A search for the Standard Model Higgs boson produced in association with dileptonically decaying top quarks and decaying into a pair of bottom quarks with the ATLAS detector

Callum Robert Kilby

Department of Physics

Royal Holloway, University of London

A thesis submitted to the University of London

for the degree of Doctor of Philosophy

March 25, 2019

Declaration

I, Callum Robert Kilby, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the document.

Callum Robert Kilby

Abstract

A search for the Standard Model Higgs boson produced in association with a pair of top quarks, where the Higgs boson decays to a pair of bottom quarks, is presented. The search uses 36.1 fb^{-1} of p-p collision data at a centre-of-mass energy of $\sqrt{s} = 13$ TeV collected using the ATLAS detector at the Large Hadron Collider in 2015 and 2016. Focus is placed on the search for the final state where decays of both top quarks produce charged leptons, referred to as the dileptonic final state. The development and study of the statistical fit used to measure the presence of collision events containing Higgs bosons is presented. Results from the combination of the search for the dileptonic final state with the search for the related single-lepton final state are presented. This combined search measured a signal cross-section of $\mu = 0.84^{+0.64}_{-0.61}$ times the Standard Model expectation. This result was later combined with searches for final states containing other Higgs boson decays to produce an observation of the production of the Standard Model Higgs boson in association with a pair of top quarks. Developments to the portion of the ATLAS trigger system responsible for track reconstruction are also presented.

Acknowledgements

A first thank you goes to my supervisor Pedro Teixeira-Dias for his ever-thoughtful and patient supervision and guidance, even amongst myriad responsibilities as head of department. Thanks also goes to Michele Faucci-Giannelli for providing day-to-day guidance and supervision. Further thanks goes to the rest of the academics of the Royal Holloway HEP group for numerous additional guidance and insights.

Heartfelt thanks go to my Mum and Dad, Merryn Woff and Robert Kilby, for the unwavering care and support throughout all the challenges a PhD brings.

Thanks also go to the friends and colleagues I've met along the way, not least for the expert knowledge of the restaurants of Geneva.

Finally, thanks go to the STFC for providing the funding and opportunity to experience the forefront of scientific research.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The running of the Large Hadron Collider (LHC) [1] and its associated particle detecting experiments since 2009 has pushed the boundaries of measurements and searches made in experimental high energy physics. The key result of these experimental analyses so far came in 2012 with the observation of the long theorised Higgs boson by the ATLAS [2] and CMS [3] experiments. The discovery of this particle completed the picture predicted by the Standard Model (SM) of particle physics [4–7]. However, further work was needed to measure the properties of the newly observed boson and examine whether it fully agreed with the Standard Model prediction. The mass of the Higgs boson [8–10], its spin and parity [11–15], and some of its production and decay rates and coupling to particles [16] have been measured and found to be consistent with the SM predictions (see also references within the cited publications). The mass was measured to be close to 125 GeV/c^2 and the particle was measured to have zero spin and positive parity. However, the production of the Higgs boson in association with top quarks, labelled as the $t\bar{t}H$ process, had not been observed after the first run of the LHC (Run 1, 2009 – 2013). This Higgs production mechanism has therefore been one focus of recent LHC data taking and analysis.

This thesis describes the search performed using the ATLAS detector during Run 2 (2015 – 2018) of the LHC for the production of the Higgs boson in association with top quarks in the case that the Higgs boson decays to a pair of bottom quarks, labelled $t\bar{t}H(b\bar{b})$. Focus is placed on the final state where both top quarks decay leptonically because this is the channel of the analysis that the author contributed to. As well, developments by the author to the portion of the ATLAS trigger system which reconstructs tracks of charged particles, referred to as the inner detector trigger, for triggers targeting

hadronically decaying tau leptons, is described.

This thesis is organised as follows:

- Chapter 2 describes the theory of the Standard Model of particle physics in terms of the fundamental particles and forces. The generation of mass through electroweak symmetry breaking and the associated Higgs boson is then described. This is followed by a discussion of the importance of the coupling of the Higgs boson to top quarks. Finally the mechanisms for producing Higgs bosons at the LHC are described with focus on the $t\bar{t}H$ mechanism of relevance to this thesis;

- Chapter 3 gives an overview of the LHC and the ATLAS detector. The different sub-systems of ATLAS are described individually based upon the role they play in selecting and reconstructing collision events;

- Chapter 4 explains the generic (i.e. non-analysis-specific) methods used to produce simulated LHC collisions and their interactions with the ATLAS detector. This chapter also describes the generic methods used to reconstruct physics objects based on measurements made by the ATLAS detector. The portion of the chapter dedicated to simulation is split between the simulation of collision events and the simulation of interactions with the ATLAS detector. The portion of the chapter dedicated to reconstruction from measurements made by ATLAS is split into the different physics signatures that are reconstructed;

- Chapter 5 details the inner detector trigger portion of the ATLAS trigger system and the developments made to this sub-system for tau lepton triggers by the author. The tau lepton triggers in question are described along with a novel 'two-stage tracking' method which is used in these triggers. The author's work is then described in three areas of development. Firstly, in improving selection efficiency associated with the novel two-stage tracking method. Secondly, in understanding reductions in the efficiency of track-finding in a low-p_{T} region and if changes were required to eliminate such reductions. Thirdly, in parametrising the novel two-stage tracking method based on track seeds used in the track finding and estimating associated reductions in computation time;

- Chapter 6 presents the ATLAS $t\bar{t}H(b\bar{b})$ analysis in Run 2 of the LHC and describes the author's work as part of this analysis. The challenges faced by the analysis are described, which are followed by an overview of the analysis and methods used to mitigate these challenges. Detailed descriptions of the definitions, datasets, and systematic uncertainties used in the analysis are then given. Contributions by the author are presented as sub-chapters in these descriptions: firstly in the choice of triggers used by the analysis in 2016 ATLAS data, secondly in the estimation of the category of events containing fake or non-prompt leptons, and thirdly in the estimation of systematic uncertainties which affect distribution shapes and event acceptance associated with the choice of parton distribution functions in simulated datasets. A detailed description of the statistical fit used in the analysis is then given along with presentation of a significant part of the author's work in the development of this fit. Different sub-chapters of this section show different stages of the fit and studies performed by the author as part of its development. Finally, results are presented from the published ATLAS $t\bar{t}H(b\bar{b})$ analysis as well as mention of the use of these results in published evidence for the $t\bar{t}H$ process from a combination of a number of ATLAS $t\bar{t}H$ searches.

Chapters 2, 3, 4 do not contain any work by the author and instead summarise the background theoretical and experimental results and methods required to describe the specific work covered by this thesis. Information is sourced from the relevant theoretical and experimental literature and citations are provided as appropriate.

Chapter 5 is built on methods developed previously by the ATLAS Collaboration, however the work presented here is entirely the author's own.

The ATLAS $t\bar{t}H(b\bar{b})$ analysis described in Chapter 6 was worked on by a team of analysers of which the author was one member; as such the overall analysis methods and analysis definitions, datasets, and systematic uncertainties can be taken to have been produced by the analysis team as a whole and not the author's direct contribution. The work presented in Sections 6.4.1, 6.5.1, 6.6.3 is entirely the author's own, except for the implementation of the PDF systematic uncertainties in the statistical fit at the end of Section 6.6.3 which was performed by others in the analysis group but is included to show

the final outcome of the author's work. The work on developing the statistical fit presented in Section 6.7 is the author's own, however these developments were performed concurrently with others in the analysis group and sometimes involved implementation of recommendations from the work of others in the group (for example, implementing systematic uncertainties which were derived by others). As such, this section includes both the author's own work and the author's own implementations in the statistical fit of developments by others. On the other hand all results presented in this section (i.e. outputs of statistical fits) are entirely the author's own.

In general if no citation is provided for a figure or table then it can be assumed to have been produced by the author.

The natural unit convention of $c = \hbar = 1$ is used hereafter in this thesis, which results in energy, mass, and momentum all being quoted in units of eV.

Chapter 2

Theory and Motivation

2.1 The Standard Model of Particle Physics

The Standard Model of particle physics (SM) is the theory that describes the so-far observed elementary particles of the Universe as well as their non-gravitational interactions [4–7]. The theory can be used to predict the rates and kinematics of particle interactions as well as the rates of particle decays. Bound states of multiple elementary particles can also be predicted and described according to the interactions described in the Standard Model. The theory also includes antiparticles as copies of the elementary particles with the sign of all physical charges and the chirality of the particles inverted.

The Standard Model is constructed from Lorentz invariant Lagrangian equations describing relativistic quantum fields and their interactions. These Lagrangian equations are renormalisable, where divergences from higher order interaction terms are contained in effective couplings at a specified renormalisation scale, therefore allowing finite calculations in terms of the renormalised parameters [17]. The fundamental particles of the Standard Model are described as quantised excitations of these fields. The fundamental constituents of matter are described by spinor fields associated with an intrinsic spin quantum number of $\frac{1}{2}$. The associated half-integer spin particles are labelled as 'fermions' and obey Fermi-Dirac statistics. Fermions are arranged into three 'generations' where fermions of a similar type have identical quantum numbers across generations but different masses. The non-gravitational fundamental forces are described by vector fields which are included in the Standard Model by requiring that the theory is invariant under global and local changes in phases of the fields, referred to as 'global gauge invariance'

and 'local gauge invariance' respectively. These vector fields are associated with an intrinsic spin quantum number of 1. Forces identified with the vector fields are described as being mediated by integer spin particles labelled as 'gauge bosons' which obey Bose-Einstein statistics. Finally a scalar field with a non-zero vacuum state is included which generates masses for fermions and for massive gauge bosons as described in Section 2.2. This scalar field is associated with an intrinsic spin quantum number of 0. The associated spin-zero boson is labelled as the 'Higgs boson'.

A fundamental requirement on the formulation of the Standard Model is that the laws of physics remain the same at all times and over all points in space. To impose this requirement variations in the 'gauge' of the theory, that is the choice of definitions of the fundamental fields, must leave the Lagrangian equations unchanged. The Lagrangian equations can then be said to be invariant under gauge transformations. General gauge variations can be described as phase transformations of the fields according to fundamental symmetries, given by the generic form:

$$\psi \rightarrow \psi' = e^{iqT\alpha}\psi \tag{2.1}$$

where ψ is the field, α is the change in phase, T is the generator of the fundamental symmetry[1], and q is the coupling strength of the field to the symmetry transformation. For the simplest case of a $\mathrm{U}(1)$ group symmetry the change in phase can be given by $e^{iq\alpha}$. To examine the effect of such a transformation the minimal Lagrangian which describes a massless spin-zero non-interacting field, ϕ, can be written:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^*(\partial^\mu \phi) \tag{2.2}$$

where $\mu \in \{0, 1, 2, 3\}$ which labels the spacetime coordinates, and $\partial_\mu = \partial/\partial x_\mu$. Technically this is an equation for Lagrangian density, $\mathcal{L}$, which can be integrated over the spatial coordinates $x_{1,2,3}$ to produce the Lagrangian. For the sake of brevity Lagrangian densities will simply be referred to as Lagrangians in this thesis.

Inserting $\phi \rightarrow \phi' = e^{iq\alpha}\phi$ into this Lagrangian will reproduce the same form in terms of ϕ' if α does not depend on the spacetime coordinates x_μ. Gauge transformations of

[1] In this context a generator is the operator representing an infinitesimal transformation according to the symmetry considered

this form are described as global transformations. This is because they represent a universal change in the gauge of fields in the Lagrangian. It can therefore be said that this Lagrangian is invariant under global gauge transformations. However if α is dependent on x_μ as $\alpha(x)$ then terms of the form $\partial_\mu \alpha(x)$ are obtained and the form of the Lagrangian is altered by the gauge transformation. Gauge transformations of this form are described as local transformations due to the dependence on the time and spatial position they are evaluated at.

To enforce local gauge invariance gauge fields are introduced which transform under local gauge transformations in such a way that $\partial_\mu \alpha(x)$ terms in the Lagrangian are removed and the overall form of the Lagrangian is invariant. These gauge fields are defined as vector fields, A_μ, which transform under associated gauge transformations as:

$$A_\mu \rightarrow A'_\mu = A_\mu - \partial_\mu \alpha(x) \tag{2.3}$$

To achieve local gauge invariance for the Lagrangian equations partial derivatives, ∂_μ, are replaced with the 'covariant derivative', D_μ, which contains the gauge fields:

$$D_\mu \equiv \partial_\mu + iqA_\mu \tag{2.4}$$

Through use of the full definition of D_μ in Lagrangians for fermion fields terms are obtained which represent the coupling of fermions to the gauge fields. By also adding Lagrangian terms for free gauge fields the gauge fields can be associated with gauge bosons and finally identified as describing fundamental forces in the Standard Model. The Standard Model therefore describes fundamental particles and their interactions in a gauge invariant way.

Specifically the Standard Model is invariant under transformations of the symmetry group:

$$\mathrm{SU(3)_C} \times \mathrm{SU(2)_L} \times \mathrm{U(1)_Y} \tag{2.5}$$

The strong nuclear force

The $SU(3)_C$ component corresponds to quantum chromodynamics (QCD) which describes the strong nuclear force. This force acts on fermions which have 'colour' charge and is mediated by eight massless gauge bosons labelled as 'gluons' which themselves carry colour charge. Colour-carrying fermions are labelled as 'quarks'. The gluons are identical except for the combination of colour charges carried by each gluon. Colour is described by three charges $C \in \{r, g, b\}$ as well as the colour conjugate charges $\bar{C} \in \{\bar{r}, \bar{g}, \bar{b}\}$ referred to as anti-colour. Colour-charge-neutral states are formed from the combinations of charges rgb, $\bar{r}\bar{g}\bar{b}$, or combinations of colour and associated anti-colour for example $r\bar{r}$. Quarks carry a single colour charge and gluons are formed out of combinations of colour anti-colour states.

A fundamental property of the $SU(3)_C$ group is that it is non-Abelian which results in coupling interactions between gluon fields. A consequence of this self-interaction is that the renormalised strong force coupling strength, α_s, increases at low energy and large distances and decreases at high energy and short distances. The result of this behaviour is that at long distances the energy contained in the strong field increases, and therefore colour-carrying particles must obey 'colour confinement' to avoid large quantities of energy being generated between distant particles. Colour confinement is defined as the requirement that colour-carrying particles are contained in bound states which are overall colourless. At short scales, such as within bound states, strong force interactions are significantly weaker and colour-carrying particles are said to be 'asymptotically free'.

A consequence of colour confinement is that when colour-carrying particles are produced from particle collisions such as those in the LHC, energy will build up in the strong field between particles which are said to be colour connected. This field energy will eventually produce additional particles and gluons which will move generally along the directions of the original particles and will eventually form colourless bound states. This process is referred to as 'hadronisation' and the resulting ensembles of particles are referred to as 'hadronic jets'.

A further consequence of the running of α_s with energy scale is on the calculability of particle interactions and decays involving the strong force. When the coupling

is small at high energy scales perturbative calculations are possible. However as energy scales decrease and the coupling increases perturbation theory is no longer valid and other more limited calculation methods, such as lattice QCD [18] (chapter 17 and references therein), or calculations based around approximations must be used. Lattice QCD discretises spacetime into a grid with finite spacing, which regularizes the theory and allows numerical evaluation of non-perturbative calculations. These calculations are limited by their very computationally intensive nature, and by the associated difficulty in approaching the continuum calculation limit by for example reducing the size of the grid spacing. There are also different methods of discretising the theory; all approaches must reproduce the same results in the limit that the grid spacing becomes zero, but can give different results for finite grid spacing.

Electroweak symmetry, the weak nuclear force, and the electromagnetic force

The $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$ component of Equation 2.5 corresponds to electroweak interactions. The $\mathrm{U}(1)_\mathrm{Y}$ group corresponds to couplings to 'hypercharge', Y, mediated by a massless boson B. The $\mathrm{SU}(2)_\mathrm{L}$ group corresponds to couplings to 'weak isospin' of which there are three components associated with the three fundamental transformations of this group. There are three massless bosons, $W^{1,2,3}$, which mediate $\mathrm{SU}(2)_\mathrm{L}$ interactions. An important feature of the $\mathrm{SU}(2)_\mathrm{L}$ group of the Standard Model is that its interactions are chirality dependent. The fields of $\mathrm{SU}(2)_\mathrm{L}$ only couple to fermions (anti-fermions) with left-handed (right-handed) chirality.

However, measurements from experiment show that particle interactions are actually governed by the weak nuclear force and the electromagnetic force. Furthermore, the gauge bosons of the weak force are measured to be massive. To introduce massive gauge bosons and produce the observed weak fields and electromagnetic field the $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$ symmetry is broken into a $\mathrm{U}(1)_{EM}$ symmetry according to electroweak symmetry breaking and the Higgs mechanism described in Section 2.2. To arrive at the weak and electromagnetic fields new fields are defined from mixtures of the electroweak fields according to:

$$W^{\pm} = \frac{1}{\sqrt{2}}(W^1 \mp iW^2) \tag{2.6}$$

$$Z^0 = \cos\theta_W \, W^3 - \sin\theta_W \, B \tag{2.7}$$

$$A = \sin\theta_W \, W^3 + \cos\theta_W \, B \tag{2.8}$$

where $W^{\pm}$ and Z^0 are the massive gauge bosons of the weak force, A is the massless gauge boson, labelled as the photon, of the electromagnetic force, and θ_W is the weak mixing angle. The $W^{\pm}$ bosons carry electric charge and therefore define 'charged current' weak interactions, whereas the Z^0 does not carry electric charge and therefore defines 'neutral current' weak interactions. Charged current weak interactions act as "flavour-changing" currents which can change fermionic fields from one 'flavour' to another, where flavour refers to the species of fermion field in the collection of Standard Model fermions. For quarks this includes changing flavour across fermion generations. The probabilities of these flavour-changing interactions are given by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [19, 20]:

$$\begin{aligned} V_{\text{CKM}} &= \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \\ &= \begin{pmatrix} 0.97446 \pm 0.00010 & 0.22452 \pm 0.00044 & 0.00365 \pm 0.00012 \\ 0.22438 \pm 0.00044 & 0.97359^{+0.00010}_{-0.00011} & 0.04214 \pm 0.00076 \\ 0.00896^{+0.00024}_{-0.00023} & 0.04133 \pm 0.00074 & 0.999105 \pm 0.000032 \end{pmatrix} \end{aligned} \tag{2.9}$$

where values are quoted from the Particle Data Group's 2018 review of particle physics [18]. The indices of V_{ij} label the two quarks involved in the charged-current coupling interaction in question; the definitions of the different quarks are given below.

Flavour-changing neutral currents are also possible in the Standard Model via processes which include virtual $W^{\pm}$ bosons in quantum loops. However, these processes are suppressed by the GIM mechanism [7]. The GIM mechanism allows for such interactions due to interference effects between similar processes which contain virtual quarks of different flavours and different masses; however, the interaction rates are highly reduced

either because of the difference in scale between the mass of the quarks and the mass of the $W^{\pm}$ boson, or due to the inclusion of off-diagonal CKM matrix elements.

Fermions

The fermionic content of the standard model is arranged into three generations of particles. Fermions are further categorised in two groups: quarks and leptons. Quarks carry fractional electric charges (in units of electron charge), carry colour charge, and experience flavour-changing charged current weak interactions across different generations of quarks. All quarks also carry a baryon number of $\frac{1}{3}$ which is a strictly conserved quantity in particle interactions. Leptons carry integer electric charge (again in units of electron charge) or no electric charge, do not carry colour charge, and also carry a lepton number of 1. Lepton number is considered individually for each generation of leptons, resulting in the lepton number of each generation being a conserved quantity in particle interactions; however, this generational lepton number conservation is violated in neutrino mixing and neutrino oscillations, as is discussed below. All fermions carry weak isospin and therefore experience the weak force.

Quarks are further categorised into two types: up-type quarks and down-type quarks. Flavour changing charged-current weak interactions occur through the coupling of up-type to down-type quarks. Up-type quarks carry electric charge of $+\frac{2}{3}$ while down-type quarks carry electric charge of $-\frac{1}{3}$. Organised by generation the up-type quarks are: the up quark, u, the charm quark, c, and the top quark, t. Similarly down-type quarks are: the down quark, d, the strange quark, s, and the bottom quark, b. Each flavour of quark is actually three different fields which are identical except for the colour charge carried. The quarks are also distinguished by the scale of their masses: up, down, and strange quarks are considered light-flavour quarks while charm, bottom, and top quarks are considered heavy-flavour quarks. The scale of the top quark mass is in fact large enough for top quarks to decay before hadronisation occurs, a feature that distinguishes the top quark from other quarks.

Bound states can be formed from quarks as either[2] mesons, $q\bar{q}$, or baryons, qqq. Bound states formed of quarks are collectively referred to as hadrons. Due to the constituent

[2]Additional states such as pentaquarks, $qqqq\bar{q}$, are also possible however they are not relevant to the work in this thesis.

nature of quarks in these bound states quarks are also referred to as partons; gluons are also referred to as partons due to the presence of gluons from the binding energy of the strong force in bound states.

Leptons are also categorised into two types: charged leptons and neutrinos. Flavour changing charged-current weak interactions occur through the coupling of charged leptons to neutrinos. Charged leptons carry electric charge of -1 while neutrinos do not carry electric charge. Organised by generation the charged leptons are: the electron, e, the muon, μ, and the tau, τ. Similarly the neutrinos are: the electron neutrino, ν_e, the muon neutrino, ν_μ, and the tau neutrino, ν_τ. Due to the fact that neutrinos only interact via the weak force neutrinos have very low interaction rates and therefore typically pass through matter without interacting.

In the Standard Model neutrinos are massless and only have left-handed chirality (right-handed for anti-neutrinos). However, it was experimentally observed in 1998 [21] that neutrinos can oscillate between neutrino flavours, for example $\nu_e \rightarrow \nu_\mu$. To oscillate between flavours in this manner, neutrinos must have a non-zero mass, allowing neutrino mass states to be formed of mixtures of neutrino flavour states. Because of this mixing neutrinos which are produced as one flavour can be found at a later time as a different flavour. As a result, lepton number for a specific lepton flavour is not conserved. Massive neutrinos may be added to the Standard Model in a number of ways, such as the inclusion of massive right-handed neutrinos which only interact gravitationally (labelled as 'sterile' neutrinos), which could generate neutrino masses through the 'seesaw mechanism' [22].

Summarised content of the Standard Model

The summary of all particles in the Standard Model, as well as the list of forces mediated by the gauge bosons, is shown in Table 2.1.

