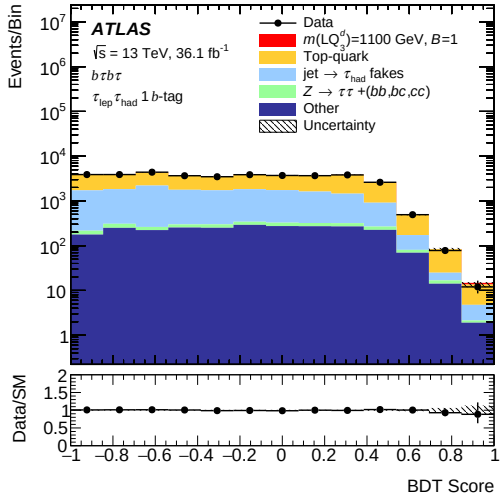
(b) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 400$ GeV



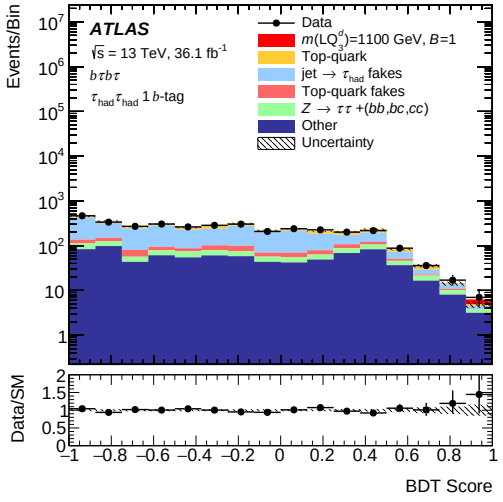
(c) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 800$ GeV



(d) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 800$ GeV

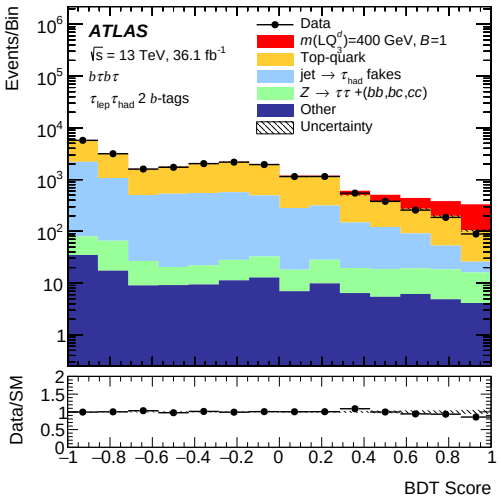


(e) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 1100$ GeV

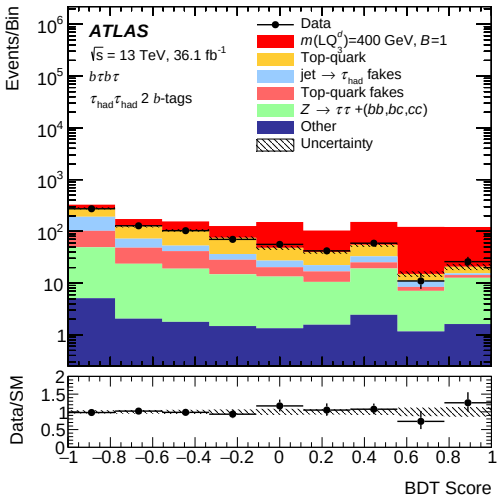


(f) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 1100$ GeV

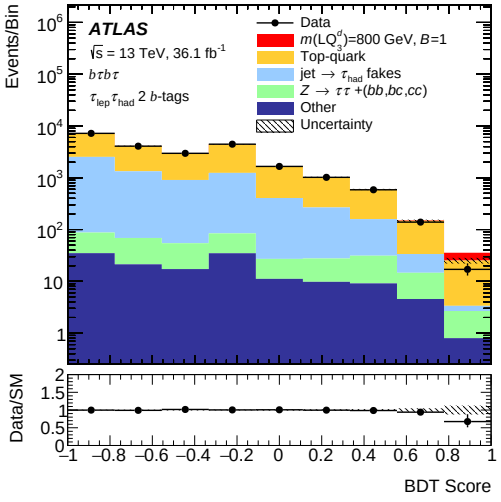
FIGURE C.3: BDT output distributions for the down-type leptoquark signal with $m(\text{LQ}_3^d) = 400$ GeV (top), $m(\text{LQ}_3^d) = 800$ GeV (middle) and $m(\text{LQ}_3^d) = 1100$ GeV (bottom), after performing the combined fit. Distributions are shown in the one b-tag signal region for the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels on the left and right respectively.



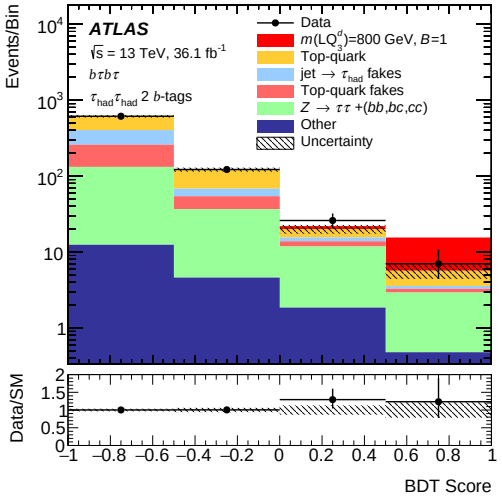
(a) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 400$ GeV



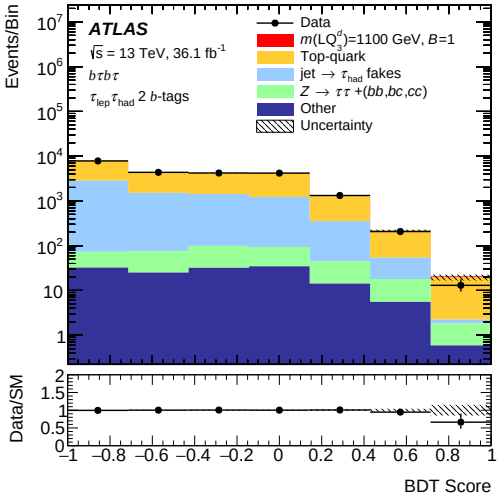
(b) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 400$ GeV



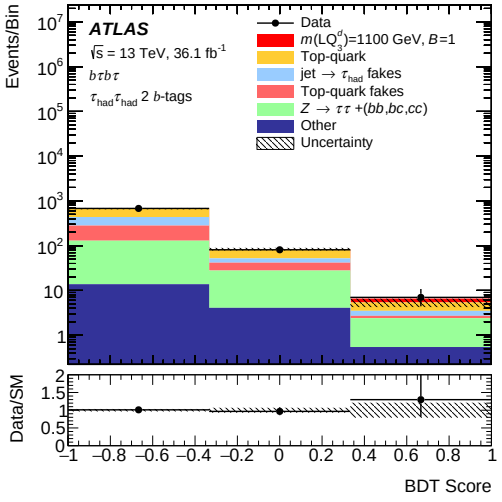
(c) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 800$ GeV



(d) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 800$ GeV



(e) $\tau_{\text{lep}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 1100$ GeV



(f) $\tau_{\text{had}}\tau_{\text{had}}$, $m(\text{LQ}_3^d) = 1100$ GeV

FIGURE C.4: BDT output distributions for the down-type leptoquark signal with $m(\text{LQ}_3^d) = 400$ GeV (top), $m(\text{LQ}_3^d) = 800$ GeV (middle) and $m(\text{LQ}_3^d) = 1100$ GeV (bottom), after performing the combined fit. Distributions are shown in the two b-tag signal region for the $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ channels on the left and right respectively.

List of Figures

List of Tables

Bibliography

[1] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.* **13** (1964) 321–323.

[2] P. W. Higgs, "Broken symmetries, massless particles and gauge fields," *Phys. Lett.* **12** (1964) 132–133.

[3] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.* **13** (1964) 508–509.

[4] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles," *Phys. Rev. Lett.* **13** (1964) 585–587.

[5] **ATLAS** Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett.* **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[6] **CMS** Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys. Lett.* **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[7] **ATLAS and CMS** Collaboration, "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments," *Phys. Rev. Lett.* **114** (2015) 191803, arXiv:1503.07589 [hep-ex].

[8] A. Banfi *et al.*, "Lecture notes for the 2016 HEP School for Experimental High Energy Physics Students," Tech. Rep. RAL-TR-2016-007, 2016. https://stfc.ukri.org/files/hep-summer-school-2016-lecture-notes/.

[9] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Addison-Wesley, Reading, USA, 1995. http://www.slac.stanford.edu/~mpeskin/QFT.html.

[10] P. Langacker, *The Standard Model and Beyond; 2nd Ed.* CRC Press, Boca Raton, USA, 2017. https://cds.cern.ch/record/2256595.

[11] **ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group** Collaboration, S. Schael *et al.*, "Precision electroweak measurements on the Z resonance," *Phys. Rept.* **427** (2006) 257–454, arXiv:hep-ex/0509008 [hep-ex].

[12] **Particle Data Group** Collaboration, "Review of particle physics," *Phys. Rev. D* **98** (2018) 030001. https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[13] P. A. M. Dirac, "The quantum theory of the electron," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **117** no. 778, (1928) 610–624. http://www.jstor.org/stable/94981.

[14] C. D. Anderson, "The positive electron," *Phys. Rev.* **43** (1933) 491–494. https://link.aps.org/doi/10.1103/PhysRev.43.491.

[15] E. Noether, "Invariant Variation Problems," *Gott. Nachr.* **1918** (1918) 235–257, arXiv:physics/0503066 [physics].

[16] O. W. Greenberg, "Spin and Unitary Spin Independence in a Paraquark Model of Baryons and Mesons," *Phys. Rev. Lett.* **13** (1964) 598–602.

[17] M. Y. Han and Y. Nambu, "Three Triplet Model with Double SU(3) Symmetry," *Phys. Rev.* **139** (1965) B1006–B1010.

[18] M. Gell-Mann, "Symmetries of baryons and mesons," *Phys. Rev.* **125** (1962) 1067–1084.

[19] D. J. Gross and F. Wilczek, "Ultraviolet behavior of non-abelian gauge theories," *Phys. Rev. Lett.* **30** (1973) 1343–1346. https://link.aps.org/doi/10.1103/PhysRevLett.30.1343.