2.2 Electroweak Symmetry Breaking and the Higgs Mechanism

To introduce mass to the gauge bosons of the Standard Model which have been measured to be massive ($W^\pm$, Z^0) it is not possible to simply include terms in the Lagrangian representing the rest masses of the bosons. These terms would be of the form $m^2 A_\mu A^\mu$

TABLE 2.1: Fundamental particles of the Standard Model. Names and symbols are given for each particle as well as the particle mass, m, taken from Reference [18]. Charges are quoted in units of electron charge. Shown in a) are fermions (with spin $= \frac{1}{2}$) of the Standard Model, ordered by generation. Shown in b) are gauge bosons (with spin $= 1$) of the Standard Model, ordered by the fundamental force they mediate. Shown in c) is the scalar boson (with spin $= 0$) of the Standard Model.

a)

Electric Charge	Generation		
	I	II	III
	Quarks		
$+\frac{2}{3}$	Up (u), $m = 2.2$ MeV	Charm (c), $m = 1.28$ GeV	Top (t), $m = 173$ GeV
$-\frac{1}{3}$	Down (d), $m = 4.7$ MeV	Strange (s), $m = 95$ MeV	Bottom (b), $m = 4.18$ GeV
	Leptons		
-1	Electron (e), $m = 0.51$ MeV	Muon (μ), $m = 106$ MeV	Tau (τ), $m = 1.78$ GeV
0	Electron neutrino (ν_e), $m < 2$ eV	Muon neutrino (ν_μ), $m < 0.19$ MeV	Tau neutrino (ν_τ), $m < 18.2$ MeV

b)

Electric Charge	Force			
	EM	Weak		Strong
0	Photon (γ), $m = 0$	-	Z boson (Z^0), $m = 91.2$ GeV	Gluon (g), $m = 0$
± 1	-	W boson ($W^\pm$), $m = 80.4$ GeV	-	-

c)

Electric Charge	Particle
0	Higgs boson (H), $m = 125.18$ GeV

which is not invariant under the gauge transformation $A_\mu \rightarrow A_\mu - \partial_\mu \alpha(x)$. It is instead necessary to define a new field, the 'Higgs field', in such a way that mass terms for vector bosons can be generated after local gauge invariance has been imposed [23–26]. This new field, Φ, is an $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ doublet of complex scalar fields defined acording to the Lagrangian:

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}(\partial_\mu \Phi)^\dagger (\partial^\mu \Phi) - V(\Phi) \\ &= \frac{1}{2}(\partial_\mu \Phi)^\dagger (\partial^\mu \Phi) + \frac{1}{2}\mu^2 \Phi^2 - \frac{1}{4}\lambda \Phi^4 \end{aligned} \tag{2.10}$$

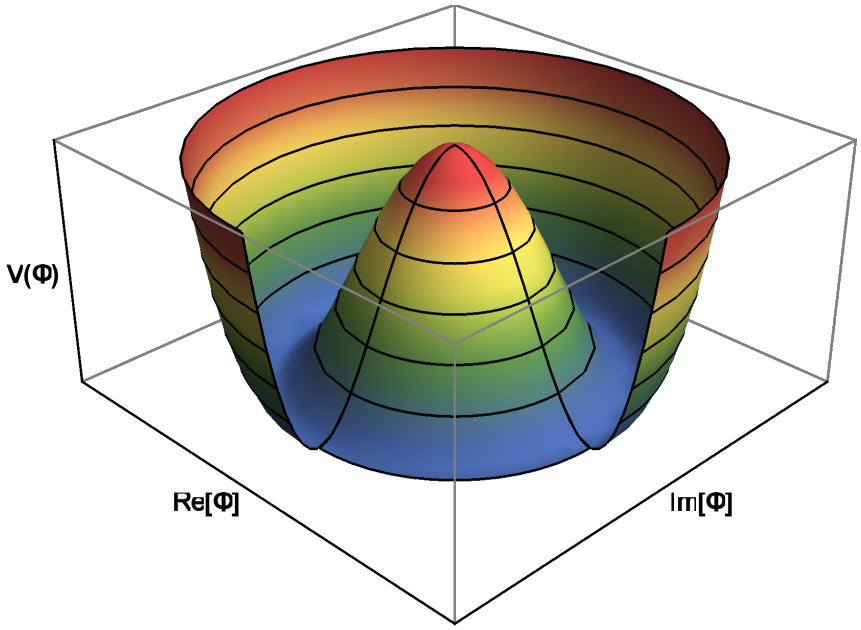


FIGURE 2.1: The shape of the Higgs potential, $V(\phi)$, in terms of its real and imaginary components. A quarter of the potential beyond the minimum is not drawn to allow the shape of the potential to be more clearly seen.

with the doublet of complex scalar fields defined as charged and uncharged complex fields:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.11}$$

An important feature of the potential in Equation 2.10, defined as $V(\Phi) = -\frac{1}{2}\mu^2\Phi^2 + \frac{1}{4}\lambda\Phi^4$ and labelled as the 'Higgs potential', is that the minimum is not at $\Phi = 0$. Instead there are minima for $\sqrt{\Phi^2} = \pm\sqrt{\mu^2/\lambda}$ which is equivalent to a continuous minimum in terms of the field components with a magnitude of $\sqrt{\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2} = \sqrt{2\mu^2/\lambda}$. The shape of this potential is shown in terms of the real and imaginary components of Φ in Figure 2.1. A local maximum can be seen at $\Phi = 0$ and the continuous minimum can be seen as a circle around this local maximum. An important consequence of this potential is that the vacuum state for Φ corresponds to a non-zero value of the field. A general vacuum expectation value for Φ can be written with the choice that at the minimum $\phi_3 = \sqrt{2}v \equiv \sqrt{2\mu^2/\lambda}$ and $\phi_{1,2,4} = 0$:

$$\langle\Phi\rangle = \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.12}$$

The field components of Φ can be redefined in terms of field fluctuations about this vacuum ground state according to $\phi_3 = v + h$ and $\phi_{1,2,4} = \xi_{1,2,4}$ where h and $\xi_{1,2,4}$ are scalar fields representing the fluctuations about $\langle\Phi\rangle$. Due to the definition of these fields about a single chosen vacuum expectation value, the $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$ symmetry is broken[3]. Variations in the field h represent fluctuations in the potential well around the chosen minimum, however variations in $\xi_{1,2,4}$ would result in the choice of a different minimum (with the same magnitude) and therefore these variations do not respect the symmetries of associated gauge transformations.

By inserting the new definitions of $\phi_{1,2,3,4}$ into the Lagrangian in Equation 2.10 a mass term associated with field h is obtained of the form $2\lambda v^2 h^2$. This therefore corresponds to a mass:

$$m_h = \sqrt{2\lambda} v \tag{2.13}$$

However no mass terms are obtained for the $\xi_{1,2,4}$ fields. These fields are therefore associated with massless bosons which are referred to as 'Goldstone bosons' following the theorem that there will exist a massless spin-zero boson for each symmetry that is broken at the vacuum state [27]. The massive field h, resulting from the one unbroken symmetry, is associated with a massive spin-zero boson labelled as the 'Higgs boson', H.

By imposing that the Lagrangian in Equation 2.10 is invariant under local $\mathrm{SU}(2)_\mathrm{L} \times \mathrm{U}(1)_\mathrm{Y}$ gauge variations the partial derivatives are replaced with covariant derivatives containing the electroweak gauge fields:

$$D_\mu \equiv \partial_\mu + \frac{i}{2} g' Y B_\mu + \frac{i}{2} g_W \sigma^a W_\mu^a \tag{2.14}$$

where $a \in \{1, 2, 3\}$, B_μ is the $\mathrm{U}(1)_\mathrm{Y}$ gauge field, W_μ^a are the $\mathrm{SU}(2)_\mathrm{L}$ gauge fields, σ^a are the Pauli matrices which are generators for the $\mathrm{SU}(2)_\mathrm{L}$ group, g' is the $\mathrm{U}(1)_\mathrm{Y}$ coupling strength, g_W is the $\mathrm{SU}(2)_\mathrm{L}$ coupling strength, and Y is the hypercharge of the scalar field.

By inserting the definitions of $\phi_{1,2,3,4}$ in terms of h and $\xi_{1,2,4}$ into this locally invariant Lagrangian, terms are obtained describing the massive boson h and the massless Goldstone bosons $\xi_{1,2,4}$. However due to the presence of the gauge fields in the Lagrangian,

[3]Technically the symmetry has not been broken and has instead been "hidden" by the new field definitions.

terms are obtained which describe the coupling of the gauge fields to h and $\xi_{1,3,4}$, and significantly terms are obtained which include only the vacuum expectation value and the gauge fields. By choosing an appropriate gauge using transformations associated with the symmetries that are broken about the vacuum state the field Φ can be defined as:

$$\Phi' = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix} \tag{2.15}$$

In this gauge the definition of h remains the same, however $\xi_{1,2,4} = 0$. The structure of the Lagrangian remains the same, but all terms containing $\xi_{1,2,4}$ are removed. By defining new gauge boson fields formed of mixtures of B_μ and W^a_μ as $W^\pm$ and Z^0 fields (the same mixtures[4] described in Section 2.1), Lagrangian terms are obtained describing the coupling of the new $W^\pm$ and Z^0 gauge fields to h, as well as terms which only contain the vacuum expectation value and the $W^\pm$ and Z^0 fields. In this gauge the latter terms have the form $(\frac{1}{2} g_W v)^2 |W^\pm|^2$ and $(\frac{1}{2} v \sqrt{g_W^2 + g'^2})^2 (Z^0)^2$. These terms can be identified as mass terms of the physical massive gauge bosons $W^\pm$ and Z^0. These gauge fields are associated with the three broken symmetries about the vacuum state and the degrees of freedom previously associated with the Goldstone bosons are instead included as longitudinal polarisations of the $W^\pm$ and Z^0 bosons. The remaining unbroken symmetry is associated with the massless photon field A. In this way the electroweak symmetry $\mathrm{SU(2)_L \times U(1)_Y}$ has been broken into a single symmetry $\mathrm{U(1)_{EM}}$ of the electromagnetic force. In doing so the measured masses of the $W^\pm$ and Z^0 bosons have been generated from the vacuum expectation value, a process labelled as the 'Higgs mechanism'. A number of self-interaction and gauge field interaction terms have also been generated for the Higgs boson. These terms define interactions of the form: HHH; $HHHH$; HVV; and $HHVV$, where $V = W^\pm$ or Z^0.

From the Lagrangian terms for the massive gauge bosons, the masses of the $W^\pm$ bosons are:

$$m_W = \frac{1}{2} g_W v \tag{2.16}$$

and the mass of the Z^0 boson is:

[4] Definitions of $\cos\theta_W = g/\sqrt{g^2 + g'^2}$ and $\sin\theta_W = g'/\sqrt{g^2 + g'^2}$ are used here, however this does not change the definitions of the new gauge fields.

$$m_Z = \frac{1}{2} v \sqrt{g_W^2 + g'^2} \tag{2.17}$$

Mass terms for fermions in the Standard Model can also be generated by the symmetry breaking of the Higgs field vacuum state. This is necessary because it is not possible to simply insert mass terms for fermions in the Standard Model. These mass terms would have the form $m\overline{\psi}\psi = m(\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L)$ where L and R label the left-handed and right-handed chiral components of the fermion fields and $\overline{\psi}$ is the Dirac adjoint spinor given by $\overline{\psi} = \psi^\dagger\gamma^0$, where γ^0 is the time-like Dirac matrix. Terms of this form would be invariant under local gauge transformations which treat left-handed and right-handed components the same, but due to the chiral nature of $\mathrm{SU(2)_L}$ the mass terms will not be invariant under $\mathrm{SU(2)_L}$ gauge transformations.

Interaction terms between fermions and the Higgs field are inserted into the Standard Model Largrangian with the form $y_f\overline{\psi}_{f,L}\Phi\psi_{f,R}$ (and similar hermitian conjugate terms), where y_f is the coupling strength and f labels the fermion under consideration. Interactions of this form are referred to as 'Yukawa couplings'. These terms are invariant under local $\mathrm{SU(2)_L}$ gauge transformations. When Φ is again defined in terms of h and $\xi_{1,2,4}$ in a gauge where $\xi_{1,2,4} = 0$ and the Yukawa coupling terms are expanded out, Lagrangian terms are obtained of the form $\frac{1}{\sqrt{2}}y_f\, v\, \overline{\psi}_{f,L}\psi_{f,R}$ (and similar hermitian conjugate terms). Fermion masses have therefore been generated through couplings to the Higgs field and have the form:

$$m_f = \frac{1}{\sqrt{2}} y_f\, v \tag{2.18}$$

Interaction terms for couplings between fermions and the Higgs boson are also obtained representing interactions of the form ffH with coupling strength $\frac{1}{\sqrt{2}}y_f$. By rearranging Equation 2.18 to be a definition of y_f it is seen that the coupling strengths are linearly dependent on fermion mass. As such the Higgs boson will couple more strongly to heavier fermions.

2.3 The Coupling of the Higgs Boson to Top Quarks

The strength of the coupling of the Higgs boson to top quarks, labelled as the 'Yukawa top coupling' y_t, has important cosmological implications as well as potential implications on the scale of new physics [28]. This is due to a dependence on Yukawa couplings in the running of the renormalised Higgs self-interaction coupling, $\lambda(\mu)$, as a function of the renormalisation scale, μ (note that here μ refers to a chosen energy scale and not the "mass" coupling from Equation 2.10). If $\lambda(\mu)$ becomes negative an additional minimum can be generated in the Higgs field, corresponding to another vacuum state. For a negative $\lambda(\mu)$ term of sufficient magnitude the additional potential minimum can occur with a lower ground state than the original minimum. The original vacuum can then be considered metastable, with a possibility of the Higgs field quantum tunnelling into the true ground state of the second minima [29, 30].

The Yukawa top coupling has a significant role in determining whether $\lambda(\mu)$ can acquire negative values. The running of $\lambda(\mu)$ with respect to Yukawa couplings goes as:

$$\frac{\mathrm{d}\lambda(\mu)}{\mathrm{d}\ln\mu} \propto -y_f^4 \tag{2.19}$$

Therefore if Yukawa couplings are of sufficient size to contribute more dominantly than positive terms in the running of $\lambda(\mu)$, negative values of $\lambda(\mu)$ can be produced. Due to the dependence of Yukawa couplings on particle mass and the fact that the top quark mass is far greater than the mass of other fermions, the dominant contribution to Equation 2.19 will be from the Yukawa top coupling. The dependence of $\lambda(\mu)$ on the renormalisation scale for different values of y_t can be seen in Figure 2.2.

A possible avoidance of vacuum metastability is the inclusion of new physics which adds additional positive terms to the running of $\lambda(\mu)$, thereby cancelling the effect of the negative fermion terms. This becomes particularly significant if y_t is measured to be large. In this scenario the additional minima in the Higgs potential can occur at relatively low energies, resulting in vacuum lifetimes that are smaller than the age of the Universe. In this scenario it is clear that new physics must be included at energy scales which cancel the effect of the Yukawa top term. In this way the energy scale expected for new physics could be estimated.

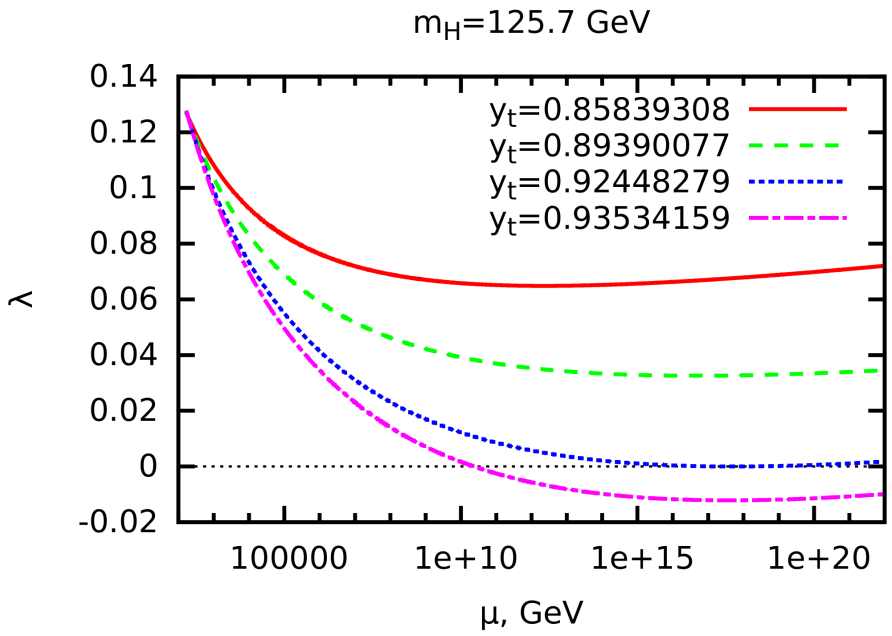


FIGURE 2.2: Dependence of the renormalised Higgs self-coupling parameter, λ, on the renormalisation scale, μ. Distributions are presented for a Higgs boson mass of 125.7 GeV and for different choices of the coupling strength between the Higgs boson and top quarks, y_t. [28]

2.4 Higgs Production at the LHC and the $t\bar{t}H$ Production Mechanism

In p-p collisions, such as at the LHC, the Higgs boson is dominantly produced through four production mechanisms: gluon–gluon fusion (ggF), vector boson fusion (VBF), production in association with a vector boson (WH and ZH), and production in association with a top quark pair ($t\bar{t}H$). Example tree level Feynman diagrams are shown for these processes in Figure 2.3. Production of the Higgs boson in association with a single top quark (tH) is also possible, although sub-dominant compared to other production mechanisms. Production in association with pairs of bottom quarks is also possible however this process is difficult to separate from background processes and is not relevant to the work presented in this thesis. The cross-section of each mechanism is shown as a function of centre-of-mass energy for p-p collisions in Figure 2.4.

The $t\bar{t}H$ production mechanism has the useful feature that the production cross-section is directly proportional to the Yukawa top coupling that is included via the vertex between top quarks and the Higgs boson. Therefore measurement of the $t\bar{t}H$ production cross-section allows for direct measurement of y_t. This is in contrast to the ggF production mechanism which includes a loop of fermions. Top quarks dominantly contribute to

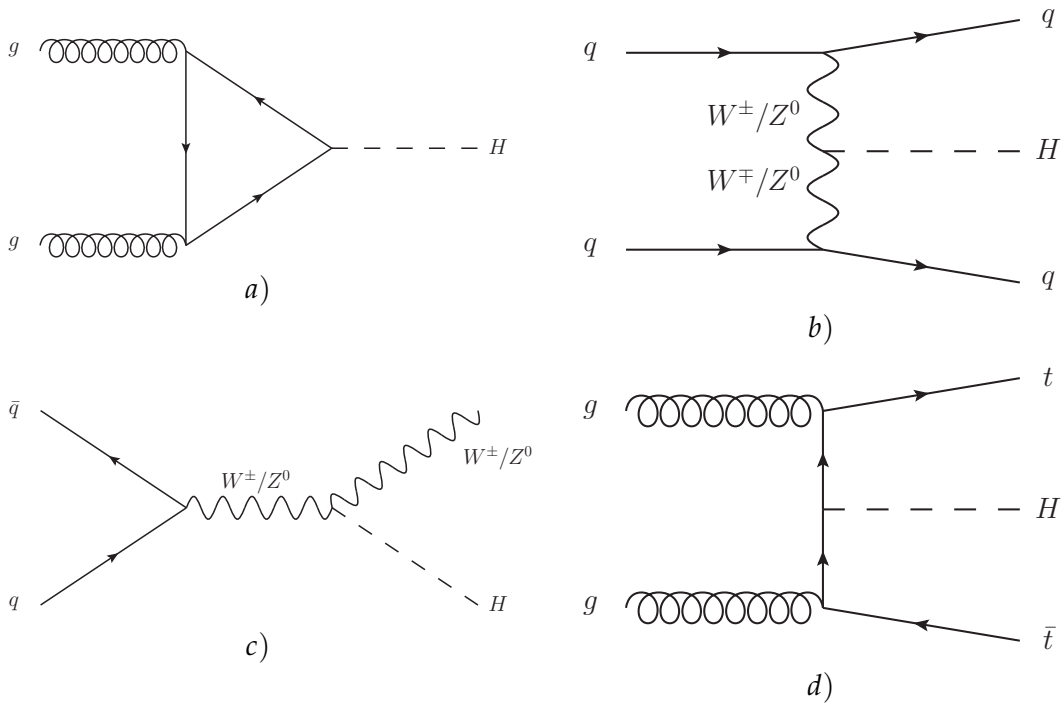


FIGURE 2.3: Example tree level Feynman diagrams for the dominant Higgs boson production mechanisms in p-p collisions. The gluon-gluon fusion mechanism is shown in a), the vector boson fusion mechanism is shown in b), the associated vector boson mechanism is shown in c), and the associated top quark pair mechanism is shown in d).

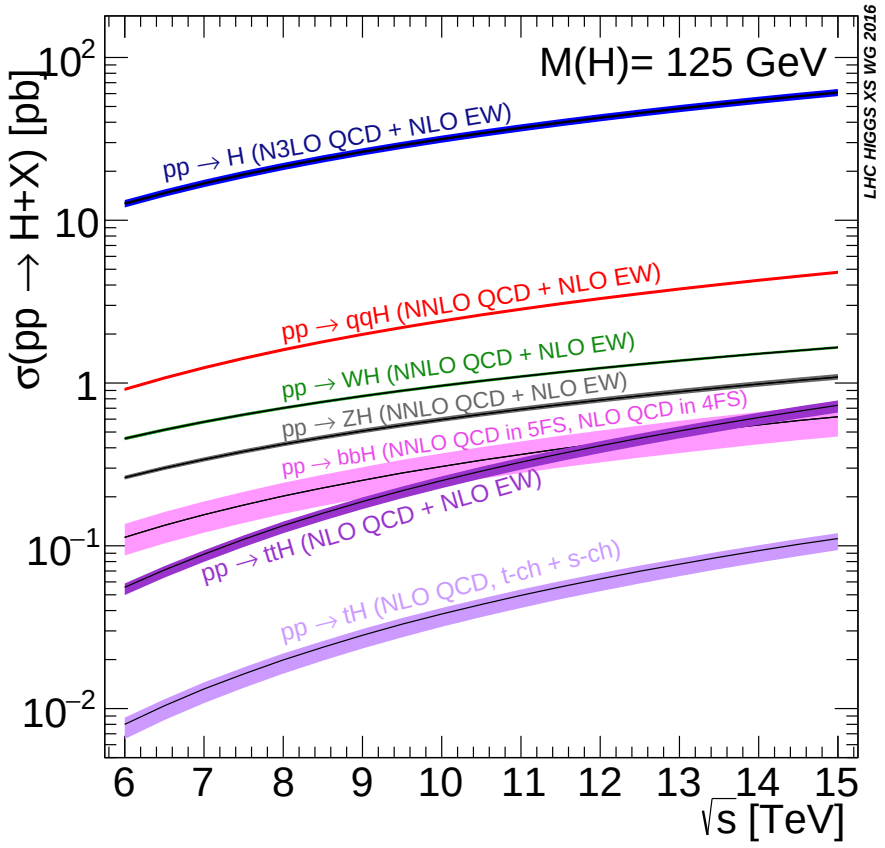


FIGURE 2.4: Higgs boson production cross-sections as a function of centre-of-mass energy [31].

this loop, meaning that the ggF cross-section is dependent on y_t. However, because other fermions also contribute to the loop, undiscovered Beyond the Standard Model (BSM) particles could also contribute to the ggF cross-section. Therefore measurements of y_t from measurements of ggF must be made under assumptions about contributions to the fermion loop.

During Run 1 of the LHC, searches for $t\bar{t}H$ production were performed by the ATLAS [32–35] and CMS collaborations [36, 37] targeting a number of different decays of the Higgs boson. The $t\bar{t}H$ production mechanism was not observed, however the combination of results from each collaboration (including other results such as measurements of ggF production) measured the significance of $t\bar{t}H$ signal to be 4.4σ, for an expected significance of 2.0σ assuming Standard Model $t\bar{t}H$ production [16]. This corresponded to a measured signal strength relative to the Standard Model prediction, $\mu_{t\bar{t}H}$, of $\mu_{t\bar{t}H} = 2.3^{+0.7}_{-0.6}$ (where $\mu_{t\bar{t}H} = 1$ corresponds to the Standard Model prediction). The individual result from ATLAS measurements was $\mu_{t\bar{t}H} = 1.9^{+0.8}_{-0.7}$, and the individual result from CMS measurements was $\mu_{t\bar{t}H} = 2.9^{+1.0}_{-0.9}$. The corresponding measurement for the Yukawa top coupling from the combined ATLAS and CMS results was 0.87 ± 0.15 relative to the Standard Model prediction, assuming no BSM particles contribute to loops, such as in ggF production. If no assumptions are made about the contribution of BSM particles to loops then the Yukawa coupling is measured to be 1.4 ± 0.2 relative to the Standard Model prediction.

Of particular relevance to the work in this thesis are results from the ATLAS and CMS Run 1 searches for $t\bar{t}H$ production in the channel that the Higgs boson decays as $H \to b\bar{b}$. The ATLAS search measured a signal significance of 1.4σ for an expected significance of 1.1σ [32], while the CMS search did not find evidence of a signal [36]. This corresponded to measurements of $\mu_{t\bar{t}H} = 1.5 \pm 1.1$ and $\mu_{t\bar{t}H} = 1.2^{+1.6}_{-1.5}$ respectively. 95% confidence limits were measured, with the ATLAS search determining $\mu_{t\bar{t}H} < 3.4$ and the CMS search determining $\mu_{t\bar{t}H} < 4.2$.

Chapter 3

The Large Hadron Collider and the ATLAS Detector

3.1 The Large Hadron Collider

The LHC is a circular high energy particle collider formed of a 27 km circumference ring located at the CERN facility near Geneva [1]. The LHC predominantly performs proton-proton collisions, labelled as p-p collisions, and operates at the frontiers of both centre-of-mass energy and collision luminosity, with a design collision centre-of-mass energy of 14 TeV and design instantaneous luminosity of 10^{34} cm^{-2} s^{-1}. These parameters allow the LHC to probe processes occurring at energies which were unreachable with previous generations of high energy particle colliders whilst collecting enough data to search for very rare processes and perform precision measurements. The LHC also performs collisions involving heavy ions; however, the studies and results presented in this thesis use data recorded only from p-p collisions. The use of protons allows for much higher beam energies than if electrons or positrons were used due to the decreased impact of synchrotron radiation, but also brings associated complications. The collision energy of the colliding partons is not fixed and is typically much lower than the collision energy of the beams; this is due to the sub-structure of the proton resulting in the momentum of each proton being spread across its constituent partons which are directly involved in the collision interactions. Related to this is the fact that proton-proton interactions will involve non-perturbative soft QCD interactions which are highly challenging to model theoretically.

To achieve cutting edge beam energies in the LHC the beams of protons are injected

from a chain of lower energy particle colliders referred to as the injection chain. This chain is built from older colliders at the CERN complex and is shown in drawing form in Figure 3.1. The collection of protons that will form proton beams is first formed by ionising hydrogen gas from a bottle of hydrogen using a hot cathode and accelerating the protons towards an anode, which provides an initial beam energy of 90 keV. The initial beam is then focused and accelerated further through a radio frequency quadrupole (RFQ) electrode chamber up to a beam energy of 750 keV and then injected into the first accelerating machine. This accelerator is the LINAC 2, which is a linear accelerator formed of 30 m of radio frequency accelerating cavities; LINAC 2 accelerates the proton beams to 50 MeV. The proton beams are then injected into the proton synchrotron booster (PSB) which splits the beams into four 157 m circumference rings and accelerates them to a beam energy to 1.4 GeV before merging them again. The beams then enter the proton synchrotron (PS), which is the first machine that was used for collisions in the past. This machine accelerates the beams around a 628 m circumferences ring to a beam energy of 25 GeV and defines the timing between proton bunches in the beams. The beams are then injected into the larger super proton synchrotron (SPS) accelerator which uses a 6.9 km ring to increase the beam energy to 450 GeV. The beams are then finally injected clockwise and counter-clockwise into the LHC where they are accelerated to the final collision energy.

To accelerate, bend, and focus the beams the LHC uses eight superconducting radio frequency accelerating cavities per beam with each delivering 2 MV of acceleration, 1,232 dipole magnets, each at a field strength of 7.7 T when beams are at collision energy, and 392 main quadrupole magnets. The beams are divided into bunches of protons spaced by 25 ns intervals, with a full bunch structure corresponding to 2,808 bunches with approximately 10^{11} protons per bunch. The high number of protons per bunch combined with the short interval between bunches, as well as focusing of the beams at the interactions points, creates the high instantaneous luminosity. For the work discussed in this thesis the beam energy was set to 6.5 TeV per beam, and hence a collision energy of 13 TeV.