[20] S. L. Glashow, "Partial Symmetries of Weak Interactions," *Nucl. Phys.* **22** (1961) 579–588.

[21] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc.* **C680519** (1968) 367–377.

[22] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[23] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction," *Prog. Theor. Phys.* **49** (1973) 652–657.

[24] M. Gell-Mann, "The interpretation of the new particles as displaced charge multiplets," *Nuovo Cim.* **4** no. S2, (1956) 848–866.

[25] T. Nakano and K. Nishijima, "Charge Independence for V-particles," *Prog. Theor. Phys.* **10** (1953) 581–582.

[26] J. Goldstone, "Field Theories with Superconductor Solutions," *Nuovo Cim.* **19** (1961) 154–164.

[27] **ATLAS** Collaboration, "Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton–proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment," Tech. Rep. ATLAS-CONF-2018-031, CERN, 2018. http://cds.cern.ch/record/2629412.

[28] **LHC Higgs Cross Section Working Group** Collaboration, "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector," arXiv:1610.07922 [hep-ph].

[29] **ATLAS** Collaboration, "Evidence for the Higgs-boson Yukawa coupling to τ leptons with the ATLAS detector," *JHEP* **04** (2015) 117, arXiv:1501.04943 [hep-ex].

[30] **ATLAS** Collaboration, "Evidence for the Higgs-boson Yukawa coupling to τ leptons with the ATLAS detector," *JHEP* **04** (2015) 117, arXiv:1501.04943 [hep-ex].

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

[32] **CMS** Collaboration, "Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector," *Phys. Lett.* **B779** (2018) 283–316, arXiv:1708.00373 [hep-ex].

[33] **ATLAS** Collaboration, "Cross-section measurements of the Higgs boson decaying to a pair of τ leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," Tech. Rep. ATLAS-CONF-2018-021, CERN, 2018. http://cds.cern.ch/record/2621794.

[34] **ATLAS** Collaboration, "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector," *Phys. Lett.* **B786** (2018) 59–86, arXiv:1808.08238 [hep-ex].

[35] **CMS** Collaboration, "Observation of Higgs boson decay to bottom quarks," *Phys. Rev. Lett.* **121** no. 12, (2018) 121801, arXiv:1808.08242 [hep-ex].

[36] D. A. Dicus *et al.*, "Interference effects and the use of Higgs boson pair production to study the Higgs trilinear self coupling," *Phys. Rev.* **D92** no. 9, (2015) 093003, arXiv:1504.02334 [hep-ph].

[37] S. Borowka *et al.*, "Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass Dependence," *Phys. Rev. Lett.* **117** no. 1, (2016) 012001, arXiv:1604.06447 [hep-ph].

[38] **ATLAS** Collaboration, "Standard Model Summary Plots Spring 2019," Tech. Rep. ATL-PHYS-PUB-2019-010, CERN, 2019. http://cds.cern.ch/record/2668559.

[39] **WMAP** Collaboration, "Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results," *Astrophys. J. Suppl.* **208** (2013) 20, arXiv:1212.5225 [astro-ph.CO].

[40] V. C. Rubin, N. Thonnard, and W. K. Ford, Jr., "Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/," *Astrophys. J.* **238** (1980) 471.

[41] R. Massey, T. Kitching, and J. Richard, "The dark matter of gravitational lensing," *Rept. Prog. Phys.* **73** (2010) 086901, arXiv:1001.1739 [astro-ph.CO].

[42] G. D. Starkman *et al.*, "Opening the window on strongly interacting dark matter," *Phys. Rev. D* **41** (1990) 3594–3603. https://link.aps.org/doi/10.1103/PhysRevD.41.3594.

[43] **Supernova Search Team** Collaboration, A. G. Riess *et al.*, "Observational evidence from supernovae for an accelerating universe and a cosmological constant," *Astron. J.* **116** (1998) 1009–1038, arXiv:astro-ph/9805201 [astro-ph].

[44] A. D. Sakharov, "Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe," *Pisma Zh. Eksp. Teor. Fiz.* **5** (1967) 32–35.

[45] **KamLAND** Collaboration, "First results from KamLAND: Evidence for reactor anti-neutrino disappearance," *Phys. Rev. Lett.* **90** (2003) 021802, arXiv:hep-ex/0212021 [hep-ex].

[46] **SNO** Collaboration, "Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory," *Phys. Rev. Lett.* **89** (2002) 011301, arXiv:nucl-ex/0204008 [nucl-ex].

[47] F. Bazzocchi and M. Fabbrichesi, "Little hierarchy problem for new physics just beyond the LHC," *Phys. Rev.* **D87** no. 3, (2013) 036001, arXiv:1212.5065 [hep-ph].

[48] J. R. Ellis *et al.*, "Supersymmetric Relics from the Big Bang," *Nucl. Phys.* **B238** (1984) 453–476.

[49] A. Djouadi and J. Quevillon, "The MSSM Higgs sector at a high M_{SUSY}: reopening the low tanβ regime and heavy Higgs searches," *JHEP* **10** (2013) 028, arXiv:1304.1787 [hep-ph].

[50] G. C. Branco *et al.*, "Theory and phenomenology of two-Higgs-doublet models," *Phys. Rept.* **516** (2012) 1–102, arXiv:1106.0034 [hep-ph].

[51] S. Liebler *et al.*, "The hMSSM approach for Higgs self-couplings revisited," *Eur. Phys. J.* **C79** no. 1, (2019) 65, arXiv:1810.10979 [hep-ph].

[52] A. Djouadi *et al.*, "The post-Higgs MSSM scenario: Habemus MSSM?," *Eur. Phys. J.* **C73** (2013) 2650, arXiv:1307.5205 [hep-ph].

[53] T. Kaluza, "Zum Unitätsproblem der Physik," *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* **1921** (1921) 966–972, arXiv:1803.08616 [physics.hist-ph].

[54] O. Klein, "Quantum Theory and Five-Dimensional Theory of Relativity (in German and English)," *Z. Phys.* **37** (1926) 895–906.

[55] L. Randall and R. Sundrum, "A Large mass hierarchy from a small extra dimension," *Phys. Rev. Lett.* **83** (1999) 3370–3373, arXiv:hep-ph/9905221 [hep-ph].

[56] L. Randall and R. Sundrum, "An Alternative to Compactification," *Phys. Rev. Lett.* **83** (1999) 4690–4693, arXiv:hep-th/9906064 [hep-th].

[57] M. Gabella, "The randall-sundrum model,". https://www.researchgate.net/publication/239928060_The_Randall-Sundrum_Model.

[58] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, "Phenomenology of the Randall-Sundrum Gauge Hierarchy Model," *Phys. Rev. Lett.* **84** (2000) 2080, arXiv:hep-ph/9909255 [hep-ph].

[59] K. Agashe *et al.*, "Warped Gravitons at the LHC and Beyond," *Phys. Rev.* **D76** (2007) 036006, arXiv:hep-ph/0701186 [hep-ph].

[60] J. Baglio *et al.*, "The measurement of the Higgs self-coupling at the LHC: theoretical status," *JHEP* **04** (2013) 151, arXiv:1212.5581 [hep-ph].

[61] K. Agashe, R. Contino, and A. Pomarol, "The minimal composite Higgs model," *Nucl. Phys.* **B719** (2005) 165–187, arXiv:hep-ph/0412089 [hep-ph].

[62] V. Sanz and J. Setford, "Composite Higgs Models after Run 2," *Adv. High Energy Phys.* **2018** (2018) 7168480, arXiv:1703.10190 [hep-ph].

[63] S. Dimopoulos and L. Susskind, "Mass Without Scalars," *Nucl. Phys.* **B155** (1979) 237–252.

[64] E. Farhi and L. Susskind, "Technicolor," *Phys. Rept.* **74** (1981) 277.

[65] V. Angelopoulos *et al.*, "Search for new quarks suggested by the superstring," *Nuclear Physics B* **292** (1987) 59 – 92. http://www.sciencedirect.com/science/article/pii/0550321387906377.

[66] B. Schrempp and F. Schrempp, "Light leptoquarks," *Physics Letters B* **153** no. 1, (1985) 101 – 107. http://www.sciencedirect.com/science/article/pii/0370269385914509.

[67] H. Georgi and S. L. Glashow, "Unity of All Elementary Particle Forces," *Phys. Rev. Lett.* **32** (1974) 438–441.

[68] W. Buchmüller, R. Rückl, and D. Wyler, "Leptoquarks in lepton-quark collisions," *Physics Letters B* **191** no. 4, (1987) 442 – 448. http://www.sciencedirect.com/science/article/pii/037026938790637X.

[69] **Belle** Collaboration, Y. Sato *et al.*, "Measurement of the branching ratio of $\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau$ relative to $\bar{B}^0 \to D^{*+}\ell^-\bar{\nu}_\ell$ decays with a semileptonic tagging method," *Phys. Rev.* **D94** no. 7, (2016) 072007, arXiv:1607.07923 [hep-ex].

[70] **Belle** Collaboration, M. Huschle *et al.*, "Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle," *Phys. Rev.* **D92** no. 7, (2015) 072014, arXiv:1507.03233 [hep-ex].

[71] **Belle** Collaboration, S. Hirose *et al.*, "Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$," *Phys. Rev. Lett.* **118** no. 21, (2017) 211801, arXiv:1612.00529 [hep-ex].

[72] **BaBar** Collaboration, J. P. Lees *et al.*, "Evidence for an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays," *Phys. Rev. Lett.* **109** (2012) 101802, arXiv:1205.5442 [hep-ex].

[73] **BaBar** Collaboration, J. P. Lees *et al.*, "Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons," *Phys. Rev.* **D88** no. 7, (2013) 072012, arXiv:1303.0571 [hep-ex].

[74] **LHCb** Collaboration, "Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$," *Phys. Rev. Lett.* **115** no. 11, (2015) 111803, arXiv:1506.08614 [hep-ex].

[75] L. Di Luzio and M. Nardecchia, "What is the scale of new physics behind the B-flavour anomalies?," *Eur. Phys. J.* **C77** no. 8, (2017) 536, arXiv:1706.01868 [hep-ph].

[76] U. Aydemir *et al.*, "B-decay anomalies and scalar leptoquarks in unified Pati-Salam models from noncommutative geometry," *JHEP* **09** (2018) 117, arXiv:1804.05844 [hep-ph].

[77] L. Breiman *et al.*, *Classification and Regression Trees*. Wadsworth, 1984.