The beams are focused and overlapped at four interaction points around the ring of the accelerator, which defines the positions of the four main particle detection experiments: the two general purpose detectors of ATLAS and CMS, and the two lower rate specialised detectors LHCb and ALICE. When the two colliding beams overlap a number of proton-proton interactions will occur due to the large quantity of protons in each

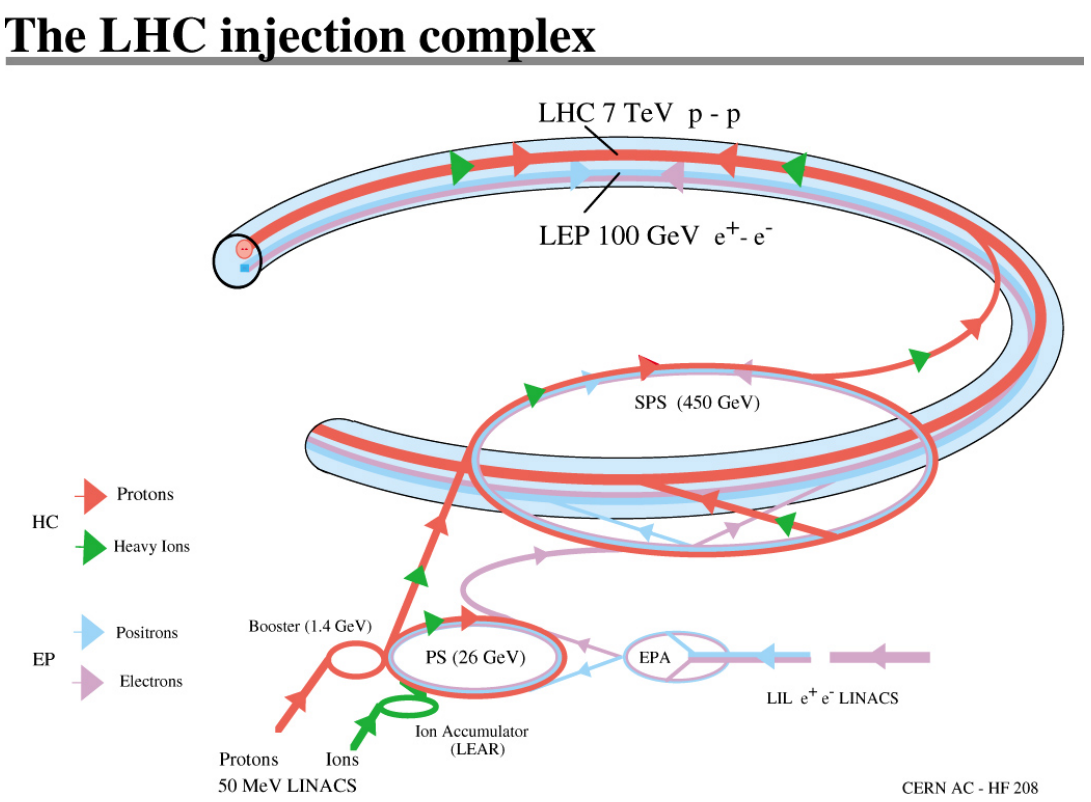


FIGURE 3.1: Drawing of the chain of particle accelerators which inject particle beams into the LHC. The paths shown for electrons and positrons refer to the Large Electron Positron (LEP) collider which predated the LHC, and are therefore not relevant. [38]

bunch, which results in multiple interaction vertices within the detectors. Due to the substructure of the proton, in most cases these interactions will be at relatively soft scales and will not produce physics of interest, however in some cases the collision will produce an interaction of interest. The other interactions are labelled as 'pile-up', and present a challenge in accurately reconstructing the interaction of interest due to the need to separate detector measurements of this interaction from those of pile-up interactions from the same collision. The point of collision for the interaction of interest is referred to as the primary vertex, while the vertices of pile-up interactions are referred to as pile-up vertices. Pile-up is quantified with the parameter μ. μ is defined as the number of p-p interactions in a given bunch crossing. Pile-up can also originate from bunch crossings before or after the bunch crossing that produces the interaction of interest due to electronic readout or detecting material relaxation that can take longer than the time between bunch crossings. Pile-up interactions originating from the same bunch crossing as the interaction of interest are referred to as 'in-time' pile-up while pile-up interactions from other bunch crossings are referred to as 'out-of-time' pile-up.

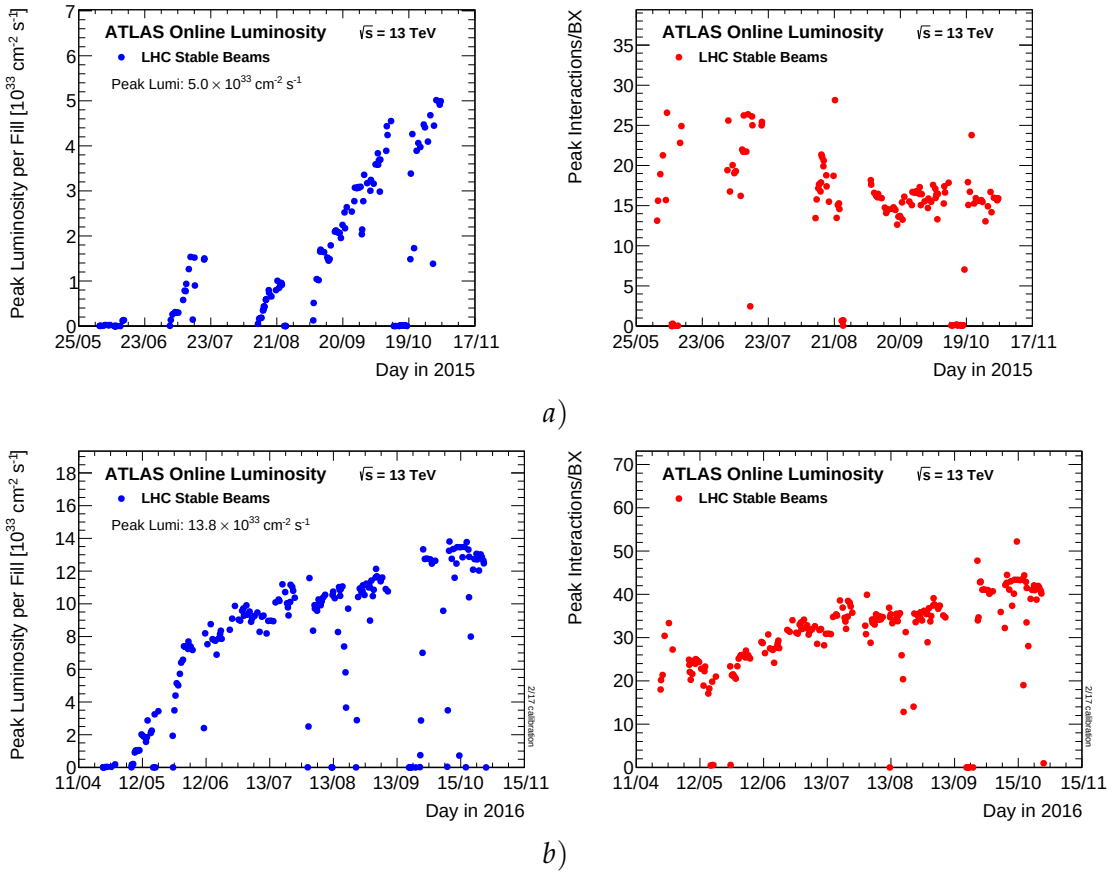


FIGURE 3.2: Distributions of peak instantaneous luminosity per LHC fill and peak number of collision vertices per LHC bunch crossing as a function of day in a specified year of data taking. Values for 2015 LHC collisions are shown in *a*) and values for 2016 LHC collisions are shown in *b*). [39]

The peak instantaneous luminosity per LHC fill and peak number of p-p interaction vertices per bunch crossing are shown for 2015 and 2016 $\sqrt{s} = 13$ TeV LHC p-p collisions delivered to the ATLAS detector in Figure 3.2. The values for the peak number of p-p interaction vertices per bunch crossing are calculated by averaging over all colliding bunches in an LHC fill and selecting the highest value. The distributions of integrated luminosity as a function of the mean number of p-p interaction vertices per bunch crossing are shown in Figure 3.3 for the full set of 2015 and 2016 $\sqrt{s} = 13$ TeV p-p collisions delivered to ATLAS. The integrated luminosity of the 2015 set is 4.2 fb^{-1} delivered to ATLAS (3.9 fb^{-1} recorded by ATLAS) and the integrated luminosity of the 2016 set is 38.5 fb^{-1} delivered to ATLAS (35.6 fb^{-1} recorded by ATLAS).

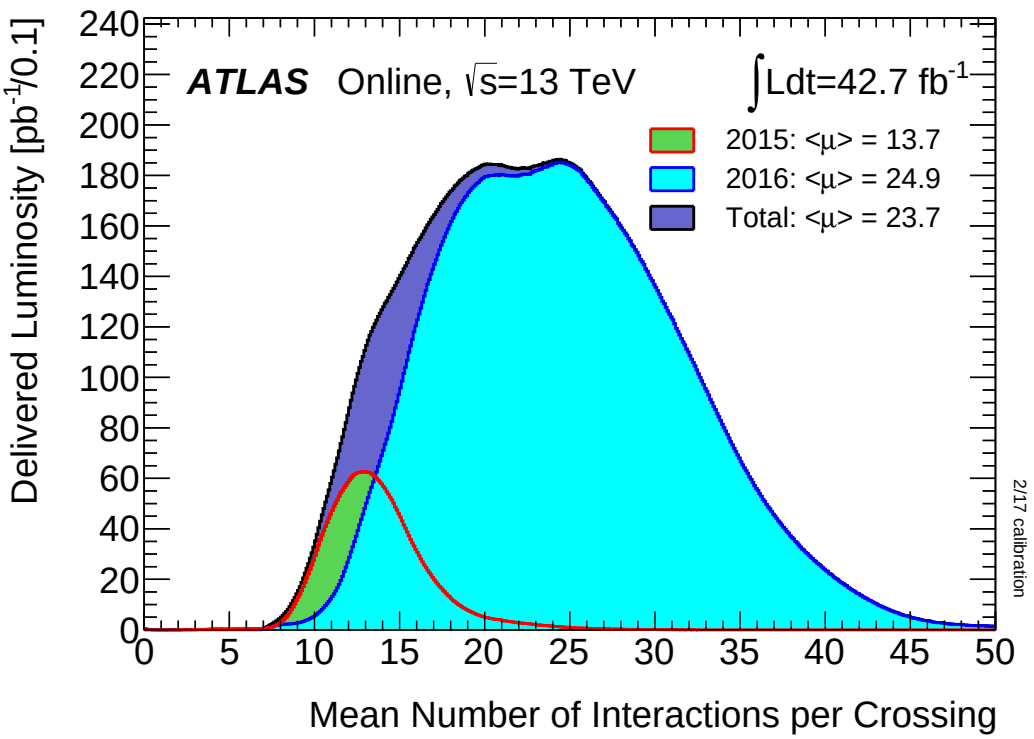


FIGURE 3.3: Distribution of integrated luminosity as a function of the mean number of collision vertices per LHC bunch crossing. Results are shown individually for the full set of 2015 (green area with red line) and 2016 (blue area with blue line) LHC collisions as well as the sum of both datasets (purple area with black line). [39]

3.2 The ATLAS Detector

The ATLAS detector is the largest general purpose particle detector in the world [40]. An overview of the ATLAS detector including a cut-away of the different sub-detector layers as well as the magnet system is shown in Figure 3.4. The detector is formed of a cylindrical barrel plus two thin barrel end-cap layers which are situated approximately 7 m along the longitudinal direction from the end-caps of the main barrel at each end; the interaction point of the LHC beams is approximately the centre of the main barrel. The overall dimensions of the detector are 44 m in length and 25 m in height and the detector weighs approximately 7000 tonnes. ATLAS is comprised of a number of sub-detector systems arranged in sequential layers. These sub-detectors are formed of either barrel or end-cap modules. To allow measurement of the momentum of charged particles the innermost layers are immersed in a 2 T magnetic field generated by a central solenoid aligned with the longitudinal axis of the barrel. Another set of toroidal magnets are used for exclusively measuring the momentum of muons; these toroids are arranged into 0.5 T barrel toroids or 1 T toroids in two end-cap modules. As well large computational resources are required to process the data collected by ATLAS and determine which events to store to ensure that data readout is within physical bandwidth and storage capacities. These computational resources are formed of multiple computer farms located near the detector and form the ATLAS trigger and data acquisition system.

The ATLAS detector uses a cartesian coordinate system where the origin lies at the interaction point between the LHC beams. The positive x-axis direction points to the centre of the LHC and the positive y-axis points upwards. The z-axis points along the beamline, which defines the x-y plane as the area used for transverse measurements (for example the transverse momentum, p_{T}, of a particle, or the transverse energy, E_{T}, of a particle). Polar coordinates are also used, with the azimuthal angle, ϕ, measured in the transverse plane around the beam axis and the polar angle, θ, measured from the beam axis. The radial distance, r, from a central axis such as the beam axis, is also commonly used. ATLAS is designed to be hermetic in the r-ϕ plane which ensures that the transverse components of collision events are completely reconstructed. In physics measurements the pseudorapidity, $\eta = -\ln(\tan(\frac{\theta}{2}))$, is usually used instead of θ. This quantity is an approximation for the rapidity of an object, defined as $y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z}$, where p_z as the z-component of an

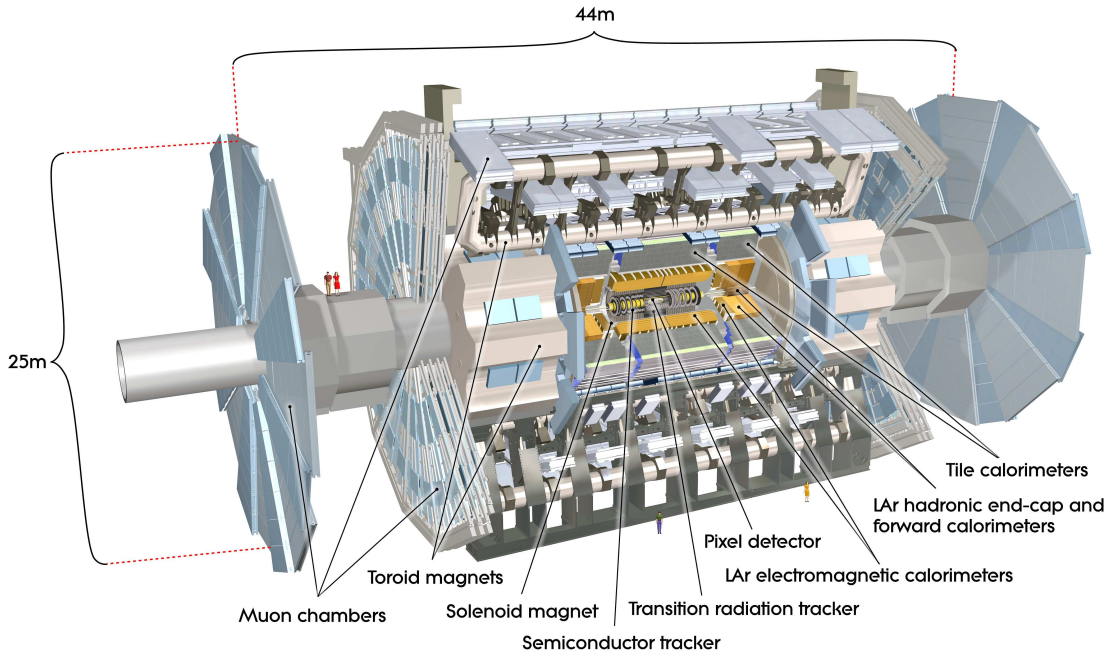


FIGURE 3.4: Overview of the layout of the ATLAS detector [40].

object's momentum. Differences in rapidity are invariant under Lorentz boosts, which allows spatial measurements that are consistent relative to each other in both the lab frame and centre-of-mass frame in question. Pseudorapidity is identical to rapidity for massless particles, or equivalently when a massive particle is highly relativistic and therefore $|\vec{p}| \gg m$, where $\vec{p}$ is the momentum of an object and m is the mass of an object. η values close to zero are referred to as 'central' pseudorapidities while larger values are referred to as 'forward' pseudorapidities. Combined differences in the η-ϕ plane are often quoted as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

Common experimental vocabulary used throughout this thesis is the description of 'objects' reconstructed by the ATLAS detector. 'Object' refers to combinations of measurements made in ATLAS modules into higher level identifications such as electrons or hadronic jets. An object may refer to lower level reconstructed information such as particle tracks or clusters of calorimeter deposits but typically the term object is used to refer to 'physics objects' such as electrons and jets.

The energy resolution and momentum resolution performance goals that motivated the design of the ATLAS detector are shown in Table 3.1.

TABLE 3.1: Energy resolution and momentum resolution goals for the ATLAS detector, categorised by the measurement type. Energy and momentum are in units of GeV. Values are taken from [40].

Measurement	Required resolution
Tracking	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$
Electromagnetic calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$
Hadronic calorimetry	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$
Muon spectrometry	$\sigma_{p_T}/p_T = 10\%$ at p_T = 1 TeV

3.2.1 Vertexing and Tracking

The sub-detectors closest to the beamline are responsible for the measurement of the paths of charged particles as tracks as well as the reconstruction of the primary vertex and resulting decay vertices, referred to as secondary vertices, using the measured tracks. The sub-detectors which perform these measurements are collectively referred to as the inner detector (sometimes abbreviated to ID). The inner detector is composed of three sub-detector layers ordered according to proximity to the beamline: the pixel detector, the semi-conducting tracker (SCT), and the transition radiation tracker (TRT). The pixel detector and SCT provide the most spatially precise measurements of track interactions, which allows for precise spatial reconstruction of tracks and vertices. The TRT provides a large number of track interactions over a large radial extent, which provides better momentum reconstruction. The TRT additionally performs electron identification through the detection of transition radiation produced by the electrons within the detector. A cutaway of the barrel layers of the inner detector is shown in Figure 3.5 and a schematic of the full inner detector is shown in Figure 3.6. The entirety of the inner detector is encased in the central solenoidal magnet of ATLAS which provides a nearly uniform magnetic field of 2 T to bend the paths of the charged particles and allow measurements of the particles' momentum to be made. The precision track measurements made by the pixel and SCT detectors cover the pseudorapidity range $|\eta| < 2.5$ and the measurements made by the TRT cover the range $|\eta| < 2$.

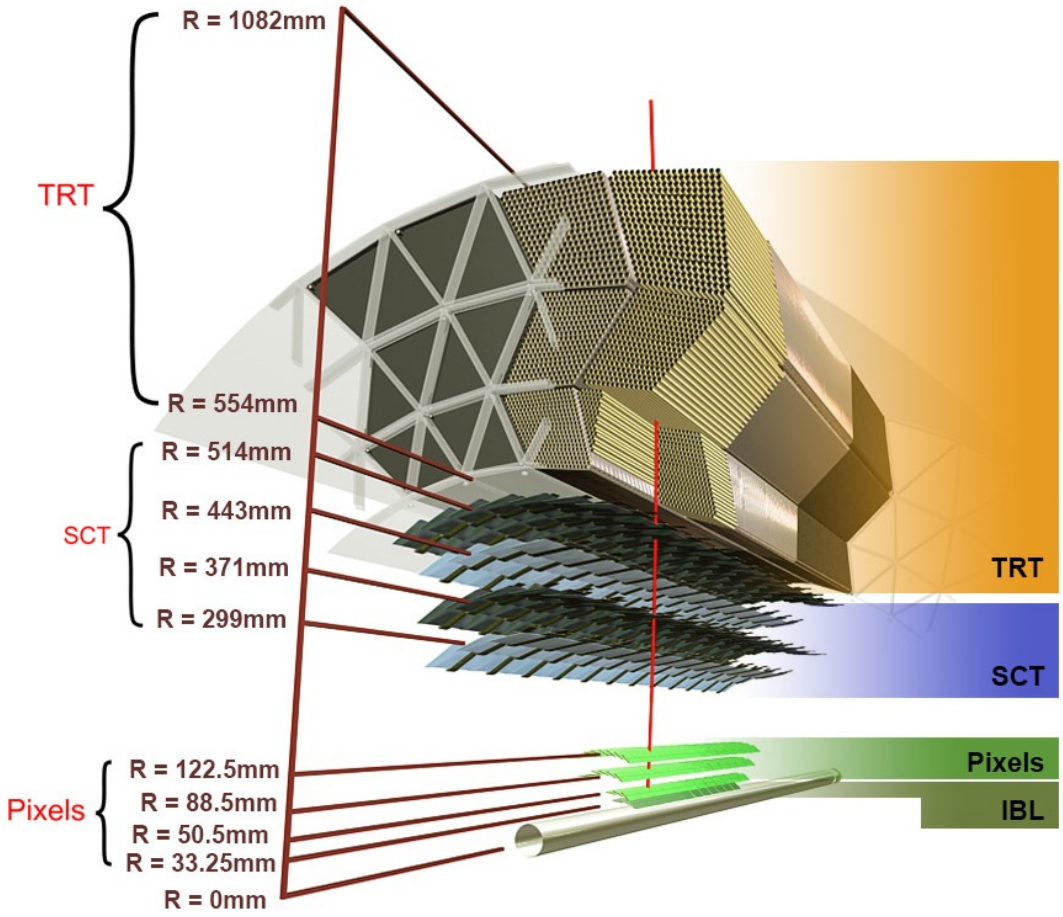


FIGURE 3.5: Cutaway of the ATLAS inner detector barrel components with the LHC beampipe shown at $R = 0$ mm. 'IBL' labels the inner-most pixel layer which was added between LHC Run 1 and Run 2. [41]

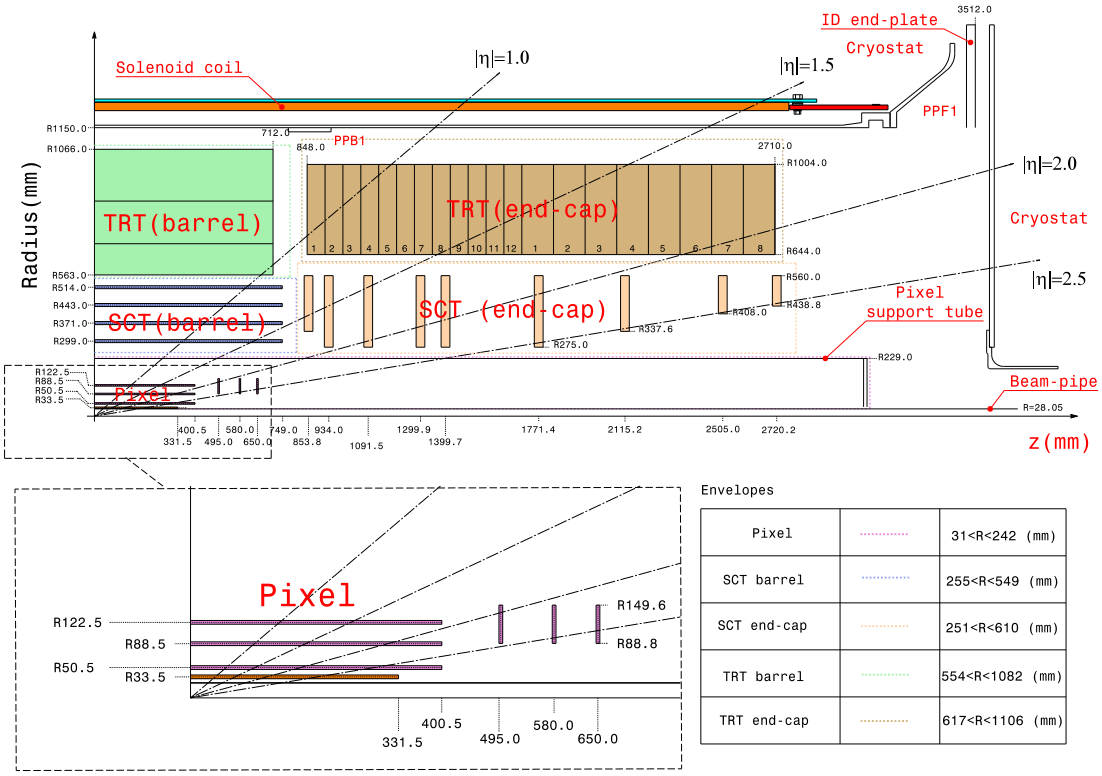


FIGURE 3.6: Schematic of the ATLAS inner detector showing barrel and endcap modules and indicating the coverage of different values of $|\eta|$ [42].

Pixel detector

The pixel detector is formed of a number of silicon pixel detector modules which provide 2D measurements along the plane of the module through the use of a grid of many individual pixels. Each pixel has an area of between 50 μm $\times$ 250 μm and 50 μm $\times$ 600 μm in the ϕ and z directions and operates by the collection of electrons which are liberated from the silicon atoms that form each pixel by the interactions of passing charged particles. These pixel modules are arranged into a number of layers arranged outwards from the interaction point. When measurements from multiple layers are combined a full 3D reconstruction of tracks is possible. The modules are arranged into four barrel layers and three endcap layers. The barrel modules are tilted in the ϕ plane and partially overlapped to ensure hermetic coverage of tracks in ϕ. The innermost barrel pixel layer is actually a newer detector layer named the insertable B layer (IBL) which was installed in 2014 prior to the start of Run 2 [43]. The IBL comprises newer pixel sensor techologies and has a mean radius of 33 mm from the centre of the LHC beampipe. The inclusion of this additional small-radius layer compared to the use of three barrel layers in Run 1 allows

for improved resolution in impact parameter[1], d_0, which improves the reconstruction of secondary vertices. This therefore improves the identification of particles containing b quarks due to their measurable decay length. The addition of detection 'hits' in another layer also improves the efficiency of finding tracks and reduces the combination of hits into fake tracks.