[78] Y. Freund, "Boosting a weak learning algorithm by majority," *Information and Computation* **121** no. 2, (1995) 256 – 285. http://www.sciencedirect.com/science/article/pii/S0890540185711364.

[79] K. Albertsson *et al.*, "Machine Learning in High Energy Physics Community White Paper," *J. Phys. Conf. Ser.* **1085** no. 2, (2018) 022008, arXiv:1807.02876 [physics.comp-ph].

[80] A. Hoecker *et al.*, "TMVA: Toolkit for Multivariate Data Analysis," *PoS* **ACAT** (2007) 040, arXiv:physics/0703039.

[81] Y. Freund and R. E Schapire, "A short introduction to boosting," *Journal of Japanese Society for Artificial Intelligence* **14** (1999) 771–780.

[82] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*, pp. 148–156. Morgan Kaufmann Publishers Inc., 1996. http://dl.acm.org/citation.cfm?id=3091696.3091715.

[83] G. Cowan *et al.*, "Asymptotic formulae for likelihood-based tests of new physics," *Eur. Phys. J.* **C71** (2011) 1554, arXiv:1007.1727 [physics.data-an].

[84] A. L. Read, "Presentation of search results: The CL(s) technique," *J. Phys.* **G28** (2002) 2693–2704.

[85] O. S. Brüning *et al.*, *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, 2004. https://cds.cern.ch/record/782076.

[86] C. Lefèvre, "The CERN accelerator complex. Complexe des accélérateurs du CERN,". https://cds.cern.ch/record/1260465.

[87] **ATLAS** Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST* **3** (2008) S08003.

[88] **CMS** Collaboration, "The CMS Experiment at the CERN LHC," *JINST* **3** (2008) S08004.

[89] **LHCb** Collaboration, "The LHCb Detector at the LHC," *JINST* **3** (2008) S08005.

[90] **ALICE** Collaboration, "The ALICE experiment at the CERN LHC," *JINST* **3** (2008) S08002.

[91] Z. Marshall, "Simulation of pile-up in the ATLAS experiment," *Journal of Physics: Conference Series* **513** no. 2, (2014) 022024. https://doi.org/10.1088%2F1742-6596%2F513%2F2%2F022024.

[92] "CERN Annual Report 2016," tech. rep., CERN, 2017. https://cds.cern.ch/record/2270805.

[93] J. Pequenao, "Computer generated image of the whole ATLAS detector,". http://cds.cern.ch/record/1095924.

[94] **ATLAS** Collaboration, *ATLAS magnet system: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, 1997. https://cds.cern.ch/record/338080.

[95] **ATLAS** Collaboration, *ATLAS central solenoid: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1997. https://cds.cern.ch/record/331067.

[96] **ATLAS** Collaboration, J. P. Badiou, J. Beltramelli, J. M. Baze, and J. Belorgey, *ATLAS barrel toroid: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1997. https://cds.cern.ch/record/331065.

[97] **ATLAS** Collaboration, *ATLAS end-cap toroids: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1997. https://cds.cern.ch/record/331066.

[98] **ATLAS** Collaboration, *ATLAS inner detector: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, 1997. https://cds.cern.ch/record/331063.

[99] **ATLAS** Collaboration, "The ATLAS Inner Detector commissioning and calibration," *Eur. Phys. J.* **C70** (2010) 787–821, arXiv:1004.5293 [physics.ins-det].

[100] **ATLAS** Collaboration, G. Ripellino, "The alignment of the ATLAS Inner Detector in Run-2," *PoS* **LHCP2016** (2016) 196.

[101] **ATLAS** Collaboration, "ATLAS pixel detector electronics and sensors," *JINST* **3** (2008) P07007.

[102] **ATLAS** Collaboration, "ATLAS Insertable B-Layer Technical Design Report," Tech. Rep. CERN-LHCC-2010-013. ATLAS-TDR-19, 2010. https://cds.cern.ch/record/1291633.

[103] M. Turala, "The atlas semiconductor tracker," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **466** no. 2, (2001) 243 – 254. http://www.sciencedirect.com/science/article/pii/S0168900201005356.

[104] **ATLAS** Collaboration, "The ATLAS Transition Radiation Tracker," Tech. Rep. ATL-CONF-2003-012, CERN, 2003. https://cds.cern.ch/record/686973.

[105] **ATLAS** Collaboration, *ATLAS calorimeter performance: Technical Design Report*. Technical Design Report ATLAS. CERN, 1996. http://cds.cern.ch/record/331059.

[106] **ATLAS** Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, 1996. https://cds.cern.ch/record/331061.

[107] **ATLAS** Collaboration, *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, 1996. https://cds.cern.ch/record/331062.

[108] **ATLAS** Collaboration, *ATLAS level-1 trigger: Technical Design Report*. Technical Design Report ATLAS. CERN, 1998. https://cds.cern.ch/record/381429.

[109] **ATLAS** Collaboration, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical Design Report ATLAS. CERN, 2003. https://cds.cern.ch/record/616089.

[110] R. Blair *et al.*, "The ATLAS High Level Trigger Region of Interest Builder," *JINST* **3** (2008) P04001, arXiv:0711.3217 [physics.ins-det].

[111] **GEANT4** Collaboration, S. Agostinelli *et al.*, "GEANT4: A Simulation toolkit," *Nucl. Instrum. Meth.* **A506** (2003) 250–303.

[112] **ATLAS** Collaboration, "The ATLAS Simulation Infrastructure," *Eur. Phys. J.* **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[113] W. Lukas, "Fast simulation for ATLAS: Atlfast-II and ISF,". https://cds.cern.ch/record/1448165.

[114] **ATLAS** Collaboration, T. Yamanaka, "The ATLAS calorimeter simulation FastCaloSim," *J. Phys. Conf. Ser.* **331** (2011) 032053.

[115] J. Alwall *et al.*, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP* **07** (2014) 079, arXiv:1405.0301 [hep-ph].

[116] H. L. Lai *et al.*, "New parton distributions for collider physics," *Phys. Rev.* **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[117] M. Bahr *et al.*, "Herwig++ Physics and Manual," *Eur. Phys. J.* **C58** (2008) 639–707, arXiv:0803.0883 [hep-ph].

[118] S. Gieseke, C. Rohr, and A. Siodmok, "Colour reconnections in Herwig++," *Eur. Phys. J.* **C72** (2012) 2225, arXiv:1206.0041 [hep-ph].

[119] S. Borowka *et al.*, "Full top quark mass dependence in Higgs boson pair production at NLO," *JHEP* **10** (2016) 107, arXiv:1608.04798 [hep-ph].

[120] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.* **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[121] **ATLAS** Collaboration, "ATLAS Run 1 Pythia8 tunes," Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, 2014. https://cds.cern.ch/record/1966419.

[122] R. D. Ball *et al.*, "Parton distributions with LHC data," *Nucl. Phys.* **B867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[123] A. Oliveira, "Gravity particles from Warped Extra Dimensions, predictions for LHC," arXiv:1404.0102 [hep-ph].

[124] T. Mandal, S. Mitra, and S. Seth, "Pair Production of Scalar Leptoquarks at the LHC to NLO Parton Shower Accuracy," *Phys. Rev.* **D93** no. 3, (2016) 035018, arXiv:1506.07369 [hep-ph].

[125] **NNPDF** Collaboration, R. D. Ball *et al.*, "Parton distributions for the LHC Run II," *JHEP* **04** (2015) 040, arXiv:1410.8849 [hep-ph].

[126] C. Borschensky *et al.*, "Squark and gluino production cross sections in pp collisions at $\sqrt{s}$ = 13, 14, 33 and 100 TeV," *Eur. Phys. J.* **C74** no. 12, (2014) 3174, arXiv:1407.5066 [hep-ph].

[127] P. Artoisenet *et al.*, "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations," *JHEP* **03** (2013) 015, arXiv:1212.3460 [hep-ph].

[128] S. Alioli *et al.*, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP* **06** (2010) 043, arXiv:1002.2581 [hep-ph].

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

[130] J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis," *JHEP* **07** (2002) 012, arXiv:hep-ph/0201195 [hep-ph].

[131] P. Z. Skands, "Tuning Monte Carlo Generators: The Perugia Tunes," *Phys. Rev.* **D82** (2010) 074018, arXiv:1005.3457 [hep-ph].

[132] M. Czakon *et al.*, "Constraints on the gluon PDF from top quark pair production at hadron colliders," *JHEP* **07** (2013) 167, arXiv:1303.7215 [hep-ph].

[133] M. Aliev *et al.*, "HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR," *Comput. Phys. Commun.* **182** (2011) 1034–1046, arXiv:1007.1327 [hep-ph].

[134] P. Kant *et al.*, "HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions," *Comput. Phys. Commun.* **191** (2015) 74–89, arXiv:1406.4403 [hep-ph].

[135] T. Gleisberg *et al.*, "Event generation with SHERPA 1.1," *JHEP* **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[136] T. Gleisberg and S. Hoeche, "Comix, a new matrix element generator," *JHEP* **12** (2008) 039, arXiv:0808.3674 [hep-ph].

[137] F. Cascioli, P. Maierhofer, and S. Pozzorini, "Scattering Amplitudes with Open Loops," *Phys. Rev. Lett.* **108** (2012) 111601, arXiv:1111.5206 [hep-ph].

[138] S. Schumann and F. Krauss, "A Parton shower algorithm based on Catani-Seymour dipole factorisation," *JHEP* **03** (2008) 038, arXiv:0709.1027 [hep-ph].

[139] S. Hoeche *et al.*, "QCD matrix elements + parton showers: The NLO case," *JHEP* **04** (2013) 027, arXiv:1207.5030 [hep-ph].

[140] C. Anastasiou *et al.*, "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO," *Phys. Rev.* **D69** (2004) 094008, arXiv:hep-ph/0312266 [hep-ph].

[141] S. Alioli *et al.*, "NLO Higgs boson production via gluon fusion matched with shower in POWHEG," *JHEP* **04** (2009) 002, arXiv:0812.0578 [hep-ph].

[142] **ATLAS** Collaboration, "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *JHEP* **09** (2014) 145, arXiv:1406.3660 [hep-ex].

[143] L. Altenkamp *et al.*, "Gluon-induced Higgs-strahlung at next-to-leading order QCD," *JHEP* **02** (2013) 078, arXiv:1211.5015 [hep-ph].

[144] B. Hespel, F. Maltoni, and E. Vryonidou, "Higgs and Z boson associated production via gluon fusion in the SM and the 2HDM," *JHEP* **06** (2015) 065, arXiv:1503.01656 [hep-ph].