SCT

Detection with the SCT works using similar methods to those of the pixel detector, however the silicon based modules are instead arranged as strip sensors with relatively large extent in z. Each strip is 50 μm in the ϕ direction and 6.4 cm along z. Two strips are connected to form a single 12.8 cm long sensor. To allow more precise determination of the z of hits from charged particles, the SCT modules are formed of two modules containing a number of SCT strips where one module is stacked on top of the other with a small rotation of 40 mrad about the central axis through the plane of the modules. This rotation defines a stereo crossing angle between the strips of each module, which allows a more precise determination of the z extent a charged particle has passed through based on the distance between crossing points of different strips in the two modules. These SCT modules are arranged into four barrel layers and nine endcap layers. The barrel modules are mounted tilted in the ϕ plane similarly to what is done in the pixel detector barrel layers. The individual SCT modules have an intrinsic resolution of 17 μm in the r-ϕ plane and 580 μm in z (r) in the barrel (endcap) modules.

TRT

In contrast to the silicon based pixel and SCT detectors the TRT is formed of many gas-filled straw tube detectors. Each straw tube is a 4 mm diameter hollow cylinder with length 144 cm (37 cm) in the barrel (endcap) modules. They contain a central tungsten wire and is filled with a gas mixture dominantly formed from Xe gas. Some straws are instead filled with an Ar-based gas mixture due to significant gas leaks that developed during operation of ATLAS and the lower cost of the Ar gas. This mixture cannot absorb transition radiation efficiently and therefore cannot be used for particle identification,

[1]The impact parameter is defined as the smallest transverse distance between a reconstructed object and the beamline.

but has similar performance for track reconstruction [44]. Charged particles ionise the gas as they pass through and the resulting electrons are collected by the central wire. The TRT is additionally composed of 19 μm diameter polypropylene fibres and 15 μm thick polypropylene foils interleaved between the straw tubes in the barrel and endcap modules respectively. Charged particles passing through the polypropylene undergo transition radiation which can then be detected by the straw tubes due to ionisation of the gas from the transition radiation photons. The energy of the transition radiation is dependent upon the mass of the charged particle, allowing separation of tracks from electrons from tracks from other charged particles based upon the number of straw measurements which pass a high-threshold value. The straws are arranged into up to 73 layers in the barrel region and 160 planes of straws in the endcaps. This results in charged particles traversing an average of 36 straws except for the region $0.8 < |\eta| < 1.0$ where the average falls to 22 due to the transition between barrel and endcap modules. The TRT straws have an intrinsic accuracy of 130 μm in the r-ϕ plane.

Tracking Performance

Results showcasing the expected performance of track reconstruction using the ATLAS inner detector are shown in Figure 3.7 for simulated muon tracks at various momenta. Results are shown for detector configurations with and without the IBL included. The improvement on the impact parameter resolution of tracks when the IBL is used can be clearly seen. On the other hand, the relative uncertainty on the measurement of track curvature is almost identical in each configuration.

3.2.2 Electromagnetic Calorimetry

Outside of the inner detector and solenoidal magnet is the calorimeter dedicated to measuring the energy of charged particles, referred to as the electromagnetic (EM) calorimeter (sometimes abbreviated to ECal). The EM calorimeter is intended to fully measure the energies of electrons (and positrons), and photons, although measurements can be made of charged hadronic particles and muons as they pass through the EM calorimeter before they reach the associated dedicated sub-detector. The EM calorimeter makes these

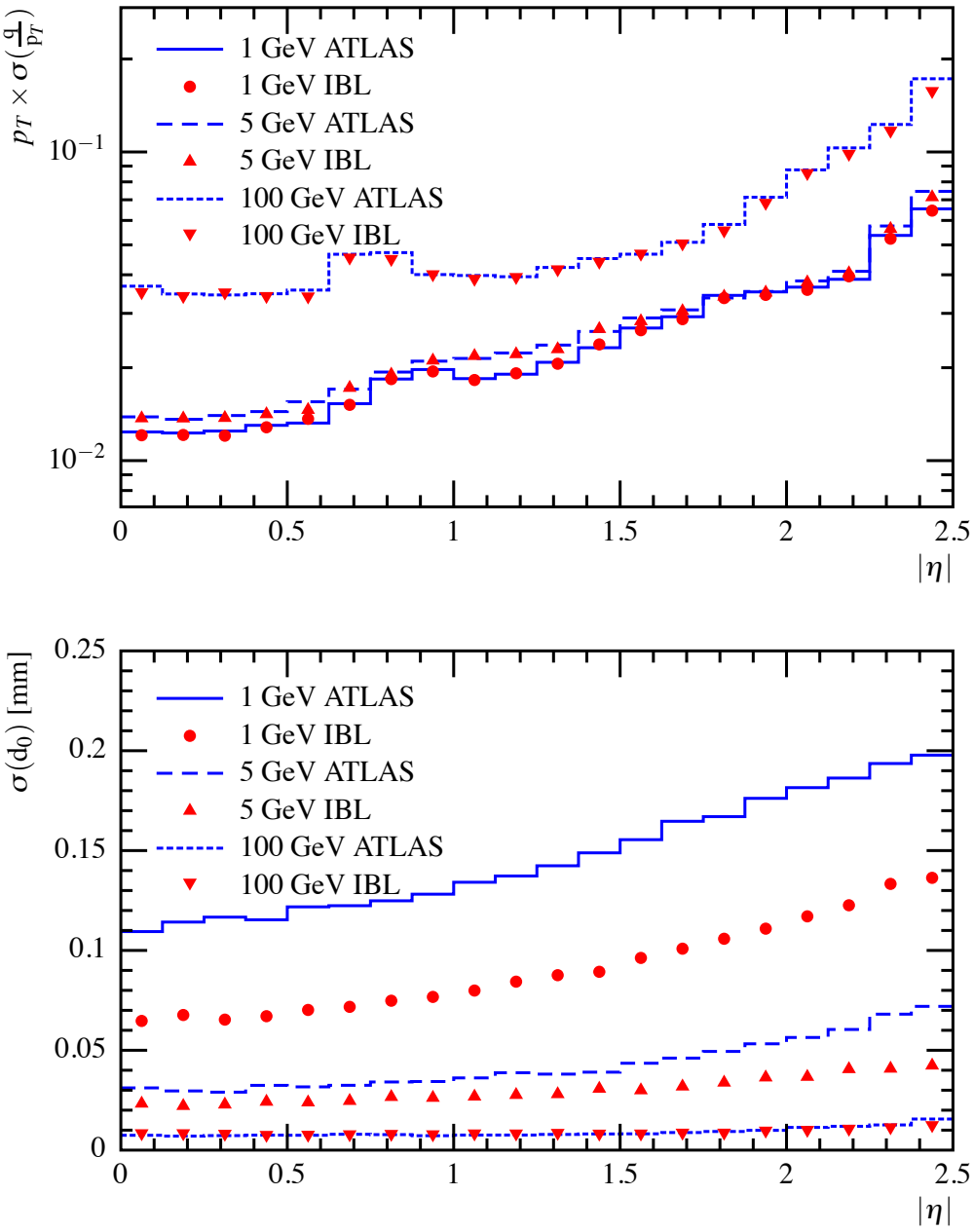


FIGURE 3.7: Expected performance of the ATLAS inner detector for track reconstruction. Results are shown for tracks from simulated 1 GeV, 5 GeV, and 100 GeV muons. Results are produced from two configurations of the ATLAS detector: without the inclusion of the IBL, labelled 'ATLAS', and with the inclusion of the IBL, labelled 'IBL'. The relative uncertainty on the curvature of a track as a function of absolute pseudorapidity is shown in the left figure, and the resolution of the transverse impact parameter of a track is shown as a function of absolute pseudorapidity in the right figure. The results were derived for simulated events containing no pile-up. [43]

energy measurements destructively by causing charged particles to lose energy in electromagnetic showers initiated by interactions with a dense material. An electromagnetic shower refers to the cascade of electrons, positrons, and photons produced by sequential electromagnetic interactions in material, which are initiated by a hard interaction. In the case of electrons this shower is primarily initiated by bremsstrahlung, while for photons it is primarily initiated by pair production. The resulting shower of particles can then be detected through interaction with detecting layers interleaved with the layers of dense material. The extent of the sub-detector can then be chosen based on the radiation length, X_0, of the detector material. Radiation length is defined as the mean distance taken for the energy of an electron to decrease to $\frac{1}{e}$ of its original value. Choosing a sub-detector size that is an appropriate multiple of the radiation length will ensure that electrons and photons completely lose their energy within the EM calorimeter and the resulting electromagnetic showers are fully contained within the EM calorimeter. The ATLAS EM calorimeter uses lead for the dense material which initiates electromagnetic showers and liquid argon (LAr) for the interleaved detecting layers. The electrons and positrons in the electromagnetic showers ionise the LAr and the resulting free electrons are cumulatively collected to determine the energy deposited by the original particle causing the electromagnetic shower. Barrel EM calorimeter modules are at least 22 X_0 thick, increasing to 33 X_0 as $|\eta|$ increases. Endcap EM calorimeter modules are at least 24 X_0 thick, increasing to 38 X_0 at forward $|\eta|$ values. A cutaway of the EM calorimeter can be seen in Figure 3.8. The EM calorimeter covers the pseudorapidity range $|\eta| < 3.2$. Additional electromagnetic reconstruction at $3.1 < |\eta| < 4.9$ is performed by the forward calorimeter (FCal), which is an additional set of LAr-copper and LAr-tungsten endcap modules.

The EM calorimeter modules are formed of interleaved lead and LAr layers arranged in an accordion shape i.e. a zig-zag pattern which is arranged with the pattern in the r-ϕ plane in the barrel modules and in the z plane in the endcap modules. This geometry ensures hermetic coverage in ϕ. As well the modules are split into multiple layers in r (z) in the barrel (endcap) modules and are segmented in η and ϕ. In the region used for precision physics defined by $|\eta| < 2.5$ the modules use three layers, except in the transition region between the barrel and endcap modules where two layers are used for $1.35 < |\eta| < 1.475$ in the barrel and $1.375 < |\eta| < 1.5$ in the endcaps. The first layer is

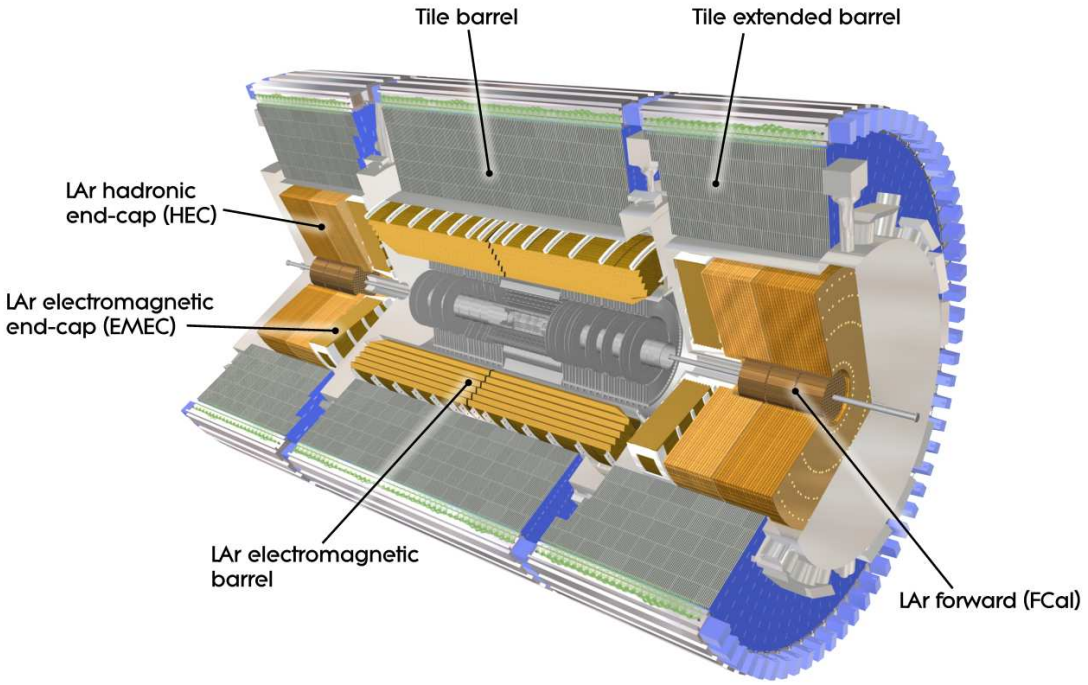


FIGURE 3.8: Cutaway of the ATLAS calorimeters showing barrel and endcap modules [40].

finely segmented in η, except for the barrel to endcap transition region, which allows precise measurement of the direction of charged particle energy deposits. η measurements for photons are made using both the first and second layers. A sketch of a barrel module is shown in Figure 3.9 where the fine segmentation of the first layer in η can be seen. The energy determinations made by the EM calorimeter modules are complemented by the use of a pre-sampling layer in the region $|\eta| < 1.8$. This is an additional thin LAr layer placed before the EM calorimeter modules which allows determination of energy lost from a particle due to interactions with material prior to the EM calorimeter.

EM Calorimetry Performance

Results showing the energy resolution of the ATLAS electromagnetic calorimeter are shown in Figure 3.10 and Figure 3.11 for the measurement of electrons. Results are shown for simulated electrons in Figure 3.10 and for electrons measured in real $\sqrt{s} = 8$ TeV data in Figure 3.11.

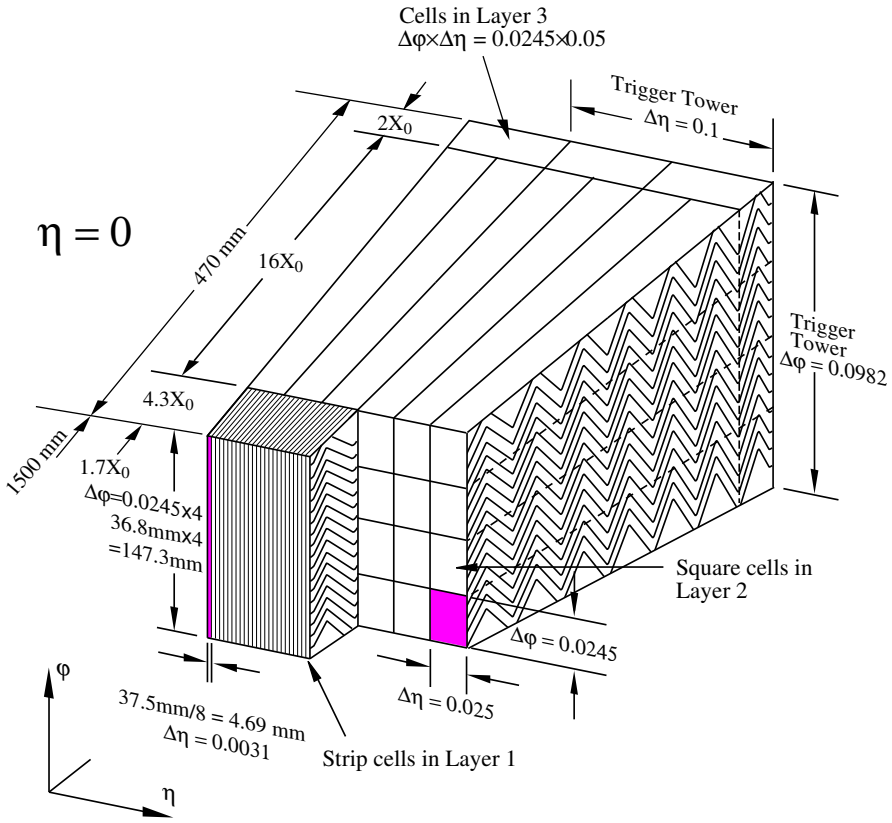


FIGURE 3.9: Sketch of a barrel module of the ATLAS EM calorimeter showing the segmentation in η and ϕ, layering in r, and accordion geometry of the interleaved lead and LAr layers [40].

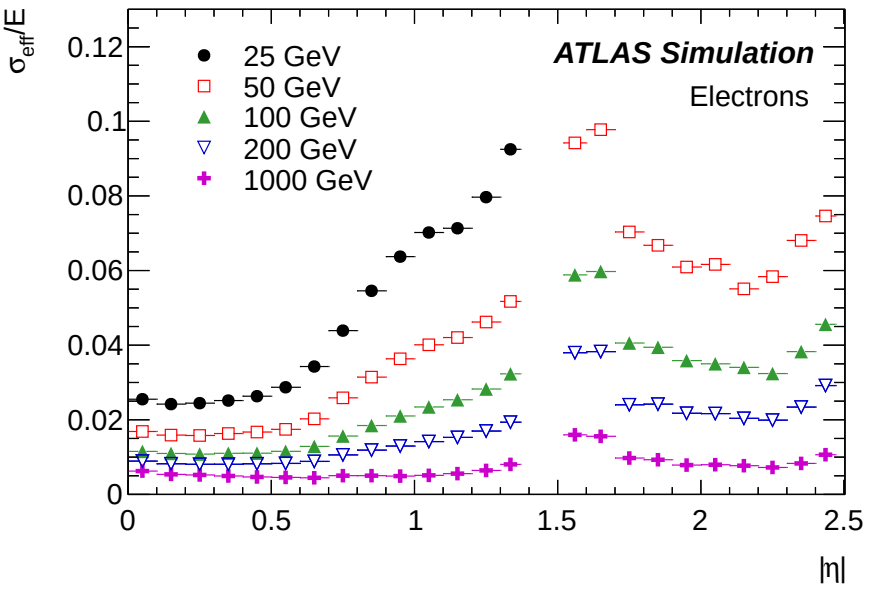


FIGURE 3.10: Relative energy resolution of the ATLAS electromagnetic calorimeter as a function of absolute pseudorapidity for simulated electrons. Results are shown for a range of electron energies. Results for 25 GeV electrons are only shown for $|\eta| < 1.37$, where they have $E_T >$ 10 GeV. [45]

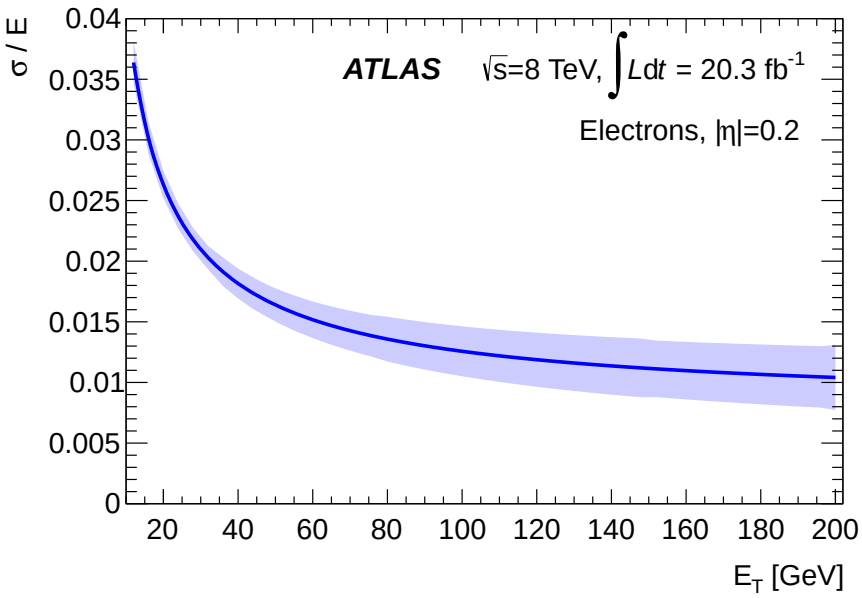


FIGURE 3.11: Relative energy resolution of the ATLAS electromagnetic calorimeter as a function of transverse energy for electrons measured in real $\sqrt{s} = 8$ TeV data, where the electrons have $|\eta| = 0.2$. The blue band indicates the uncertainty on the energy resolution. [45]

3.2.3 Hadronic Calorimetry

Beyond the EM calorimeter is a second calorimeter sub-detector which is instead dedicated to the measurement of the energy of hadronic particles, referred to as the hadronic calorimeter (sometimes abbreviated to HCal). Similarly to the EM calorimeter, measurements are made by destructively interacting hadronic particles with a dense material to initiate hadronic showers. A hadronic shower is produced from inelastic nuclear interactions which produce secondary hadrons and excited nuclei; these interaction products can then decay into further particles and undergo further inelastic interactions, resulting in a shower of hadronic particles. An EM component can also be produced due to the decays of neutral pions to photons. The products of the hadronic showers can then be detected through interactions with detecting layers interleaved with the layers of dense material. The extent of the sub-detector is chosen based on the interaction length, λ, of the detector material. Interaction length is defined as the mean distance travelled by a hadronic particle before experiencing an inelastic nuclear reaction. Choosing detector sizes that are a appropriate multiples of the interaction lengths of detector material will ensure that hadrons lose all of their energy and the resulting hadronic showers are fully contained in the hadronic calorimeter. Unlike the EM calorimeter the hadronic calorimeter uses two different combinations of dense and detecting materials and methods. The

barrel modules plus two extended barrel modules use steel as the dense material and scintillating tiles as the detecting material. These scintillating tiles absorb energy from particles passing through the material. This energy is then released as photons that are collected by optic fibres and cumulatively read out by photomultiplier tubes. The scintillating-tile-based barrel and extended barrel modules are collectively referred to as the tile calorimeter. On the other hand the endcap hadronic calorimeter modules use copper as the dense material and LAr as the detecting material. The principle of detection is therefore similar to the EM calorimeter. LAr-based detection is not used in the barrel hadronic calorimeter modules due to the cheaper cost of the steel plus scintillating tile technology. This cheaper cost allows for the largest radial coverage with the lowest cost. The combination of EM calorimeter and hadronic calorimeter material in the barrel region is approximately 9.7 λ thick (composed of approximately 7.4 λ of tile calorimeter), and approximately 10 λ thick in the end caps. The barrel tile calorimeter modules cover the region $|\eta| < 1.0$, the extended barrel tile calorimeter modules cover the region $0.8 < |\eta| < 1.7$, and the endcap LAr modules cover the region $1.5 < |\eta| < 3.2$. The FCal is also used for hadronic calorimetry and extends the coverage up to $|\eta| < 4.9$. A cutaway of the hadronic calorimeter can be seen in Figure 3.8.

The tile calorimeter modules are formed of staggered layers of scintillating tiles inserted into steel segments. The fibres from individual tiles are grouped and groups are read out by the same photomultiplier tube. The groups are defined such that the tile calorimeter readout is divided into three radial sections and a number of sections approximately segmented in η. The structure of a tile calorimeter module is shown in Figure 3.12.

Hadronic Calorimetry Performance

Results showing the transverse momentum resolution of ATLAS hadronic calorimetry are shown in Figure 3.13. Results are shown for real $\sqrt{s} = 7$ TeV data and for simulation, and are derived using two independent methods for measuring jet momentum resolution. Descriptions of the methods used can be found in Reference [46].

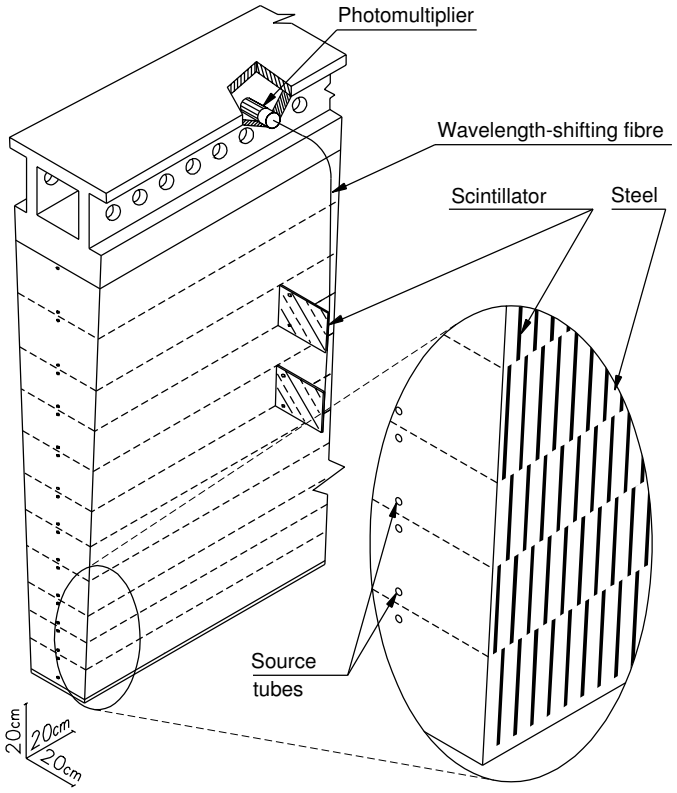


FIGURE 3.12: Schematic of a tile calorimeter module of the ATLAS hadronic calorimeter [40].