[145] R. V. Harlander *et al.*, "Soft gluon resummation for gluon-induced Higgs Strahlung," *JHEP* **11** (2014) 082, arXiv:1410.0217 [hep-ph].

[146] R. V. Harlander, S. Liebler, and T. Zirke, "Higgs Strahlung at the Large Hadron Collider in the 2-Higgs-Doublet Model," *JHEP* **02** (2014) 023, arXiv:1307.8122 [hep-ph].

[147] O. Brein, R. V. Harlander, and T. J. E. Zirke, "vh@nnlo - Higgs Strahlung at hadron colliders," *Comput. Phys. Commun.* **184** (2013) 998–1003, arXiv:1210.5347 [hep-ph].

[148] D. J. Lange, "The EvtGen particle decay simulation package," *Nucl. Instrum. Meth.* **A462** (2001) 152–155.

[149] W. Lampl *et al.*, "Calorimeter Clustering Algorithms: Description and Performance," Tech. Rep. ATL-LARG-PUB-2008-002, CERN, 2008. http://cds.cern.ch/record/1099735.

[150] **ATLAS** Collaboration, "Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data," Tech. Rep. ATLAS-CONF-2016-024, CERN, 2016. https://cds.cern.ch/record/2157687.

[151] T. Cornelissen *et al.*, "Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)," Tech. Rep. ATL-SOFT-PUB-2007-007,

ATL-COM-SOFT-2007-002, CERN, 2007.
https://cds.cern.ch/record/1020106.

[152] **ATLAS** Collaboration, "Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung," Tech. Rep. ATLAS-CONF-2012-047, CERN, 2012. https://cds.cern.ch/record/1449796.

[153] T. G. Cornelissen *et al.*, "The global χ^2 track fitter in ATLAS," *Journal of Physics: Conference Series* **119** no. 3, (2008) 032013. http://stacks.iop.org/1742-6596/119/i=3/a=032013.

[154] **ATLAS** Collaboration, "Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data," *Eur. Phys. J.* **C77** no. 3, (2017) 195, arXiv:1612.01456 [hep-ex].

[155] **ATLAS** Collaboration, "Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung," Tech. Rep. ATLAS-CONF-2012-047, CERN, 2012. http://cds.cern.ch/record/1449796.

[156] **ATLAS** Collaboration, "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV," *Eur. Phys. J.* **C76** no. 5, (2016) 292, arXiv:1603.05598 [hep-ex].

[157] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm," *JHEP* **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[158] R. Atkin, "Review of jet reconstruction algorithms," *Journal of Physics: Conference Series* **645** no. 1, (2015) 012008. http://stacks.iop.org/1742-6596/645/i=1/a=012008.

[159] **ATLAS** Collaboration, "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1," *Eur. Phys. J.* **C77** (2017) 490, arXiv:1603.02934 [hep-ex].

[160] **ATLAS** Collaboration, "Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector," *Phys. Rev.* **D96** no. 7, (2017) 072002, arXiv:1703.09665 [hep-ex].

[161] **ATLAS** Collaboration, "Tagging and suppression of pileup jets with the ATLAS detector," Tech. Rep. ATLAS-CONF-2014-018, CERN, 2014. https://cds.cern.ch/record/1700870.

[162] "Pile-up subtraction and suppression for jets in ATLAS," Tech. Rep. ATLAS-CONF-2013-083, CERN, Aug, 2013. https://cds.cern.ch/record/1570994.

[163] **DØ** Collaboration, "Observation of single top quark production." https://www-d0.fnal.gov/Run2Physics/top/singletop_observation/singletop_observation_updated.html, 2009. Accessed 18th July 2018.

[164] **ATLAS** Collaboration, "Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run," Tech. Rep. ATL-PHYS-PUB-2016-012, CERN, 2016. https://cds.cern.ch/record/2160731.

[165] **ATLAS** Collaboration, "Performance of b-Jet Identification in the ATLAS Experiment," *JINST* **11** no. 04, (2016) P04008, arXiv:1512.01094 [hep-ex].

[166] G. Piacquadio and C. Weiser, "A new inclusive secondary vertex algorithm for b-jet tagging in atlas," *Journal of Physics: Conference Series* **119** no. 3, (2008) 032032. http://stacks.iop.org/1742-6596/119/i=3/a=032032.

[167] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting," *Nucl. Instrum. Meth.* **A262** (1987) 444–450.

[168] **ATLAS** Collaboration, "Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV," arXiv:1805.01845 [hep-ex].

[169] **ATLAS** Collaboration, "Expected performance of the ATLAS b-tagging algorithms in Run-2," Tech. Rep. ATL-PHYS-PUB-2015-022, CERN, 2015. https://cds.cern.ch/record/2037697.

[170] **ATLAS** Collaboration, "Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV," arXiv:1805.01845 [hep-ex].

[171] **ATLAS** Collaboration, "Identification and energy calibration of hadronically decaying τ leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV," *Eur. Phys. J.* **C75** no. 7, (2015) 303, arXiv:1412.7086 [hep-ex].

[172] **ATLAS** Collaboration, "Reconstruction, Energy Calibration, and Identification of Hadronically Decaying τ Leptons in the ATLAS Experiment for Run-2 of the LHC," Tech. Rep. ATL-PHYS-PUB-2015-045, CERN, 2015. https://cds.cern.ch/record/2064383.

[173] **ATLAS** Collaboration, "Measurement of the τ lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV," Tech. Rep. ATLAS-CONF-2017-029, CERN, 2017. https://cds.cern.ch/record/2261772.

[174] **ATLAS** Collaboration, "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV," arXiv:1802.08168 [hep-ex].

[175] A. Elagin *et al.*, "A New Mass Reconstruction Technique for Resonances Decaying to $\tau\tau$," *Nucl. Instrum. Meth.* **A654** (2011) 481–489, arXiv:1012.4686 [hep-ex].

[176] **ATLAS** Collaboration, "Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV," Tech. Rep. ATLAS-CONF-2018-006, CERN, 2018. https://cds.cern.ch/record/2314418.

[177] **ATLAS** Collaboration, "Number of tracks vs. mu with full 2016 data," Tech. Rep. IDTR-2016-015, CERN, 2016. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2016-015.

[178] **CMS** Collaboration, "Measurement of neutral strange particle production in the underlying event in proton-proton collisions at $\sqrt{s} = 7$ TeV," *Phys. Rev. D* **88** no. CMS-QCD-11-010. CMS-QCD-11-010. CERN-PH-EP-2013-086, (2013) 052001. 35 p. https://cds.cern.ch/record/1551364.

[179] **ATLAS** Collaboration, "Study of the material of the ATLAS inner detector for Run 2 of the LHC," *JINST* **12** no. 12, (2017) P12009, arXiv:1707.02826 [hep-ex].

[180] **ATLAS** Collaboration, "Tracking Studies for b-tagging with 7 TeV Collision Data with the ATLAS Detector," Tech. Rep. ATLAS-CONF-2010-070, CERN, 2010. https://cds.cern.ch/record/1281352.

[181] **ATLAS** Collaboration, "Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run," Tech. Rep. ATL-PHYS-PUB-2017-013, CERN, 2017. http://cds.cern.ch/record/2273281.

[182] **ATLAS** Collaboration, "Primary vertex reconstruction with the ATLAS detector," Tech. Rep. ATL-PHYS-PROC-2016-163. 12, CERN, 2016. https://cds.cern.ch/record/2222390.

[183] L. A. Harland-Lang *et al.*, "Parton distributions in the LHC era: MMHT 2014 PDFs," *Eur. Phys. J.* **C75** no. 5, (2015) 204, arXiv:1412.3989 [hep-ph].

[184] S. Dulat *et al.*, "New parton distribution functions from a global analysis of quantum chromodynamics," *Phys. Rev.* **D93** no. 3, (2016) 033006, arXiv:1506.07443 [hep-ph].

[185] J. Butterworth *et al.*, "PDF4LHC recommendations for LHC Run II," *J. Phys.* **G43** (2016) 023001, arXiv:1510.03865 [hep-ph].

[186] **ATLAS** Collaboration, "Search for resonances decaying to a W or Z boson and a Higgs boson in the $\nu\nu bb$, $\ell\nu bb$ and $\ell\ell bb$ final states with $\sqrt{s} = 13$ TeV ATLAS data in the context of models with 2 Higgs doublets or additional heavy vector triplets," Tech. Rep. ATL-COM-PHYS-2016-479, CERN, 2016. https://cds.cern.ch/record/2151842.

[187] **ATLAS** Collaboration, "Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector," *Phys. Lett.* **B786** (2018) 59–86, arXiv:1808.08238 [hep-ex].

[188] **ATLAS** Collaboration, "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector," *Phys. Lett.* **B784** (2018) 173–191, arXiv:1806.00425 [hep-ex].

[189] **ATLAS** Collaboration, "Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC," *Eur. Phys. J.* **C76** no. 12, (2016) 653, arXiv:1608.03953 [hep-ex].

[190] **ATLAS** Collaboration, "Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC," *Phys. Rev. Lett.* **117** no. 18, (2016) 182002, arXiv:1606.02625 [hep-ex].

[191] **ATLAS** Collaboration, M. Aaboud *et al.*, "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data," *JINST* **14** no. 03, (2019) P03017, arXiv:1812.03848 [hep-ex].

[192] **ATLAS** Collaboration, "Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV," Tech. Rep. ATL-PHYS-PUB-2015-015, CERN, 2015. https://cds.cern.ch/record/2037613.

[193] **ATLAS** Collaboration, "Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Phys. Rev. Lett.* **121** no. 19, (2018) 191801, arXiv:1808.00336 [hep-ex].

[194] **ATLAS** Collaboration, "Combination of searches for Higgs boson pairs in pp collisions at 13 TeV with the ATLAS experiment.," Tech. Rep. ATLAS-CONF-2018-043, CERN, 2018. http://cds.cern.ch/record/2638212.

[195] **ATLAS Collaboration** Collaboration, "Measurement prospects of the pair production and self-coupling of the Higgs boson with the ATLAS experiment at the HL-LHC," Tech. Rep. ATL-PHYS-PUB-2018-053, CERN, 2018. https://cds.cern.ch/record/2652727.

[196] **ATLAS** Collaboration, "Searches for third-generation scalar leptoquarks in $\sqrt{s}$ = 13 TeV pp collisions with the ATLAS detector," arXiv:1902.08103 [hep-ex].