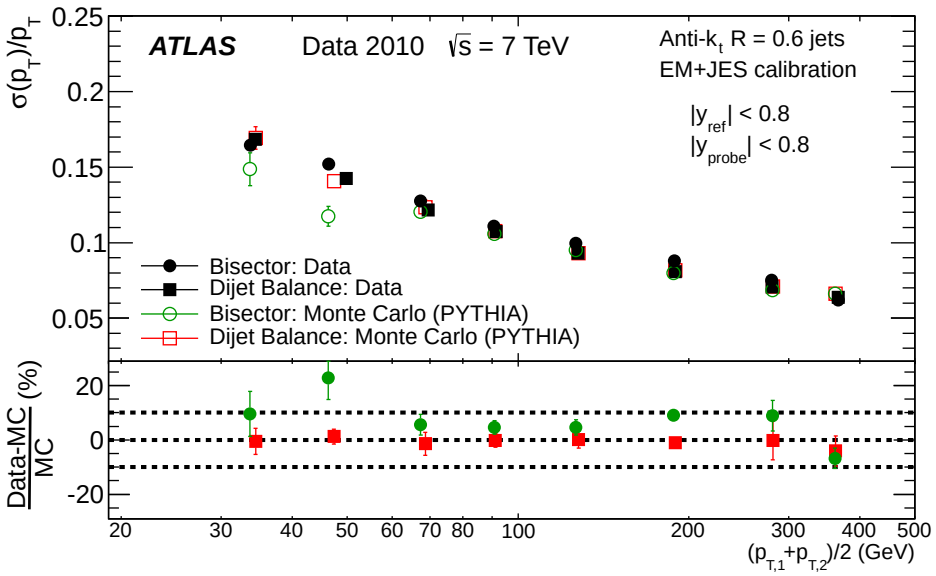


FIGURE 3.13: Relative resolution of the transverse momentum of jets measured in simulation and in real $\sqrt{s} = 7$ TeV data, where the jets are required to be central in rapidity. Results are shown for two independent methods of measuring transverse momentum resolution. Results are shown as a function of the average momentum of the two jets used in the measurement methods. [46]

3.2.4 Muon Spectrometry

The final sub-detector is dedicated to the reconstruction of the paths of muons and reconstruction of their momentum, and is referred to as the muon spectrometer (sometimes abbreviated to MS). The positioning of the muon spectrometer as the outer layers of ATLAS reflects the minimum ionising particle (MIP) nature of muons. Muons will pass through calorimeter layers with little energy loss while other particles will mostly be fully contained before the outer radius of the hadronic calorimeter (except for weakly interacting particles such as neutrinos). The muon spectrometer is composed of modules which use a number of different detection methods. These modules are organised into a number of different barrel and endcap arrangements. Four different types of modules are used: monitored drift tubes (MDT), cathode strip chambers (CSC), resistive plate chambers (RPC), and thin gap chambers (TGC). MDT and CSC modules are responsible for precision spatial measurements of the path of muons passing through the muon spectrometer, while RPC and TGC modules are responsible for the triggering of collision events of interest through rapid detection of the presence of high energy muons. TGCs also provide a secondary measurement of ϕ to complement the measurements from the precision modules. The muon spectrometer makes use of toroidal magnets to bend the paths of muons to allow for measurements of their momentum. 0.5 T toroid magnets with magnetic fields aligned in the azimuthal direction are arranged in a barrel formation which covers the region $|\eta| < 1.4$. Two smaller endcap modules contain a number of 1 T toroid magnets which are arranged within the endcaps in similar fashion to the barrel magnets and cover the region $1.6 < |\eta| < 2.7$. The region $1.4 < |\eta| < 1.6$ is covered by a combination of the barrel and endcap magnetic fields. A cutaway of the muon spectrometer and toroid magnets is shown in Figure 3.14.

The MDTs function similarly to the straw tubes of the TRT in the ATLAS inner detector: drift tubes are formed of 30 mm tubes with a central tungsten wire, which are filled dominantly with Ar gas. The gas is ionised by the passage of muons and the resulting electrons are collected and read out by the central wire. The time taken for the electrons to be collected by the wire is used to determine the position of the detected muon. MDT modules are formed of one or two groupings of three or four stacked layers of a number

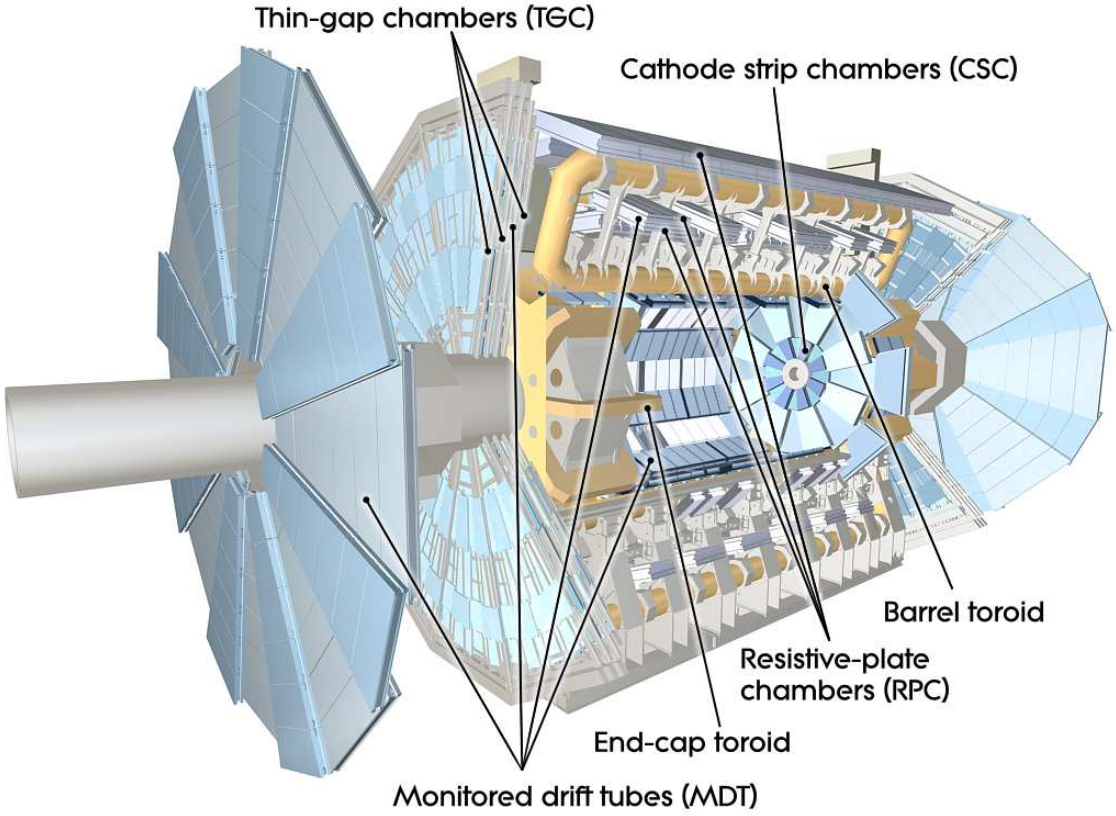


FIGURE 3.14: Cutaway of the ATLAS muon system and toroidal magnets [40].

of drift tubes. These modules are arranged into three barrel layers and three endcap layers. The innermost endcap layer covers the region $|\eta| < 2.0$ while the outer two endcap layers cover $|\eta| < 2.7$.

The CSCs are formed of many tungsten wires arranged between two planes of copper strips; this arrangement creates a uniform electric field near to the copper strips. The space between the two planes is filled with a gas mixture dominantly formed from Ar gas. Muons ionise this gas and the positive ions can be collected and read out by the copper strips (the electrons are collected by the wires and are not used for measurements). The strips in one plane are arranged perpendicularly to the strips in the other plane, which allows spatial measurements along two axes. CSC modules are formed of four stacked CSC planes. These modules are used in a single endcap layer in the region $2.0 < |\eta| < 2.7$. CSC modules are used in this region instead of MDT modules due to higher rate conditions which exceed the safe operation levels of MDT modules.

The RPCs are formed of two parallel plates made of a highly resistive plastic laminate with a primarily $C_2H_2F_4$ gas mixture in between the plates and metallic strips on

the outer faces of the plates. A strong electric field is placed between the resistive plates which results in an avalanche of electrons being emitted in the gas when a muon passes through. The charge avalanche rapidly creates a capacitive electric field felt by the metallic strips on the outer faces of the resistive plates which results in a measurable signal. The metallic strips on one plate are oriented perpendicularly to those on the other plate, which allows spatial measurements along two axes. RPC modules are formed of two partially overlapping stacked RPC planes which have an overlap of 65 mm. Three barrel layers of RPC modules are used, which are arranged around the MDT barrel layers: two RPC layers are on the inside and outside faces of the second MDT barrel layer and the third RPC layer is either on the inside or outside face of the third MDT barrel layer depending on the radial position of the MDT layer. These layers cover the region $|\eta| < 1.05$.

The TGCs are formed of similar modules as the CSCs i.e. a number of wires between two charged planes. In the case of TGC modules the distance between the wires and the strips is smaller than the distance between the wires. The result of this is that the uniform region of the electric field is smaller, which results in electrons from gas ionisation drifting to the wires more quickly. The wires of the TGCs are read out as well as metallic strips outside of the charged planes. TGC modules are formed of either a doublet or triplet of stacked TGC planes. The TGC modules are arranged into four endcap layers: one layer on the inside face of the first MDT endcap layer, one layer on the inside face of the middle MDT endcap layer, and two layers on the outside face of the middle MDT endcap layer. These layers cover the region $1.05 < |\eta| < 2.4$. TGCs are used in this region due to higher radiation levels.

Muon Spectrometry Performance

Results showing the invariant mass resolution of pairs of muons measured with the ATLAS muon spectrometer are shown in Figure 3.15. Results are shown for real $\sqrt{s} =$ 13 TeV data and for simulation, and are measured for muon pairs identified as originating from Z bosons and J/ψ mesons. The observed invariant mass resolution was measured to correspond to a relative muon transverse momentum of 1.7% and 2.3% at central pseudorapidity, and 2.3% and 2.9% for measurements made with endcaps, for $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ decays respectively [47].

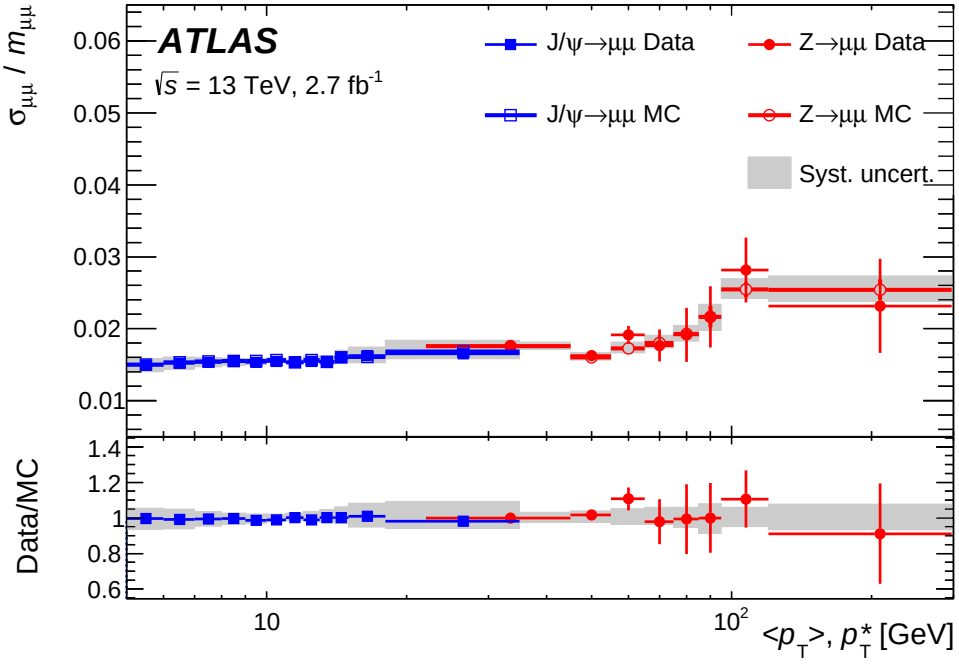


FIGURE 3.15: Relative resolution for the invariant mass of pairs of muons measured in simulation and in real $\sqrt{s} = 13$ TeV data. Results are shown for muon pairs identified as originating from Z bosons and J/ψ mesons. Results are shown as a function of the average momentum of the two muons used in the measurements, where $<p_{\mathrm{T}}>$ is the average momentum defined for $J/\psi \to \mu\mu$ events and p_{T}^{*} is the average momentum defined for $Z \to \mu\mu$ events. [47]

3.2.5 Triggering and Data Acquisition

To manage the high rate of LHC collisions and reduce the amount of data which is processed and which is saved to disk, ATLAS employs a trigger system, referred to as the ATLAS trigger, to partially reconstruct events during data taking and make decisions about which events should be recorded. The ATLAS trigger therefore forms a part of the hardware and software architecture which is used for data acquisition (DAQ). Processing which is done during data taking is referred to as being performed 'online', while processing done after an event has been recorded is referred to as being performed 'offline'. Objects which are reconstructed in a given stage are often referred to as an 'online object' or 'offline object' as appropriate. For example an 'online track' is a track reconstructed by track finding algorithms within the ATLAS trigger while an 'offline track' is a track reconstructed by offline track finding algorithms.

For an LHC beam operating with the design bunch spacing of 25 ns ATLAS will be subject to a collision event rate of 40 MHz. Many of these collisions will not produce events of interest and can therefore be discarded; this therefore reduces the rate of events

to be processed and recorded. The ATLAS trigger must be able to rapidly make these determinations under the stringent requirement of the 25 ns time between bunch crossings. To do this ATLAS employs a level 1 (L1) trigger system formed of a mixture of dedicated detector components and processing hardware to rapidly measure basic event properties. However due to the high instantaneous luminosity of approximately 10^{34} cm^{-2} s^{-1} the selection criteria for the trigger decisions must be further tuned to ensure that the final output event rate is below the tolerable thresholds of the data readout rate and offline storage space and offline processing requirements. The maximum L1 trigger output rate is approximately 100 kHz, which is defined by the limitations of the detector readout capability. This rate must be reduced to a final output rate of approximately 1.5 kHz. To achieve this further reduction in rates, more detailed event reconstruction must be performed than what can be achieved in the L1 trigger. To do this ATLAS employs a high level trigger (HLT) system formed of a dedicated computer farm that performs offline-like software-based reconstruction of events which pass the L1 trigger. More detailed criteria based on this more complete event reconstruction can then be used to make final trigger decisions and reduce the event rate to tolerable levels. A functional diagram showing the path of information through the DAQ and trigger systems from the inputs from detector modules to the readout to storage is shown in Figure 3.16. The rate of events and rate of data at each stage is also shown. FE refers to the front end electronics responsible for the read out from each detector system. ROD refers to the read out driver electronics responsible for sending detector measurements to the read out service. FTK refers to a fast tracker hardware module that was in the process of being commissioned during 2015 and 2016 data taking but was not used in triggers for data taking.

The L1 trigger performs coarse event reconstruction using the dedicated trigger modules in the muon spectrometer (RPCs and TGCs) and reduced-granularity information from the EM calorimeter and hadronic calorimeter. In this way the L1 trigger can make rapid reconstructions of muons, electrons and photons[2], hadronic jets, and hadronically decaying tau leptons. Missing transverse energy can also be measured based on the sum of the transverse energies of reconstructed objects. Requirements can also be made based on the 'isolation' of how spatially close a given measurement is from other measurements

[2]No distinction can be made between the EM calorimeter deposits of electrons and photons in the L1 trigger due to the lack of charged particle track measurements.

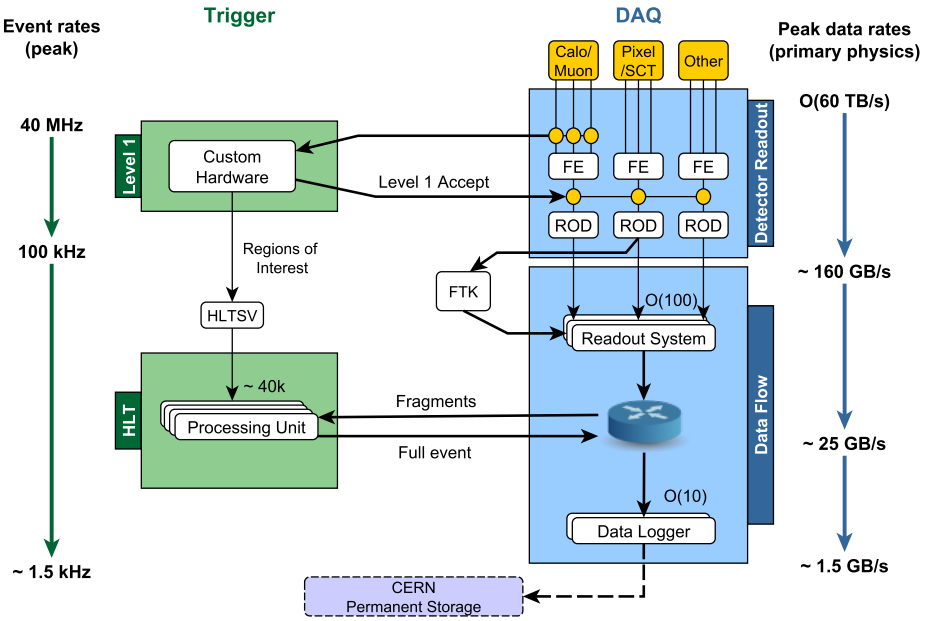


FIGURE 3.16: Functional diagram which shows an overview of the ATLAS trigger and data acquisition systems as well as the reduction in event and data rates over the course of trigger and data acquisition processing. The top of the diagram represents detector measurements made for each LHC bunch crossing and the bottom of the diagram represents the final readout of information to permanent storage. [48]

made in the L1 trigger. The time taken for an L1 trigger decision is less than 2.5 μs, with a target time of 2.0 μs [49].

The HLT trigger performs full granularity reconstruction from the muon spectrometer and calorimeters and also performs track reconstruction of measurements made by the inner detector. The portion of the HLT which is dedicated to track reconstruction is referred to as the inner detector trigger. Reconstructed tracks can then be matched to measurements in other sub-detectors to provide full reconstruction of physics objects. The HLT can therefore separate electrons from photons and tag jets which contain b quarks. It is worth noting that the inner detector trigger does not make any decisions to keep or reject an event; these decisions are made based on the fully reconstructed physics objects which make use of the tracks. The HLT can also make decisions based on the quality of reconstructed objects, as well as spatial quantities such as $\Delta\eta$ between two objects. More detailed object isolation requirements can also be consdered in the HLT decision. For the full granularity reconstruction performed by the HLT to be possible within the time constraints of online running the reconstruction is performed within reduced spatial

volumes referred to as regions of interest (RoI). The η and ϕ directions of these RoIs are given by the measurements made in the L1 trigger and the spatial extent is defined appropriately to the object being reconstructed. The time taken for an HLT trigger decision is approximately 300 ms [49].

The running of the ATLAS trigger is based on menus of many individual 'trigger chains', often referred to simply as triggers. Each trigger defines a selection on one or more objects reconstructed by the L1 trigger followed by further selections on these or additional objects in the HLT trigger. For example a single electron trigger would require an EM calorimeter deposit reconstructed in the L1 trigger to pass a specified energy threshold. Events passing this L1 trigger decision would then undergo HLT event reconstruction such that the electron is fully reconstructed. A tighter p_{T} threshold, as well as identification criteria, could then be placed on the reconstructed electron to finally accept the event in question. An event which passes both the L1 and HLT trigger decisions is described as having passed the trigger in question. Depending on the trigger selections an event can pass multiple individual triggers in a trigger menu.

Chapter 4

Event Simulation and Object Reconstruction

4.1 Event Simulation

To analyse the particle collision events recorded by ATLAS and to test the theoretical understanding of high energy physics processes, it is usually necessary to utilise simulated collision events which can be analysed in the same manner as real data. Producing simulated events requires theoretical modelling of a number of aspects of p-p collisions and their products. Measurements of the simulated events using ATLAS must also be considered to produce reconstructed data that can be compared to reconstructed data arising from real collision events. The interaction of the produced simulated particles with detector components must therefore also be considered in a dedicated simulation of the ATLAS detector. Both collision event generation and detector simulation use Monte Carlo (MC) simulation methods. These methods use random generation of variables which are distributed according to functions of interest, for example the matrix element of a collision process or the modelling of the interaction of particles with detector material, to provide many samples which are distributed with appropriate probability density.

To improve agreement between predictions from simulated data and results from real data, event weights are applied to simulated events. These weights cover a number of sources of corrections, that are broadly split into two types:

- The first type is associated with the number of simulated events generated with various amounts of pile-up. The proportion of events with different values of μ may

be found to be incorrect when compared against μ measured in real data. Event weights can therefore be applied to reproduce the correct proportions of pile-up;

- The second type is associated with mismodelling of the reconstruction and identification of objects. For example, simulation may predict an inaccurate rate for electrons to pass an isolation requirement compared to what is measured in real data. Event weights can therefore be applied to to reproduce the correct selection efficiency for such a require

Weights are also calculated as part of Monte Carlo event generation and are applied to the resulting simulated datasets.

4.1.1 Event Generation

The generation of simulated collision events is split into two different types of calculation: the simulation of the matrix element process and the simulation of the resulting parton shower and hadronisation processes. The matrix element process is the hard scatter interaction of partons from a proton in each colliding beam. Performing these calculations requires knowledge of the distributions of partons in each proton to correctly predict process cross-sections and kinematics. The description of these parton distributions are encoded in parton distribution functions (PDFs) which give the probability for a parton of a given type to be found with some fraction of the proton's total momentum[1]. With the PDFs known, the matrix element calculations of the parton-parton interactions are performed at fixed-order in perturbation theory. As a result the production of additional quarks or gluons from beyond the order of the calculation must be handled separately. This is achieved by using the parton shower simulation.

Parton shower processes are the higher order emissions of additional gluons or the splittings of gluons into quarks from initial or final state particles from the matrix element process. The production of these additional partons is performed by scaling the probabilities for a given type of parton to split into additional partons as a function of energy[2]. The parton shower simulation calculates splittings until the energy scale of the partons has dropped below approximately 1 GeV. At this point hadronisation of free coloured partons

[1] A more detailed description of PDFs is given in the PDF uncertainty studies in Section 6.6.3.

[2] This scaling is done using the DGLAP equations that describe the same parton splitting processes happening within the proton, as mentioned in Section 6.6.3.

into colourless hadrons due to the self-interacting nature of QCD begins to occur. Because the energy scales are low, QCD calculations become non-perturbative and approximations must be used for how additional particles and final colourless states are produced. The two models used are the Lund string model [50] and the cluster model [51]. These calculations are also performed within parton shower simulation software packages. For brevity "parton shower and hadronisation" is often shortened to "parton shower" in this thesis when describing the simulation packages used. The final step of the event generation is to decay any particles which do not live long enough to interact with the ATLAS detector. After this the list of stable final-state particles can be defined to be put through the detector simulation. The record of the particles and interactions produced by the matrix element and parton shower simulations is referred to as the 'truth' record.

Individual datasets of simulated events are produced using combinations of matrix element simulators and parton shower simulators provided by a number of different simulation groups. The matching of a matrix element simulation package to a parton shower simulation package is itself a technical challenge, and which may be handled differently between different combinations of packages. As well the simulation settings and parametrisations provided by the groups (for example the maximum energy scale for the emission of additional radiation) may be found to give poor agreement when compared to real data events in physics analyses. Therefore an additional step in the simulation of events may be employing sets of 'tunes' of simulation parameters which give better real-world agreement. These tunes are produced centrally in the ATLAS Collaboration, with input from the authors of the simulation packages.

Additional simulation is used to model the interaction of the proton remnants after the hard scatter has taken place, as well as the soft scatter interactions of other protons in the colliding bunches i.e. pile-up interactions. These non-hard-scatter processes are collectively referred to as the underlying event (UE) and are overlaid on the produced hard scatter simulated event to produce the full description of the colliding beams.

4.1.2 Detector Simulation

The simulation of the interaction with the ATLAS detector of particles produced by collision events is performed using the GEANT4 simulation package [52]. This package geometrically models the material of the ATLAS detector to the level of detail of wiring in the apparatus and models the interaction of particles passing through this matter (including bending by the magnetic fields of the ATLAS magnets), the response of the detector to the particle interactions, and the decays of particles within the detector volume. The modelling of the detector is regularly tuned according to alignment measurements to ensure the detector description matches the real-world condition of ATLAS. The output of the detector simulation is equivalent to measurements made by ATLAS of real data events, which allows the measurements of both real events and simulated events to be digitised and reconstructed in exactly the same manner.

The full simulation of interactions with the detector is computationally costly which can be a limiting factor when a large number of simulated events are required for physics analyses. To reduce the computational cost while retaining a good description of interactions with the detector, a fast simulation labelled as ATLFAST II (AFII) is used instead [53]. This simulation performs the same modelling in the inner detector and muon spectrometers as the full simulation, however the response of the calorimeters is modelled using parametrised showers. This parametrisation is done with the FASTCALOSIM simulation [54] which reproduces longitudinal shower properties but gives only average lateral shower properties. The AFII fast simulation is mainly used for additional simulated datasets which are used to study systematic effects.

4.2 Object Reconstruction

Measurements made in the ATLAS detector are combined and calibrated into identified physics objects using a number of different methods (and utilising measurements in different parts of the ATLAS detector) appropriate to each object type. In general only the reconstruction methods and types of reconstructed objects that are relevant to work described in this thesis are included in this section.

4.2.1 Track and Vertex Reconstruction

Tracks are used in the reconstruction of all objects relevant to the work presented in this thesis. Vertex reconstruction is essential for the removal of objects originating from pile-up interactions and for measurement of the properties of longer lived particles (such as hadrons containing b quarks) through the reconstruction of secondary vertices. For most objects tracks are only reconstructed in the inner detector, however for muons separate tracks are reconstructed in the inner detector and in the muon spectrometer. The muon spectrometer tracks are discussed in Section 4.2.3.

Tracks are reconstructed by forming track seeds from a limited number of measurements in the pixel and SCT layers and extrapolating them into track candidates using additional measurements in the pixel and SCT layers; candidates which pass ambiguity solving and quality criteria are extrapolated into measurements made in the TRT before performing track fitting to produce final tracks [55, 56]. The first step in this process is clustering together measurements made by multiple elements of pixel or SCT detector modules to form 3D space-points. Combinations of three space-points are then chosen as track seeds. The direction of the seeds then defines spatial windows in which to attempt combinations with space-points in additional detector layers. A track candidate is then built by applying a Kalman filter [57] to the candidate measurements which works by calculating predicted track parameters along the direction of the combined measurements and updating these parameters by iterating through the detector layers. This algorithm also takes into account scattering of tracks by layers of material in the detector. The resulting track candidates are then scored based on criteria such as χ^2 of the track candidate and the number of missing hits in active detector layers, referred to as 'holes'. Ambiguities of space-points which are used in multiple track candidates are then resolved with a bias toward keeping space-points in higher scoring tracks. Tracks are re-scored and quality criteria such as a minimum number of measurement clusters are applied to reject poor quality tracks. The resulting track candidates are then extrapolated into the TRT to provide many additional hits. The tracks are then refitted with a higher resolution fit to provide final reconstructed tracks.

Tracks are used to reconstruct collision vertices by iteratively refitting candidate vertices until tracks are rejected from a given vertex and are instead used to repeat the procedure with other vertices [58, 59]. This procedure works by first defining a seed position for a first vertex based on the beam spot[3] position in the transverse plane and the mode of the z values of the reconstructed tracks at closest approach to the beam spot. The iterative fit is then performed by assigning weights to each track based on χ^2 agreement with the estimated vertex position and then recomputing the vertex position with these track weights included. This process is repeated for a number of iterations and tracks which are incompatible with a vertex position at a level of seven standard deviations are not assigned to the vertex. The fitting process is then repeated for these leftover tracks until all tracks are associated with a vertex or no additional vertices can be defined. Reconstructed vertices with at least two associated tracks are considered for the primary vertex.

4.2.2 Electrons

The measurement signature of electrons is a single track pointing to a large energy deposit in the EM calorimeter. Reconstruction therefore entails clustering energy measurements in the EM calorimeter into identified electron deposits which are then matched to inner detector tracks to form the reconstructed electron [60, 61]. An important consideration in this reconstruction is energy losses of electrons passing through the inner detector due to bremsstrahlung. These energy losses affect the paths of the electrons through the tracker, meaning that they must be taken into account when reconstructing the particle tracks. Criteria on the identification of how electron-like a reconstructed electron is are defined to reduce reconstruction of other particles as electrons. Separate criteria are also defined on how isolated a reconstructed electron is from other objects to reduce the selection of electrons which are produced from the decays of other particles which originated from the hard collision interaction.

The EM calorimeter clusters are reconstructed by scanning a sliding window across groups of calorimeter modules to determine if any groups are of sufficient energy to have arisen from an electron. The window size is 0.075×0.125 in $\eta \times \phi$ space, which corresponds to 3×5 of the granularity of the second EM calorimeter layer (0.025×0.025).

[3]The central point of the overlapping LHC beams.

Calorimeter 'towers' within this window, consisting of groupings of calorimeter layers in $\Delta\eta^{\text{tower}} \times \Delta\phi^{\text{tower}} = 0.025 \times 0.025$ regions (where the $\eta \times \phi$ space of the calorimeter is formed from a grid of such regions), are then selected to seed a full electron calorimeter cluster if they have $E_{\text{T}} > 2.5$ GeV. The full clusters are then formed around the seeds with a clustering algorithm which allows for removal of duplicates from nearby seeds. A larger window is finally used to determine the kinematic properties of a cluster, with the size of the window being determined by the position of the cluster in the calorimeter.

Tracks are reconstructed using either the standard track reconstruction described in Section 4.2.1 or a modified electron-specific track reconstruction designed to accommodate energy losses due to bremsstrahlung. The standard track reconstruction accounts for energy losses due to interactions with the detector under the assumption that tracks are from charged pions, referred to as the pion hypothesis. If a track seed with $p_{\text{T}} > 1$ GeV cannot be extrapolated into a full track with at least seven hits and the seed is within $\Delta R < 0.3$ of the barycentre of an EM cluster (defined as the mean η and ϕ of the energy in the cluster, calculated from the energy-weighted sum of the positions of cluster constituents), the electron-specific track finding is tried. In this method the standard track finding is modified to instead use an electron hypothesis which allows up to 30% energy loss at intersections of the track with detector elements. The tracks found from both methods are then fitted using a χ^2 track fit which uses the same hypothesis that was used in the initial track finding. The electron hypothesis is employed in the fit by including an extra term that compensates for increases in χ^2 due to bremsstrahlung. If a track which was found under the pion hypothesis does not survive the pion fit the electron fit is tried instead. The tracks are then spatially matched to the EM calorimeter clusters by comparing the extrapolated position of a track within the second EM calorimeter layer to the position of the cluster. Consideration is given to the best spatially matching track if a number of tracks match to one cluster. Tracks are then refitted using a Gaussian sum filter [62] algorithm, which is a non-linear generalisation of the Kalman filter method, to provide precise measurements of the track properties.

Identification criteria on the reconstructed electrons are defined by combining a number of variables which have some discriminantion of whether a reconstructed electron is signal-like (i.e. originated from an electron) or background-like (i.e. originated from another particle) using a multivariate analysis (MVA) method. The method used is the

likelihood-based (LH) method which calculates the probability for an object to be signal-like or background-like based on the combination of probability density functions from the discriminating variables used. A final discriminant variable $d_{\mathcal{L}}$ is defined as the relative probability for the object to be signal-like:

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B}, \qquad \mathcal{L}_{S(B)}(\vec{x}) = \prod_{i=1}^{n} P_{s(b),i}(x_i) \tag{4.1}$$

where $\vec{x}$ is the vector of the values of the discriminating variables, $P_{s,i}(x_i)$ is the evaluation of the signal probability density function for the i^{th} discriminating variable at the value x_i, and $P_{b,i}(x_i)$ is the equivalent quantity for the background probability density function. The variables used are broadly categorised as the shapes and relative energy depositions in different layers of the EM calorimeter, parameters of the associated track, distance between extrapolated track positions and cluster position, relative energy leakage into the hadronic calorimeter, and probability of the presence of transition radiation in the TRT. The particle identification properties of the TRT are utilised in the final variable as a likelihood discriminant variable which defines the probability of a TRT hit to give a high threshold measurement, which is indicative of transition radiation. Simple cuts on the multiplicity of hits in the particle track are also used.

The identification criteria are finally defined by applying a cut at three working points: loose, medium, and tight. The working points are not exclusive of each other, which means that reconstructed electrons identified as tight are also identified as medium and reconstructed electrons identified as medium are also identified as loose. The working points are optimised using simulated datasets as a function of $|\eta|$ and E_{T}. Efficiencies of correctly identifying electrons and mis-identifying hadrons as reconstructed electrons for the three working points are shown in Figure 4.1. The reconstruction efficiencies for both truth electrons and hadrons are seen to increase at 50 GeV because the last bin used for the optimisation of electron ID is the 45-50 GeV bin.

Isolation criteria are used to remove electrons which are not 'promptly' produced from the hard collision interaction and are instead 'non-promptly' produced from the decays of hadronic particles which were produced in the hard interaction. The isolation criteria also further remove hadrons which are mis-identified as reconstructed electrons. The criteria are placed on variables which measure the energy in spatial regions

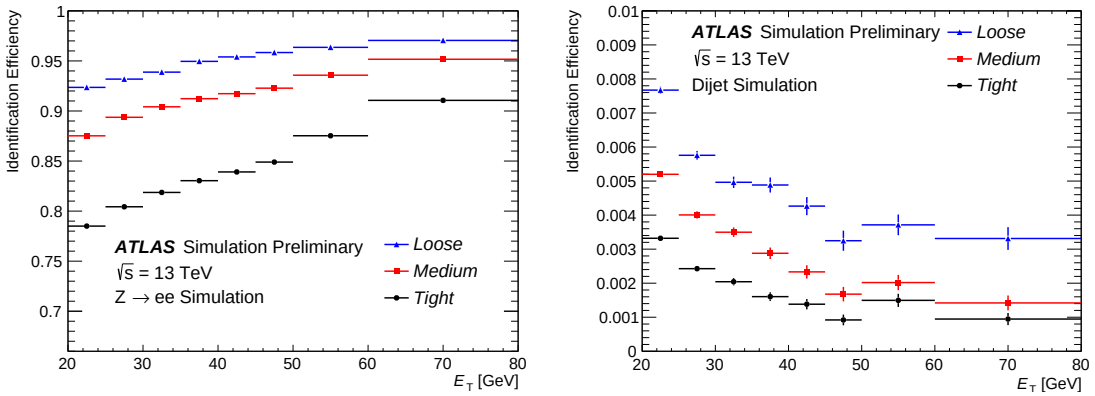


FIGURE 4.1: Efficiency of the electron reconstruction identification working points to reconstruct an object as an electron in simulated data as a function of E_{T} of the reconstructed electron. The efficiency to reconstruct truth electrons in $Z \rightarrow ee$ events is shown in the left figure and the efficiency to reconstruct hadrons (i.e. mis-identification) in dijet events is shown in the right figure. [60]

around the reconstructed electrons, which allow for determination of whether a reconstructed electron is non-isolated. Two types of isolation variables are used: calorimeter based and track based. The calorimeter based variable measures the total transverse energy in a cone around the barycentre of the electron cluster with an energy deposit of size $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ centred around the cluster barycentre subtracted. The track based variable measures the total transverse momentum of all tracks passing basic quality requirements within a cone around the track associated with the reconstructed electron, excluding tracks associated with the electron (this can include additional tracks from converted bremsstrahlung photons). The size of the cone decreases as a function of electron E_{T} for $E_{\mathrm{T}} > 50$ GeV, which is referred to as 'variable' isolation. Isolation operating points are defined using selections on these variables. Of relevance to the work in this thesis is the loose and 'Gradient' operating point. The loose operating point varies the selection on the two isolation variables as a function of E_{T} to define a fixed efficiency for the reconstruction of truth electrons. The Gradient operating point again varies the selections on the two isolation variables as a function of E_{T} but instead allows the selection efficiency to vary, resulting in lower efficiency for truth electrons at low E_{T} and higher efficiency as E_{T} increases. In this way the background rejection is improved at low E_{T} while retaining high signal acceptance at high E_{T}.

Scale factors are applied to electrons in simulated events to ensure that reconstructed electrons which originate from simulated data are found and selected with the same efficiency as reconstructed electrons from real data. The total scale factor per electron is

a combination of a number of weights that correct for mismodelling of tracking properties and shower shapes, and scale reconstruction and selection efficiencies to match those found in data. The scale factor is defined as:

$$w_{\text{electron}} = w_{\text{reconstruction}} \times w_{\text{identification}} \times w_{\text{isolation}} \times w_{\text{trigger}} \tag{4.2}$$

where w is the weight applied and the subscript refers to the portion of electron reconstruction that the corrections are derived for. 'Trigger' refers to the efficiency of finding reconstructed electrons using the online reconstruction algorithms of electron triggers as compared to the reconstructed electrons found by the offline reconstruction algorithms.

The size of w_{electron} is in general slightly below 1.0, with most simulated events given a weight between 0.8 and 1.0. A small number of events are given weights up to 1.3, and a falling distribution of events are given weights down to 0.0.

4.2.3 Muons

The measurement signature of muons is a track in the inner detector associated with a track or track segments in the muon spectrometer. Reconstruction therefore entails reconstruction of tracks in the muon spectrometer and the matching of these tracks to inner detector tracks [47]. Criteria on identification quality and isolation are employed similarly to what is used in electron reconstruction.

Tracks are reconstructed in the muon spectrometer by first forming track segments out of measurements made in individual modules. For measurements in MDT layers, hits are combined with hits from nearby RPC or TGC trigger modules using a Hough transform [63]. A straight-line fit is performed on the hits found in each layer to produce a track segment for one muon spectrometer MDT layer. Track segments in CSC layers are built using a combinatorial search which is separate for the η and ϕ layers. Full muon spectrometer tracks are then built out of hits from track segments in different layers. This is done by performing a combinatorial search starting from a single segment chosen as a seed. Seeds are first chosen from segments in the middle layers, before later also being chosen from segments in inner and outer layers. Segments can initially be used in multiple track candidates; an overlap removal algorithm then the best assignment for a segment to a single track, or allows a segment to be shared between two tracks. The hits

from a track candidate are then fitted with a global χ^2 fit. The χ^2 value is used to accept or reject track candidates. Individual hits may be rejected from a track candidate if they have a large contribution to χ^2. Additional hits may also be added to the track candidate if they are found to be consistent with the track candidate trajectory.

Inner detector track reconstruction is performed according to the standard methods described in Section 4.2.1. Four methods are then used to reconstruct the final reconstructed muons, of which the combined muon method is relevant to the work presented in this thesis. Combined muons are defined using a combination of a track in the muon spectrometer and a track in the inner detector. A combined track is reconstructed by refitting the combination of hits from both the inner detector track and muon spectrometer track. During this refit hits in the muon spectrometer may be added or removed to improve the fit quality. Most combined tracks are reconstructed in an outside-in fashion by extrapolating the muon spectrometer track inward and matching it to an inner detector track. An inside-out method is also used as a complementary approach where an inner detector track is extrapolated to the muon spectrometer and a muon spectrometer track is matched to it. The inside-out method corresponds to approximately 0.5 % of reconstructed muons selected with the medium identification criteria described below.

Identification criteria on the reconstructed muons are defined using selections on the number of hits in different layers in the inner detector and muon spectrometer as well as selections on the relative differences between inner detector and muon spectrometer tracks. Of relevance to the work described in this thesis are the loose and medium working points. Similar to electron identification working points, reconstructed muons passing the medium working point are a subset of those passing the loose working point. Efficiencies to reconstruct muons are shown in Figure 4.2, where the muons have been identified using a tag-and-probe method. In this method muons are selected with a high purity by choosing events which contain a muon pair with an invariant mass close to the Z or J/ψ mass. A 'tag' muon is required to have triggered a single-muon trigger and is required to pass selections that ensure it is a good quality muon. Measurements can then be made of the corresponding 'probe' muon, which has looser selections applied. In Figure 4.2 loose muons are only plotted with $|\eta| < 0.1$ because the selections for loose and medium muons are similar outside of this region. The difference at central η is the presence of loose muons which originate from methods that allow inner detector tracks to be

associated with measurements in the calorimeters or partial measurements in the muon spectrometer. The efficiency is lower for medium muons where reconstructed muons of this type are excluded, due to reduced instrumentation in the muon spectrometer in this region to make space for cabling and services for the calorimeters and inner detector.

The isolation requirements used for reconstructed muons are the same as those used for reconstructed electrons, with the same track-based and calorimeter-based isolation methods employed. In the track-based isolation method the inner detector track associated with the muon is excluded. In the calorimeter-based isolation method, energy clusters within $\Delta R < 0.1$ of the reconstructed muon are considered as deposits originating from the muon and are excluded.

Scale factors are applied to reconstructed muons following a method similar to what is done for reconstructed electrons. An additional scale factor is included to correct mismodelling of the association of muon tracks to the primary vertex.

4.2.4 Hadronic Jets

The measurement signature of a hadronic jet is a number of tracks pointing to a number of collimated energy deposits in the hadronic calorimeter with some energy deposits also in the electromagnetic calorimeter[4]. Reconstruction can therefore entail the clustering of either nearby tracks, or nearby calorimeter deposits, or reconstruction based on the combination of tracks and calorimeter deposits. However for the reconstructed jets relevant for the work in this thesis, reconstruction entails only clustering of nearby calorimeter deposits from all calorimeters, although matching to tracks is used for calibration of reconstructed jets and for definition of jet quality criteria. A number of calibrations and corrections are applied to the calorimeter deposits and reconstructed jets to ensure they are well reconstructed. Quality criteria are used to identify reconstructed jets which were mis-reconstructed from non-collision sources or detector noise. A matching of reconstructed jets to reconstructed tracks is also used to identify jets that come from the primary vertex and consequently identify jets which originate from pile-up vertices.

The first step in jet reconstruction is clustering measurements in the calorimeters into 3D topological clusters referred to as 'topo-clusters' [64]. Topo-clusters are intended to separate calorimeter measurements arising from jets from measurements arising from

[4]Electromagnetic energy deposits dominantly arise from the decay $\pi^0 \to \gamma\gamma$

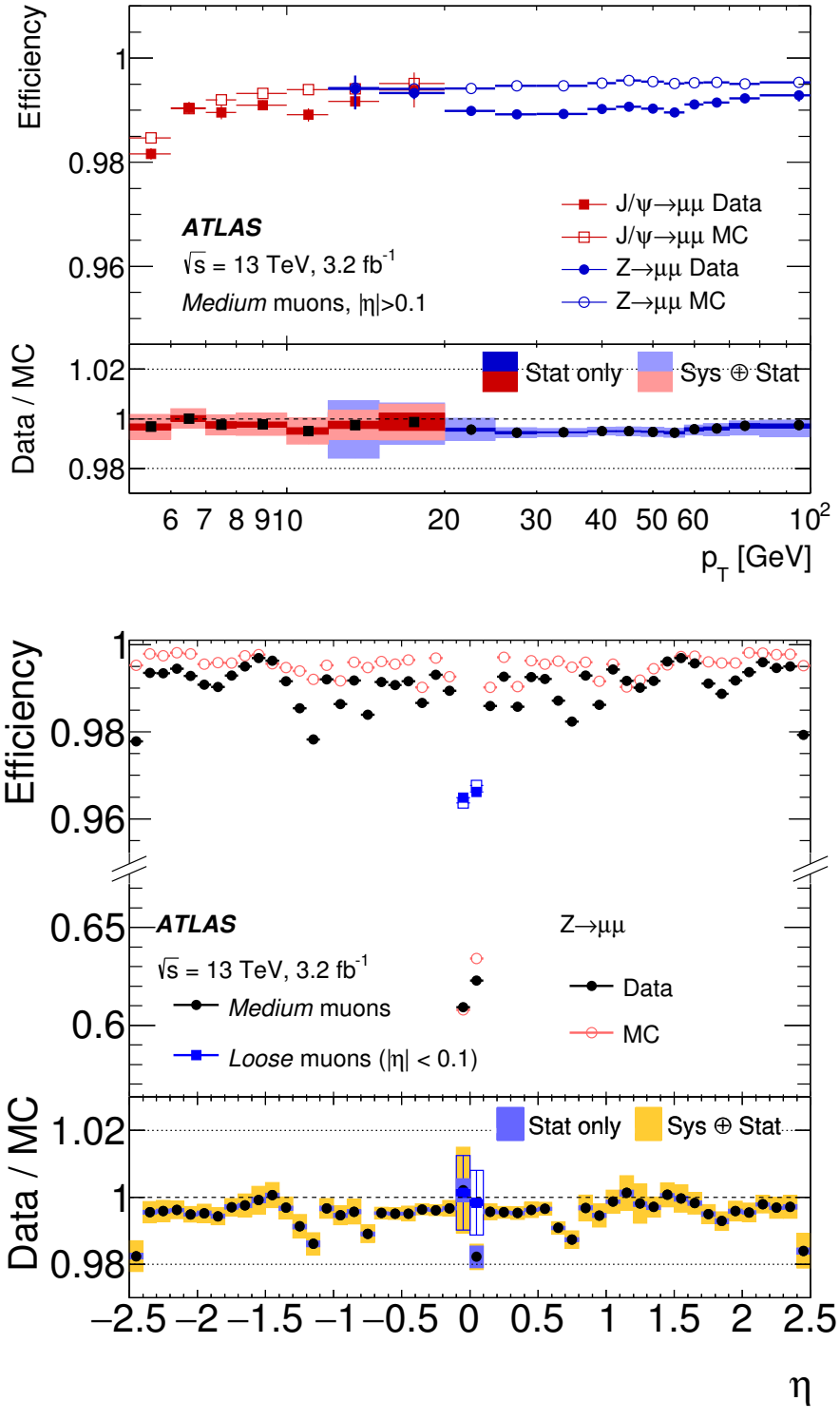


FIGURE 4.2: Efficiency of the muon reconstruction to reconstruct muons in real data and simulated data. The efficiency as a function of p_{T} of the reconstructed muon is shown in the top figure for medium muons with $|\eta| > 0.1$ in $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events. The efficiency as a function of η of the reconstructed muon is shown in the bottom figure for loose muons with $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 0.1$ and medium muons with $p_{\mathrm{T}} > 10$ GeV in $Z \to \mu\mu$ events. The y-axis scale is broken in the bottom figure to allow differences between data and simulated data to be more clearly seen. [47]

electronic noise or products of pile-up interactions (both are collectively referred to as noise). Individual topo-clusters do not necessarily represent the energy deposit of a full shower arising from a single particle but may instead represent a partial shower or merged response from a number of particles. Topo-clusters are produced using a growing-volume algorithm which uses a signal-to-expected-noise ratio for energy measured in a calorimeter cell to iteratively combine nearby calorimeter measurements. This algorithm works by initially selecting a calorimeter cell with a high signal-to-expected-noise significance and combining it with two nearby cells of appropriate significance to form a seed cluster. Neighbouring calorimeter cells which pass appropriate significance thresholds are then iteratively merged into the topo-cluster until the last set of neighbouring cells included does not pass the significance threshold. Overlapping seeds are merged into the same topo-cluster. Final topo-clusters may be too large to provide detailed measurement of the energy flow within the calorimeter and so a cluster splitting approach splits any topo-clusters which contain two or more local signal maxima calorimeter cells into multiple topo-clusters. The calorimeter energy measurements used to form the topo-clusters are measured on the EM energy scale. This refers to the fact that the ATLAS calorimeters respond differently to electromagnetic and hadronic particles of the same energy (referred to as non-compensating calorimetry). Measurements at the EM energy scale accurately reconstruct energy deposits for electrons and photons but do not include any corrections for the response of hadronic particles.

After topo-clusters have been defined jets are built from them using the anti-k_t jet clustering algorithm [65]. In this algorithm a distance variable d_{ij} is defined between two objects i and j as well as a distance d_{iB} between an object i and the beam. A generic definition of these distance variables is:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \tag{4.3}$$

$$d_{iB} = k_{ti}^{2p} \tag{4.4}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, k_t is transverse momentum (i.e. p_{T}), y is rapidity, and ϕ is azimuthal angle. R is a radius parameter that is chosen to define the size

of resulting jets, and p defines the algorithm used for jet clustering. In the anti-k_t algorithm a value of $p = -1$ is chosen and a value of $R = 0.4$ is typically used. Other jet algorithms are defined by other choices for p: the k_t algorithm uses $p = 1$ [66] and the Cambridge/Aachen algorithm uses $p = 0$ [67, 68]. These jet clustering algorithms iteratively form jets using these distance variables by choosing the smallest distance at a given step and combining objects i and j if the smallest distance is a d_{ij}. If the smallest distance in a given step is instead a d_{iB} then object i is considered a jet and the algorithm proceeds on the remaining non-jet objects. In the anti-k_t algorithm the choice of $p = -1$ ensures that the smallest distances are associated with the hardest objects. This has the effect that jets will be formed by primarily clustering soft objects to hard ones and only clustering soft objects to other soft objects once no nearby hard objects remain. The benefit of this approach is that jet definitions (including the number of jets) will be stable against the addition of soft or collinear radiation. As well the reconstructed jets are either conical in shape or cones with volumes cut out by the presence of nearby harder jets. Due to this robustness against soft and collinear radiation the anti-k_t algorithm is infrared safe and collinear safe when used to define jets in theoretical calculations and as such can be used in the full phase space of partons without divergences or cutoff scales needed to remove divergences. Jets are reconstructed with an implementation of the anti-k_T algorithm in the FastJet software package [69].

The set of reconstructed jets resulting from the anti-k_t algorithm are then calibrated to have the same jet energy scale (JES) as jets formed of stable truth particles in simulated data; they are also calibrated so that the jet four-momentum are defined with the primary vertex as the origin rather than the centre of the detector (jet energy is kept constant in this calibration) [70]. The origin correction results in improved η resolution for jets. A chart showing the sequence of calibration steps and summarising what is done in each step is shown in Figure 4.3.

The first calibration step is the correction of jet directions to originate from the primary vertex. Anti-k_t jets are reconstructed to originate from the origin of the spatial axes (i.e. the centre of the ATLAS detector) and so the origin correction recalculates the four-momentum of each jet using the reconstructed primary vertex of an event as the origin. This improves the pseudorapidity resolution of the jets.

The next calibration step is correcting for effects of pile-up [71]. This is split into a jet

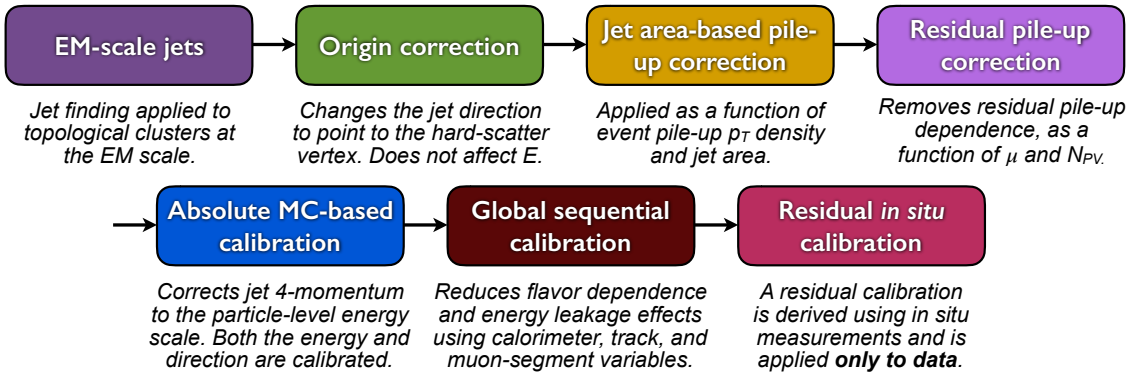


FIGURE 4.3: Sequence of calibrations applied to reconstructed jets. 'EM-scale jets' refers to the input set of reconstructed jets. [70]

area p_{T} density calibration and a residual correction. The jet-area-based correction uses knowledge of the average p_{T} density of jets to subtract the p_{T} contribution to a jet which arises from pile-up. The average p_{T} density is calculated per-event by firstly reconstructing jets from topo-clusters using the k_t algorithm with $R = 0.4$. The k_t algorithm is used because of its sensitivity to soft radiation and as such is more sensitive to calorimeter deposits from pile-up. The area A of each k_t jet in the η-ϕ plane is defined by uniformly distributing 'ghost' particles of infinitesimal momentum across the η-ϕ plane prior to jet formation. Once jets have been reconstructed the relative number of ghost particles within a jet can be counted to define A. The p_{T} density of each jet is then calculated as $\frac{p_{\mathrm{T}}}{A}$ and the average p_{T} density ρ is defined as the median of all p_{T} densities in an event. The p_{T} contribution of pile-up to reconstructed jets can then be subtracted using ρ and the area of the jet in question. This jet-area-based correction is only derived in the region $|\eta| < 2$ due to the higher occupancy of jets in the forward regions.

As a result of the central pseudorapidity range used to derive the area-based corrections there is residual dependence of reconstructed jet p_{T} on pile-up contributions in higher occupancy regions or higher occupancy jets. The dependence is observed by comparing the p_{T} of reconstructed jets to the p_{T} of jets reconstructed from truth particles in simulated data, which are referred to as truth jets. Truth jets have no dependence on pile-up due to being formed of only particles originating from the hard interaction. A reconstructed jet and truth jet are compared if they are within $\Delta R = 0.3$ and the reconstructed jet has $p_{\mathrm{T}} > 10$ GeV. The residual p_{T} dependence is split into a dependence on the number of reconstructed primary vertices, N_{PV}, which is sensitive to in-time pile-up,

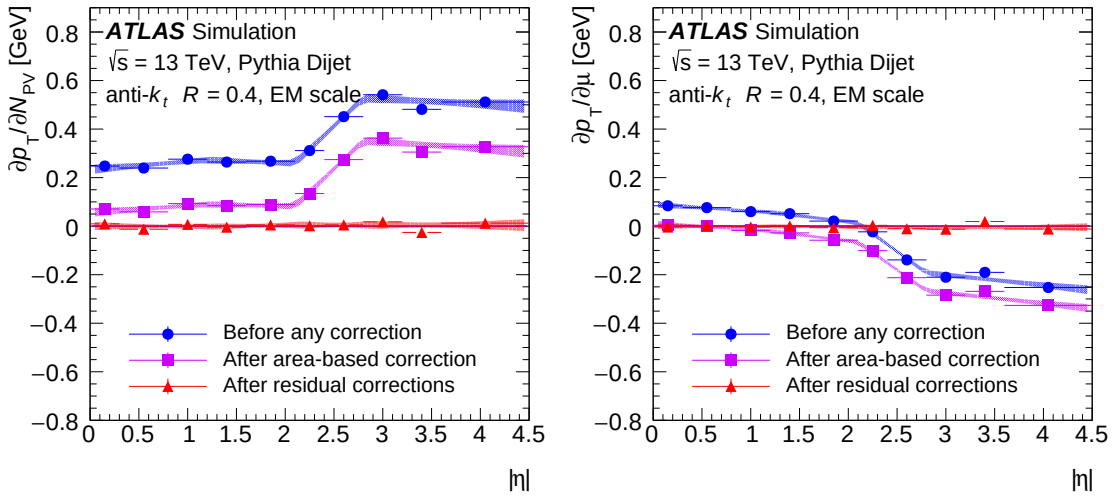


FIGURE 4.4: Dependence of reconstructed jet p_{T} on pile-up as a function of absolute pseudorapidity before and after pile-up corrections are applied. The dependence on N_{PV} (with N_{PV} averaged over μ) is shown in a) which represents the dependence on in-time pile-up. The dependence on μ (with μ averaged over N_{PV}) is shown in b) which represents the dependence on out-of-time pile-up. [70]

and a dependence on the mean number of p-p interactions per bunch crossing, μ, which is sensitive to out-of-time pile-up. A coefficient α is associated with the N_{PV} dependence and a coefficient β is associated with the μ dependence. α and β are fitted as a function of the p_{T} of the truth jet and $|\eta|$ of the reconstructed jet, and the resulting coefficients are used to subtract the residual dependence of the p_{T} of reconstructed jets on pile-up.

The combined pile-up correction to the p_{T} of reconstructed jets is defined as:

$$p_{\mathrm{T}}^{\mathrm{corr}} = p_{\mathrm{T}}^{\mathrm{reco}} - \rho A - \alpha(N_{\mathrm{PV}} - 1) - \beta\mu \tag{4.5}$$

where $p_{\mathrm{T}}^{\mathrm{reco}}$ is the initial p_{T} of the reconstructed jet, $p_{\mathrm{T}}^{\mathrm{corr}}$ is the p_{T} of the reconstructed jet after correction and A is the area of the reconstructed jet. The dependence of the reconstructed jet p_{T} on N_{PV} and μ after each correction is shown in Figure 4.4.

The next calibration step is the absolute calibration of jet energy scale and jet η to truth information in simulated data. This step is performed by reconstructing truth jets and selecting only reconstructed jets which are matched to truth jets within $\Delta R = 0.3$. The reconstructed jets are required to have no other reconstructed jet of $p_{\mathrm{T}} > 7$ GeV within $\Delta R = 0.6$ and only one truth jet of $p_{\mathrm{T}}^{\mathrm{truth}} > 7$ GeV within $\Delta R = 1.0$. The additional matching requirements enforce that the reconstructed jets are isolated, which avoids any

ambiguities when matching reconstructed jets to truth jets. To calibrate the reconstructed jets the mean of the energy response $\frac{E^{\text{reco}}}{E^{\text{truth}}}$ is measured in bins of η_{det}, which is the η of the jet from the centre of the ATLAS detector. This definition of η is used to avoid any ambiguity in which region of the detector is measuring a jet. The reconstructed jet energy scale is then calibrated by applying as a calibration factor the inverse of the mean of the energy response, parametrised by E^{reco}. The calibration of jet η then removes biases in the distribution of $\eta_{\text{det}}^{\text{reco}} - \eta_{\text{det}}^{\text{truth}}$, parametrised as a function of E^{truth} and η_{det}.

The next calibration step is the correction of dependences of jet energy scale on the longitudinal and transverse features of a jet by using global sequential corrections. Each correction in the chain of sequential corrections applies a correction factor of the inverse of the jet response using the same reconstructed jet to truth jet matching as in the absolute calibration step described above. However the corrections are scaled to keep average energy constant which means that they only remove the dependence considered in each step and do not change the overall energy scale. Three of the corrections associate reconstructed inner detector or muon tracks with reconstructed jets. This matching is performed by treating each track as a ghost particle similar to what was used in the jet area calculations in the pile-up corrections: tracks are treated as particles of negligible momentum which are clustered with other measurements as part of standard jet reconstruction. The full list in order of dependences that are corrected for are:

1. The fraction of jet energy in the first layer of the hadronic tile calorimeter;
2. The fraction of jet energy in the third layer of the electromagnetic calorimeter;
3. The number of tracks with $p_{\text{T}} > 1$ GeV associated with a jet;
4. The average p_{T} weighted ΔR between between the jet axis and tracks associated with a jet;
5. The number of muon track segments associated with a jet.

The last calibration step is *in situ* calibrations that correct for differences in jet response between real data and simulated data. These *in situ* calibrations are only applied to jets reconstructed from real data. The corrections are derived by comparing the p_{T} of a reconstructed jet to the p_{T} of another object in the same event that is chosen to be well

modelled. The response $\mathcal{R}_{\text{in situ}}$ is then defined as the average ratio of jet p_{T} to reference object p_{T}, binned by reference object p_{T}. The ratio of jet response in data and simulated data can then be defined:

$$\frac{\mathcal{R}^{\text{data}}_{\text{in situ}}}{\mathcal{R}^{\text{MC}}_{\text{in situ}}} \tag{4.6}$$

where the 'data' and 'MC' superscripts refer to real data and simulated data respectively. This ratio can then be inverted to produce correction factors for jets in data. The calibrations are derived from three sources with the calibrations from each being applied sequentially: η-intercalibration, Z+jet and γ+jet balance, and multi-jet balance. The η-intercalibration balances the p_{T} of a jet reconstructed with $0.8 < |\eta_{\text{det}}| < 4.5$ against a jet with $|\eta_{\text{det}}| < 0.8$. The jets are expected to be balanced in leading order QCD calculations, and therefore any imbalance is attributed to different calorimeter responses. This intercalibration is defined as a function of both p_{T} and η_{det}; for the remaining intercalibrations the correction is defined only as a function of p_{T}. The Z+jet and γ+jet balance compares the p_{T} of a jet with $|\eta| < 0.8$ to the p_{T} of a reconstructed Z boson (which decays to electrons or muons) or of a photon. This calibration covers jets up to a p_{T} of 950 GeV. The multi-jet balance compares the p_{T} of a high p_{T} jet to several low p_{T} jets. This calibration allows the calibration of jets up to a p_{T} of 2 TeV. The value of the ratio from Equation 4.6 is shown for each calibration source in Figure 4.5.

Jet quality criteria are defined by directly applying selections on variables that discriminate against calorimeter noise in LAr calorimeters; applying selections on measurements of the fraction of energy deposits in the different calorimeters; and applying selections on measurements of tracks associated with jets [72]. The noise discrimination variables compare the shape of electronic pulses from LAr measurements to the expected pulse shape from detector simulation. As well, the total energy of negative energy calorimeter measurements within a jet are considered (negative energy measurements are a result of calorimeter noise). The energy fraction variables measure the relative amount of a jet's energy that are measured by the electromagnetic and hadronic calorimeters. The maximum energy fraction in a single layer within a calorimeter is also considered. The track variables measure the relative fraction of the p_{T} of tracks associated with a jet

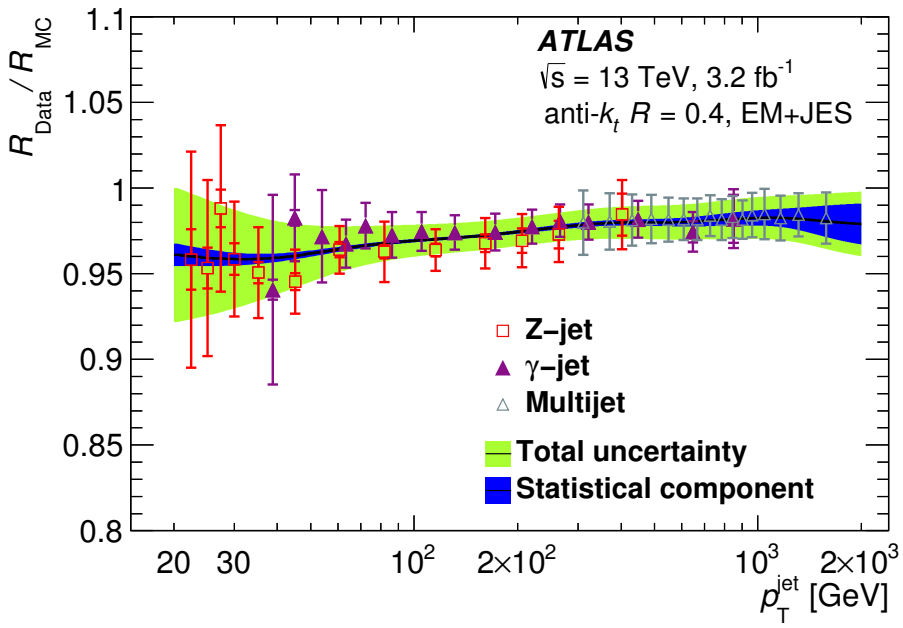


FIGURE 4.5: The value of the ratio of jet response in real data and simulated data as a function of jet p_{T}. The measurements from the three sources of *in situ* calibrations are shown. The symbol given on the y-axis, $\mathrm{R_{Data}/R_{MC}}$, is defined in the same way as the symbol in Equation 4.6. [70]

compared to the p_{T} of a jet. A number of criteria are defined using these variables, for example whether more than a quarter of a jet's energy is measured by hadronic calorimeter deposits of poor quality, and a jet passing any of the criteria is considered a poor quality jet.

To separate reconstructed jets which originate from the primary vertex from reconstructed jets which originate from pile-up vertices a multivariate method called the jet vertex tagger (JVT) is used [71]. The jet vertex tagger is a likelihood function built from two variables which each associate tracks with a reconstructed jet. The first variable compares the total p_{T} of tracks which originate from the primary vertex and are associated with a jet to the total p_{T} of tracks which originate from other vertices and are associated with the same jet. The second variable compares the total p_{T} of tracks which originate from the primary vertex and are associated with a jet to the p_{T} of the jet. Selections can be made on the output score of the likelihood function to define a jet as originating from the primary vertex or not.

4.2.5 b-jets and Jet Flavour Tagging

It is desirable to identify jets that originate from b quarks, labelled as 'b-jets', in physics analyses such as in the reconstruction of top quarks which decay $t \to bX$. Due to the relatively long lifetime of hadrons containing b quarks, generally referred to as 'b-hadrons', as well as other features, such as the high mass of the b quark, b-jets have an experimental signature that can be distinguished from jets which originate from other partons. In particular the relatively long lifetime of b-hadrons of the order of 1.5 ps results in a secondary decay vertex for the decay of b-hadrons which is measurably displaced from the primary interaction vertex by on average approximately 3 mm (the b-hadrons have a proper decay length of $c\tau \approx 450$ μm).

Jets which originate from c quarks are referred to as 'c-jets' and jets which originate from gluons or u, d, or s quarks are referred to as 'light-jets'. Consideration is also given to the separation of b-jets from jets from hadronically decaying tau leptons. The identification of b-jets and separation of jets from other sources is referred to as 'b-tagging'. Jets identified by b-tagging algorithms as containing a b-hadron are referred to as 'b-tags'.

Three basic b-tagging algorithms are defined and then combined using an MVA-based algorithm to provide best separation of b-jets and jets from other sources [73, 74]. Each basic algorithm aims to measure or reconstruct a different aspect of the decay of a b-hadron. The three basic algorithms are the impact parameter based algorithm, the secondary vertex finding algorithm, and the decay chain multi-vertex algorithm. The b-tagging algorithms use reconstructed tracks which are associated with reconstructed jets to measure or reconstruct properties of the decays of hadrons in the jet.

Impact parameter based algorithm

Impact parameter based algorithms aim to measure the presence of tracks from a secondary decay vertex by measuring impact parameters of tracks relative to the primary vertex. A general definition of the impact parameter of a secondary vertex track relative to the primary vertex can be seen in Figure 4.6. Tracks which originate from secondary vertices are expected to have large impact parameter values due to the displacement of secondary vertices from the primary vertex. Two impact parameter variables are defined:

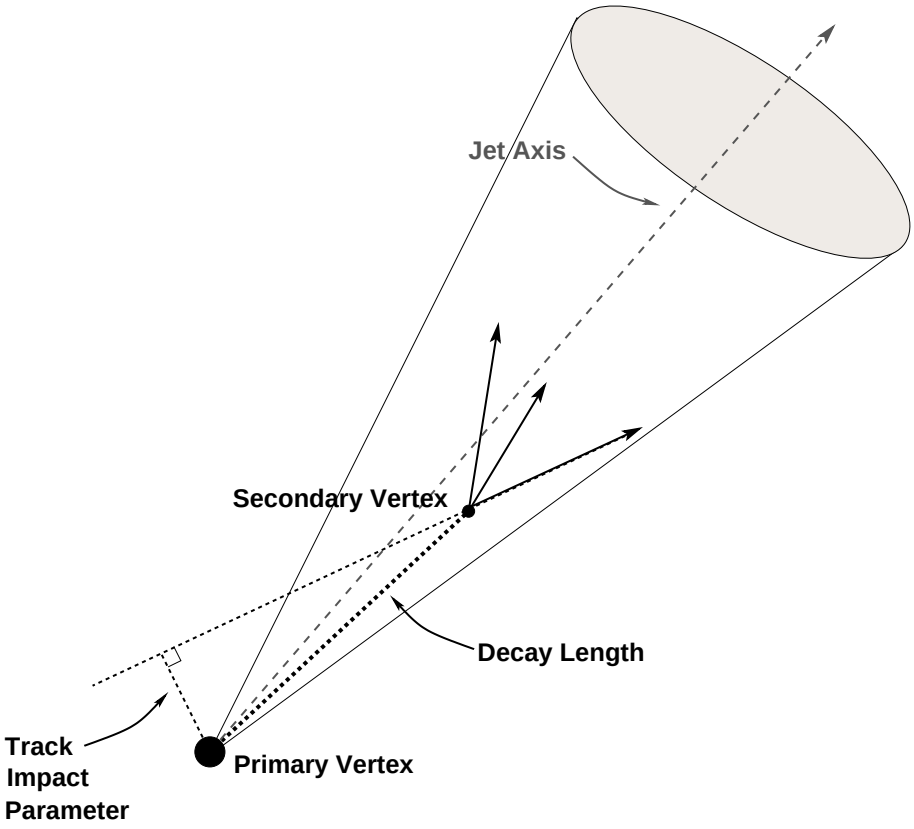
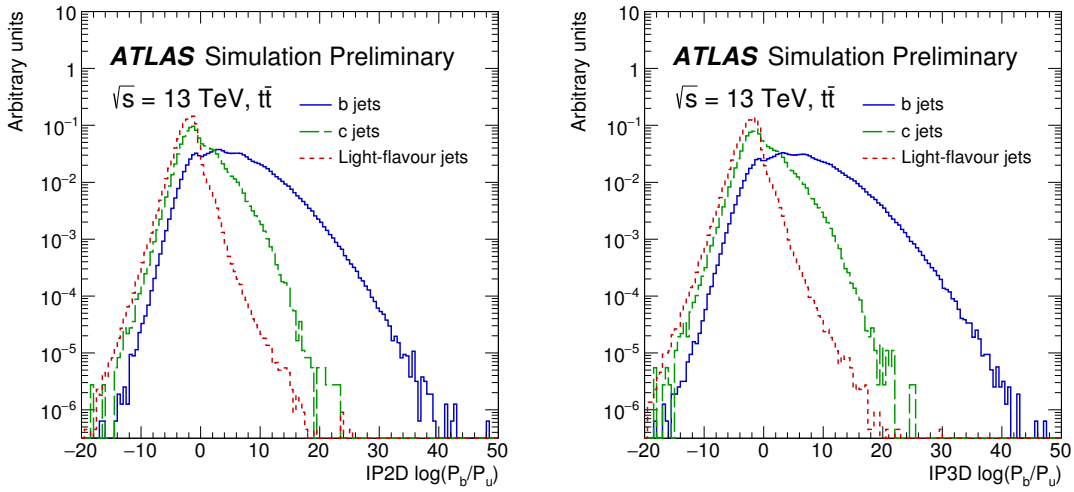


FIGURE 4.6: Drawing of a secondary vertex occurring within a jet [75].

the transverse impact parameter, d_0, which is the distance of closest approach to the primary vertex in the r-ϕ plane; and the longitudinal impact parameter, $z_0 \sin\theta$, which is the distance from the primary vertex in the r-η plane at the point of closest approach in the r-ϕ plane.

Two impact parameter based algorithms are defined: the first, referred to as the IP2D tagger, uses transverse impact parameter significance, $\frac{d_0}{\sigma_{d_0}}$, as a discriminating variable while the second, referred to as the IP3D tagger, uses a two-dimensional template to both the transverse impact parameter significance and the longitudinal impact parameter significance, $\frac{z_0 \sin\theta}{\sigma_{z_0 \sin\theta}}$. A number of probability density functions are defined per-track for the two impact parameter significance variables; the probability density functions are distinguished by different categorisations of the tracks based on hit multiplicities and based on the jet-flavour-hypothesis (b, c, or light) being considered.

A log-likelihood ratio (LLR) is then computed based on the sum of tracks to test the probability of one jet-flavour-hypothesis against another:



a) *b*)

FIGURE 4.7: The log likeihood ratio comparing the probabilities for b-jets and light-jets using the *a*) IP2D and *b*) IP3D algorithms. Distributions are shown for different flavours of jet from simulation. [73]

$$\text{LLR} = \sum_{i=1}^{N} \log\left(\frac{p_{b,i}}{p_{u,i}}\right) \tag{4.7}$$

where p is the value of a probability density function, N is the total number of tracks, i labels the track being considered, and the other subscripts label the jet-flavour-hypothesis used for a given p (u refers to light-jets). In this equation the probability density functions for b- and light-jets are used however the comparisons of b-jets and c-jets, and c-jets and light-jets are also made. The distributions of the LLR in Equation 4.7 for the IP2D and IP3D taggers are shown for different flavours of jets in simulated data in Figure 4.7. The flavour of the jet is known by matching reconstructed jets to truth jets and determining if the truth jet contains a b-hadron, or a c-hadron if no b-hadron is present. This method is used in similar plots in the remainder of this section.

Secondary vertex finding algorithm

The secondary vertex finding algorithm attempts to reconstruct a single inclusive secondary vertex within a jet by testing all pairings of reconstructed tracks associated with a jet for the presence of a vertex displaced from the primary vertex. Vertices formed from track pairs, referred to as two-track vertices, are required to be significantly displaced from the primary vertex (based upon the significance of the distance between each track

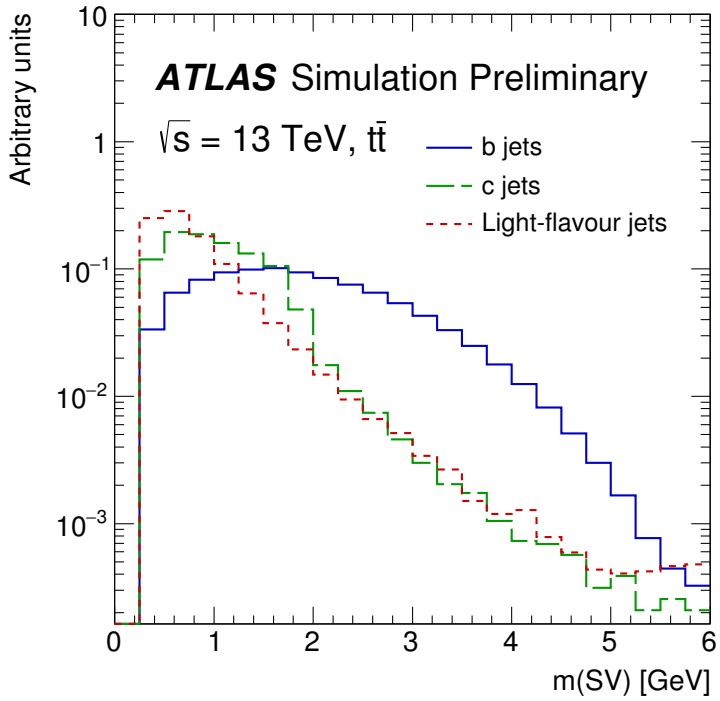


FIGURE 4.8: The invariant mass of the collection of tracks from secondary vertices reconstructed using the secondary vertex finding algorithm. Distributions are shown for different flavours of jet from simulation. [73]

and the primary vertex as well as values from impact parameter based algorithms) and the tracks must be of sufficient quality. Any track pairs which have an invariant mass compatible with other long-lived particles such as K_s mesons and Λ baryons are rejected; the invariant mass is also used to reject track pairs originating from photons. The positions of two-track vertices are also compared against the radius of the innermost inner detector layer to reject any that have originated from interactions with detector material. Surviving two-track vertices are combined into a single vertex with outlier two-track vertices iteratively removed. Properties of the reconstructed secondary vertex such as the invariant mass of the collection of tracks from the secondary vertex can then be used to discriminate between jets of different flavours. An example distribution showing the mass of the reconstructed secondary vertex for different flavours of jets is shown in Figure 4.8.

Decay chain multi-vertex algorithm

The decay chain multi-vertex algorithm attempts to reconstruct the full b-hadron decay chain within a jet, which can involve the production of c-hadrons. This is done by using a Kalman filter method to reconstruct a common line from the primary vertex which also contains b-hadron and c-hadron decay vertices. In this way the flight path of a b-hadron is approximated. This approach produces reconstructed b-hadron and c-hadron vertices if it is possible to do so according to spatial resolution; this can be done even if only a single track is associated to any of these vertices. Properties of the reconstructed decay topology can then be used to discriminate jets of different flavours.

Multivariate combination algorithm

The multivariate combination algorithm, referred to as 'MV2', combines the outputs of the three basic algorithms into a single discriminant variable with better separation of b-jets from jets of other flavours than any of the individual basic algorithms. The MV2 algorithm is a boosted decision tree[5] (BDT) that treats b-jets as the signal to be correctly identified and c-jets and light-jets as the background to reject from signal. The input variables are correlated with the reconstructed jet p_{T} and $|\eta|$; however, the p_{T} and η of the different flavours of jets are reweighted to have the same distribution during the training of the BDT (this reweighting is not used when evaluating the final output of the BDT). This is done to avoid using differences in the kinematics of the jet flavours as discriminating information while still allowing the input variables to be correlated with jet kinematics.

A few different versions of the MV2 b-tagger are trained, where each variant includes a different fraction of c-jets in the collection of background category jets during training and is otherwise identical to other variants. This allows the rejection of light-jets and the rejection of c-jets to be optimised based on how important each is to a physics analysis. Three variants are defined: MV2c00, MV2c10, and MV2c20. MC2c00 contains no c-jets in the background category, MC2c10 contains a c-jet fraction of 7% in the background category, and MV2c20 contains a c-jet fraction of 15% in the background category. Note that the 'cXX' numeric is indicative of the c-jet fraction relative to other MV2 variants and

[5] A description of the functionality of boosted decision trees is given in Section 6.3 when describing the boosted decision trees used in the ATLAS Higgs search that is the focus of this thesis.

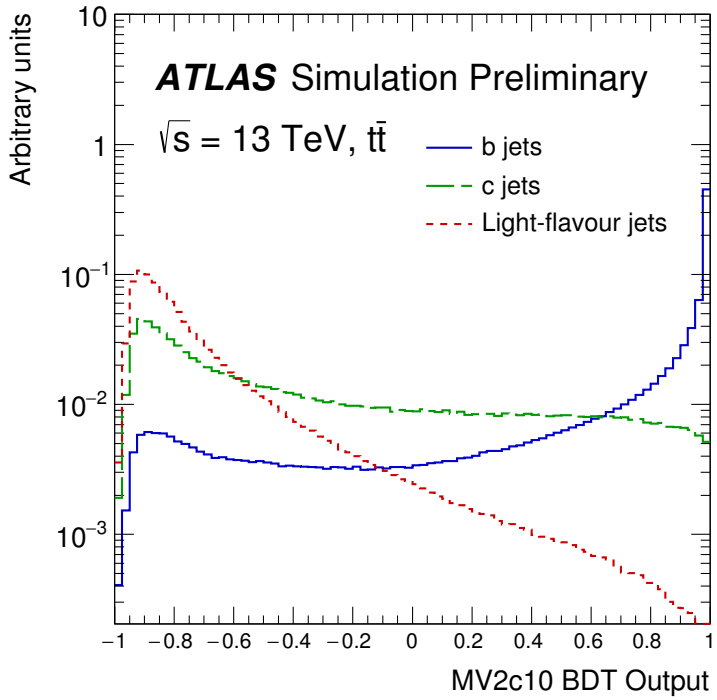


FIGURE 4.9: The output of the MV2c10 *b*-tagging algorithm. Distributions are shown for different flavours of jets from simulation. [73]

does not denote the actual *c*-jet fraction (the values reflect an older version of the tagger where the *c*-jet fractions were indeed 0%, 10%, and 20%). The MV2c10 tagger is typically used due to similar power to reject *c*-jets to the MV2c20 tagger but without the degraded light-jet rejection associated with the MV2c20 tagger. The distribution of the output of the MV2c10 *b*-tagger for different flavours of jets is shown in Figure 4.9.

b-tagging working points are defined based on selections on the MV2 output distribution. Four working points are defined: loose, medium, tight, and very tight. The working point definitions for the MV2c10 tagger and their performance are shown in Table 4.1. The values are derived from simulation for jets with $p_{\mathrm{T}} > 20$ GeV. For a given working point any reconstructed jet which passes the selection value on the MV2c10 output is considered a *b*-tag while any jet that fails the selection is considered untagged; in this way jets can be considered to be tagged (or untagged) at a given working point.

A significant challenge in the use of *b*-taggers is ensuring that the efficiency to select jets of a given flavour is the same in real data and simulated data. Dedicated calibrations

TABLE 4.1: Definitions and performance of the different b-tagging working points for the MV2c10 b-tagger. The rejection values are defined as the inverse of the efficiency to select a jet of the appropriate type. Values are taken from Reference [73].

Working point	BDT cut value	b-jet efficiency [%]	c-jet rejection	light-jet rejection	τ rejection
very tight	0.9349	60	34	1538	184
tight	0.8244	70	12	381	55
medium	0.6459	77	6	134	22
loose	0.1758	85	3.1	33	8.2

from data are performed using a number of methods [74, 76][6] which result in a scale factor to apply to simulated events to correct the b-tagging efficiencies in simulated data. Typically this is performed at a single working point at a time, which defines that all b-tags in an analysis are tagged at the same working point. However the work described in Chapter 6 makes use of an extended calibration which calibrates the ensemble of working points simultaneously, which allows jets to be b-tagged at different working points within the same event. This is referred to as 'pseudo-continuous calibration', where continuous calibration would calibrate the continuous MV2c10 score as opposed to working points at fixed cut values.

4.2.6 Taus

Tau leptons decay either leptonically into an electron or muon plus neutrinos, or hadronically into one or three charged hadrons plus additional neutral hadrons. Decay modes containing five charged hadrons are also possible however they are a very low fraction of tau decays and are neglected. The electrons and muons from leptonically decaying taus can be reconstructed in the same way as electrons and muons from other sources however the hadronic decay modes provide a hadronic jet signature that can be separated from hadronic jets originating from partons [77, 78]. The measurement signature for hadronically decaying tau leptons is therefore one or three tracks pointing to a narrow shower in the calorimeters.

[6]Most documentations of calibrations with 13 TeV data have not yet been published and instead an older paper which contains information on calibration methods and the only 13 TeV b-tagging calibration paper published so far are cited.

Reconstruction of hadronically decaying taus is seeded by jets reconstructed using the methods described in Section 4.2.4. Tau candidate objects are formed from topo-clusters and tracks within $\Delta R < 0.2$ of the barycentre of the seeding jet. Tracks are also considered within $0.2 < \Delta R < 0.4$ to provide a measure of the isolation of the tau candidate. Energy calibrations similar to those for the reconstruction of hadronic jets are applied to the tau candidates, to correct energy measurements made by the detector to the average energy of truth decay products in simulated data.

No explicit rejection of tau candidates originating from non-tau sources is done during the reconstruction of tau candidates and instead identification criteria are used to reject these tau candidates. The identification criteria are defined using a boosted decision tree which uses a number of variables defined using the tracks and energy measurements associated with a tau candidate. Two separate BDTs are trained, one for one-prong taus and the other for three-prong taus. Loose, medium, and tight identification criteria are defined by placing selections on the BDT outputs.

Tau reconstruction in tau triggers (i.e. online tau reconstruction) is performed differently due to the measurement limitations of the L1 ATLAS trigger system. In L1 tau triggers a tau candidate is formed by looking at calorimeter energy deposits in a 0.2×0.2 region in $\eta \times \phi$ space formed of four 0.1×0.1 granularity calorimeter 'towers'. L1 trigger requirements are placed on the total transverse energy of the two most energetic adjacent towers. The η and ϕ of any deposits that pass the L1 thresholds are given to the HLT which reconstructs a tau candidate using topo-clusters within $\Delta R < 0.2$ of the L1 tau candidate direction. The energy is calibrated similarly to what is described above. Multi-stage tracking methods which are described in detail in Chapter 5 are used to reconstruct tracks associated with the tau candidates. Identification BDTs for online tau reconstruction are then used to define identification working points in the same way as described above.

4.2.7 Missing Transverse Momentum

Missing transverse momentum originates from weakly interacting particles that pass through ATLAS detector material without interacting and consequently cannot be reconstructed (such as neutrinos produced from the decays of W bosons to a lepton and

neutrino). Due to the hermetic coverage of ATLAS in the ϕ plane the full transverse momentum system of the detectable collision products can be measured. As such, by considering the momentum balance of the reconstructed objects the missing transverse momentum vector, $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$, can be defined to balance the full transverse system [79]. Components of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ are defined by the negative sum of the appropriate x and y components of the momenta of all objects reconstructed in the detector as well as a 'soft term' that includes momentum from all tracks matched to the primary vertex that are not included in any reconstructed object:

$$E_{x(y)}^{\mathrm{miss}} = p_{x(y)}^{\mathrm{miss},\mu} + p_{x(y)}^{\mathrm{miss},e} + p_{x(y)}^{\mathrm{miss},\gamma} + p_{x(y)}^{\mathrm{miss},\tau} + p_{x(y)}^{\mathrm{miss,jets}} + p_{x(y)}^{\mathrm{miss,soft}} \tag{4.8}$$

where the superscript refers to the type of object considered (γ refers to reconstructed photons) and the order of the sum reflects the order of object reconstruction. The order of reconstruction is chosen to minimise double counting of energy measurements.

The objects used to define $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ are reconstructed and calibrated according to the approaches previously described in this chapter. Additionally, selections that are applied to objects as part of a physics analysis must also be applied to objects entering the defintion of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ to ensure a consistent interpretation of a given event. An overlap removal procedure is also implemented to ensure that detector measurements are not shared between reconstructed objects and multiply counted in the definition of $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ as a result. This procedure removes any objects that are considered to overlap with a higher priorty object (defined according to the order of the sum in Equation 4.8.

Chapter 5

Development of Algorithms in the Inner Detector Trigger for Tau Lepton Triggers

5.1 Introduction

Constant development and optimisation of the methods used in the ATLAS trigger system is vital to ensuring that ATLAS can continue to efficiently collect data as LHC collision conditions become increasingly strenuous. Refinement of the track finding performed as part of the HLT trigger is particularly important due to the computational cost of track reconstruction. The author has worked on the improvement of these track finding methods as part of Run 2 of the LHC with efforts towards ensuring that the methods used are efficient with respect to the selection and reconstruction of tracks of interest as well as efforts towards reducing the computation time. The developments discussed in this chapter were worked on as part of tau lepton triggers which are not used elsewhere in this thesis, however the studies performed and approaches developed could be extended to other triggers which use the two-stage tracking methods described below. More triggers, such as muon triggers, may utilise these methods in future as collision conditions continue to develop.

Other results from the inner detector trigger ATLAS technical group have been presented by the author at the 17th International Workshop on Advanced Computing and

Analysis Techniques in Physics Research [80][1] and the 26th International Symposium on Nuclear Electronics and Computing [81].

5.2 The Inner Detector Trigger

As mentioned in Section 3.2.5 the inner detector trigger is the portion of the ATLAS high level trigger which is responsible for the reconstruction of tracks from charged particles as well as the reconstruction of vertices[2] [82]. Tracks are essential in the accurate reconstruction of most physics objects reconstructed in the HLT and are therefore very important for making accurate trigger decisions of whether to permanently store a collision event based upon the reconstructed objects. However track reconstruction is also computationally expensive and can therefore lead to bottlenecks in the running of the HLT due to object reconstruction not being able to proceed until track finding has finished. Due to the online nature of the HLT this could therefore result in collision events being missed due to trigger software taking too long to process preceding events. Track reconstruction also becomes more important but also more challenging as pile-up increases and the number of tracks and vertices to be reconstructed increases as a result. Track finding within the inner detector trigger is peformed using similar methods to the offline approaches described in Section 4.2.1. Track finding is split into two stages: a pattern recognition stage labelled as the Fast Track Finder (FTF) which is followed by (and provides inputs for) a hit ambiguity solving and track scoring stage labelled as precision tracking (PT). The performance of the online tracking algorithms is comparable to those used offline. Track reconstruction in the inner detector trigger only considers hits in the pixel and SCT detector layers.

To minimise computation time and allow track finding to be performed online the inner detector trigger runs track finding in regions of interest (RoI) that enclose a reduced volume within the detector. RoI volumes are defined by an extent along the beamline, Δz, and an extent in pseudorapidity and azimuthal angle, $\Delta\eta$ and $\Delta\phi$ respectively. These widths define a roughly wedge shaped region within the detector extending from the

[1] The author prepared a poster to be presented at the conference however he was unfortunately unable to attend to present it himself.

[2] The work described in this section does not include work on vertex reconstruction and so focus is placed on track reconstruction in the inner detector trigger.

beamline. RoIs are also defined by an η value and ϕ value that determine which direction the centre of the RoI points in as well as a z value which defines the position of an RoI along the beamline. The quantities z_+ and z_- are defined as $z + \Delta z$ and $z - \Delta z$ respectively[3]; equivalent quantities are also defined for η and ϕ. The initial directions of an RoI are seeded by information from the preceding L1 trigger; for example the η and ϕ of a calorimeter cluster measured by a tau lepton trigger. The widths of RoIs are defined appropriately to the type of physics object being reconstructed. In general the widths are chosen to approximately cover the expected spatial resolution of measurements used in track finding, for example the z resolution of track reconstruction, as well as reflecting the physics needs of the object in consideration, for example choosing $\Delta\eta$ and $\Delta\phi$ widths that are wide enough to include all tracks from a jet.

Track finding studies and monitoring within the inner detector trigger are normally performed by comparing tracks found by the online tracking algorithms to a specified collection of 'reference' tracks. For example all tracks found by the offline tracking algorithms may be used as reference tracks. By attempting to spatially match reference tracks to online tracks it is possible to compare the properties of any tracks which are matched. Track finding efficiency can also be defined for the reference tracks by counting the number of reference tracks which are successfully matched to an online track and comparing this number to the total number of reference tracks (i.e. including reference tracks which are not found by the online tracking algorithms):

$$\text{Efficiency} = \frac{N_{\text{tracks}}^{\text{online, ref}-\text{matched}}}{N_{\text{tracks}}^{\text{ref}}} \tag{5.1}$$

Where the N_{tracks} is the number of tracks, the superscript "online, ref-matched" refers to tracks reconstructed by online tracing algorithms that are matched to a reference track, and the superscript "ref" refers to tracks from the reference track collection. To ensure that tracking studies are only performed relative to reference tracks which could be expected to be found within the inner detector trigger, the reference track collections are filtered to include only tracks which are within trigger RoIs in consideration.

[3]Width quantities such as Δz are usually quoted as $\Delta z = \pm w$, where w is the width, which implies the definition of w as the half-width from the RoI centre; however the definition quoted here for z_+ and z_- and equivalent quantities assume the choice of the positive width i.e. $\Delta z = +w$.

5.2.1 Two-Stage Tracking Method

To ensure that track finding for triggers which measure objects with a number of expected tracks can be run within the necessary time budget at the high pile-up conditions of Run 2 of the LHC a two-stage tracking method is employed. This method uses two RoIs of differing sizes to reduce the total volume that tracking must be run in. The two-stage tracking method is employed in tau lepton triggers (it is also employed in b-jet triggers, however the method used by the tau triggers is focused on here). The two-stage method works by first defining an RoI which is narrow in angular widths, $\Delta\eta = \Delta\phi = \pm 0.1$, but is extended along the beamline, $\Delta z = \pm 225$ mm. The $\Delta\eta$ and $\Delta\phi$ widths are chosen to reflect the granularity of calorimetry in the L1 trigger. The Δz width is the full length of the colliding region of the beamline. The η and ϕ of this first-stage RoI are taken from the calorimeter deposit measured by the L1 tau trigger. The fast track finder is then run in the first-stage RoI to produce an initial collection of tracks. From this collection a track which is determined to have most likely come from the decay of the tau lepton is used to seed the z position of a second-stage RoI. The initial selection criteria for this 'seed track' was picking the highest p_{T} track in the first-stage RoI track collection, however this method was refined as discussed in Section 5.3. The second-stage RoI is much wider in angular widths, $\Delta\eta = \Delta\phi = \pm 0.4$, however it is much narrower along the beamline about the z position seeded from the tau track, $\Delta z = \pm 7$ mm. The $\Delta\eta$ and $\Delta\phi$ widths originate from the spatial volume used to determine isolation in tau reconstruction, $\Delta R = 0.4$. The Δz width is chosen to approximately cover the z resolutions of tracks and measurements that seed tracks for tracks that are central in the RoI. The fast track finder is re-run in the second-stage RoI followed by running the precision tracking to produce the final track collection that can then be used for the reconstruction of the tau. Compared to a single-stage tracking method, which uses a single RoI with $\Delta\eta = \Delta\phi = \pm 0.4$ and $\Delta z = \pm 225$ mm, the total volume that tracking is run in is much reduced. A schematic comparison of the RoIs used in the single-stage and two-stage methods is shown in Figure 5.1.

In this way the two-stage tracking method is able to rapidly make an accurate determination of the z position of a measured tau and then perform full track reconstruction in a much shorter (in Δz) RoI than the single-stage tracking method RoI. Measurements of

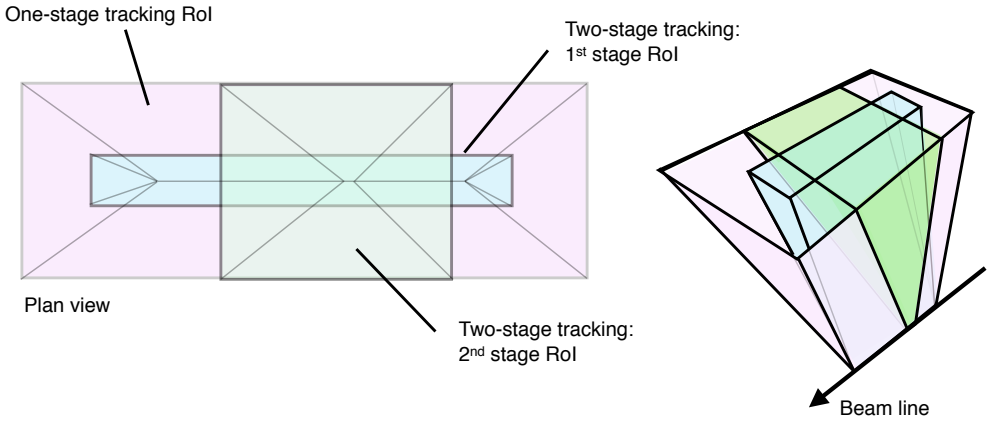


FIGURE 5.1: A schematic of the regions of interest used by tau lepton triggers in the single-stage track finding method and in the two-stage track finding method. The regions of interest are not drawn to scale. [82]

algorithm timing collected from 2015 data found that the single-stage method resulted in a mean FTF timing of 66 ms and a mean precision tracking timing of 12 ms while the two-stage method timings were a mean of 23 ms (21 ms) for the FTF in the first-(second-)stage RoI and a mean of 5 ms for precision tracking, resulting in an improvement to the total times of more than 30% (the exact timings and difference in timings are dependent on the computer used to run the algorithms and on the collision conditions) [82].

Good running of the two-stage method (i.e. without track losses and resulting losses of reconstructed taus compared to the single-stage method) is dependent on a number of assumptions and requirements:

- Firstly the track selected to seed the position of the second-stage RoI is indeed from the measured tau or near to it. If it is not then the second-stage RoI will be mispositioned and will not contain tracks from the tau. As such the full tau reconstruction cannot take place which will result in a rejection of the event by the tau trigger;

- Secondly the angular widths of the first-stage RoI must not be too narrow to fully enclose tracks from the measured tau. If they are then the tracks will not be reconstructed and the second-stage RoI could be seeded by a non-tau track and be mispositioned;

- Thirdly the Δz of the second-stage RoI must be wide enough to include tau tracks which are displaced from the z position of the tau which produced the tracks. If it

is not wide enough then tau tracks may be lost and the full tau reconstruction will be inaccurate.

5.2.2 Tau Lepton Triggers Within the Inner Detector Trigger

As discussed above track reconstruction within tau triggers is performed using the two-stage tracking method where the first-stage RoI is seeded from the L1 tau calorimeter cluster. Final tracks found using this method are then associated with the tau calorimeter cluster formed from topo-clusters in the HLT to form a tau candidate [83]. Tracks within $\Delta R < 0.2$ of the calorimeter cluster direction are considered as 'core' tracks and tracks within $0.2 < \Delta R < 0.4$ are considered 'isolation' tracks. Tau candidates with $1 \leq N_{\text{track}}^{\text{core}} \leq 3$ and $N_{\text{track}}^{\text{isol.}} \leq 1$ are accepted and tau identification criteria are computed to allow selection of final reconstructed taus. HLT trigger decisions can then be made based on the reconstructed taus.

Throughout the developments presented in this chapter studies were performed using events passing a modified tau trigger with a 25 GeV threshold, which is designed for the estimation of performance in the inner detector trigger. The modification is that any selections on the reconstructed tracks as part of the trigger decision are neglected. This change results in an inclusive collection of saved events which do not depend on the performance of the inner detector trigger. This allows full offline analysis of the inner detector trigger algorithms including events which would otherwise be rejected in the trigger due to poor tracking performance. This trigger is referred to as the `tau25_idperf` trigger.

5.3 Improving Efficiency to Correctly Position Second-Stage Regions of Interest Using a Track Quality Selection

As mentioned in Section 5.2.1, the positioning of the second-stage RoI was initially defined by the z position of the highest p_{T} track found by the fast track finder in the first-stage RoI. This choice reflected the assumption that the highest p_{T} track from a tau decay will have higher momentum than other tracks found in the first-stage volume, which are likely to be tracks from soft processes such as pile-up vertices or are incorrect track

assignments from fake coincidences of hits in the tracking detectors (referred to as fake tracks). However it was observed that in occasional cases the highest p_{T} track in a first-stage RoI would not originate from a tau. These tracks would likely be at a z position away from the position of the tau. As a result any second-stage RoIs which were seeded by the position of these non-tau high-p_{T} tracks would not reconstruct the tau tracks and therefore full tau reconstruction would not be possible.

It was also observed that in some cases where the highest p_{T} track did not originate from a tau the p_{T} of the track was very high (on the order of hundreds of GeV) and track properties associated with the track such as number of hits were indicative of a poor quality track. As such it was suspected that in some cases fake tracks were produced from coincidences of hits in the fast track finder that were reconstructed with a relatively straight trajectory. Therefore these fake tracks are reconstructed with very high p_{T} and are chosen to seed the position of the second-stage RoI instead of a tau track. To ensure that the chosen track was indeed a real track studies were performed to extend the selection of the track to include track quality requirements which would reject fake tracks while ideally keeping all leading p_{T} tracks from taus. Some events were also observed where the highest p_{T} track did not originate from a tau but had reasonable p_{T} and track quality. In these cases it appeared that by chance genuine tracks from other sources were produced with higher p_{T} than tau tracks. This effect could for example be the result of a tau decay which includes neutral hadrons that carry a significant amount of the tau momentum, leaving a charged component with lower momentum than non-tau tracks in the same RoI. It was expected that it would not be possible to separate these genuine non-tau tracks from genuine tau tracks using a quality selection in the same way it was hoped that fake tracks could be removed; as such, focus was placed on characterising and removing the fake tracks.

Studies were performed using a simulated dataset of $Z \rightarrow \tau^{+}\tau^{-}$ events produced using the PYTHIA 8 simulation package [84] and the full ATLAS detector simulation. 5000 events from this dataset were used as an input for these studies; in these events 1294 tau trigger first-stage RoIs passed a selection that they contain at least one offline track that is spatially matched to a tau truth track, and were subsequently studied. Tracks which were reconstructed by online tracking algorithms (within the inner detector trigger) were determined to have originated from taus by performing spatial matching to tracks which

were reconstructed by offline tracking algorithms, and which themselves were spatially matched to 'truth tracks' that show the trajectory of truth particles produced from the decay of taus. This matching was done by requiring that an online track and any offline track pass the criteria:

$$\sqrt{\left(\frac{\eta_{\text{online}} - \eta_{\text{offline}}}{0.4}\right)^2 + \left(\frac{\phi_{\text{online}} - \phi_{\text{offline}}}{0.4}\right)^2 + \left(\frac{z_{\text{online}} - z_{\text{offline}}}{7}\right)^2} < 1 \qquad (5.2)$$

where the subscripts "online" and "offline" refer to parameters of the online track and offline track under consideration. A similar but stricter criteria is placed on the matching of an offline track to any truth track:

$$\sqrt{\left(\frac{\eta_{\text{offline}} - \eta_{\text{truth}}}{0.02}\right)^2 + \left(\frac{\phi_{\text{offline}} - \phi_{\text{truth}}}{0.02}\right)^2 + \left(\frac{(z_{\text{offline}} - z_{\text{truth}})\sin\theta_{\text{offline}}}{2}\right)^2} < 1 \qquad (5.3)$$

where the subscripts "offline" and "truth" refer to parameters of the online track and truth track under consideration.

Matching to offline tracks ensures that only cases where tau tracks are expected to be reconstructed were considered (according to the track matching definitions used). The parameterisation of the online track to offline track matching was chosen such that a matched offline track should always be contained in a second-stage RoI that was seeded by the online track under consideration. The parametrisation of the offline track to truth track matching was chosen to be as tight as measurement resolution allows, to ensure that a matched truth track results in the offline track under consideration.

By checking if the track in first-stage RoIs that was chosen to seed the position of the second-stage RoI was matched in this way for the `tau25_idperf` trigger it was possible to determine if the chosen track was a tau track or not. In this way an efficiency of correct second-stage RoI placement could be defined as the number of first-stage RoIs where the chosen track was a tau track compared to the number of first-stage RoIs which contained a tau track in the track collection. Leading tau tracks were also considered in RoIs where the highest p_{T} track is a non-tau track. This allows inclusion of tracks which should be chosen after the removal of non-tau tracks.

Once second-stage-seeding track choices were categorised as originating from a tau

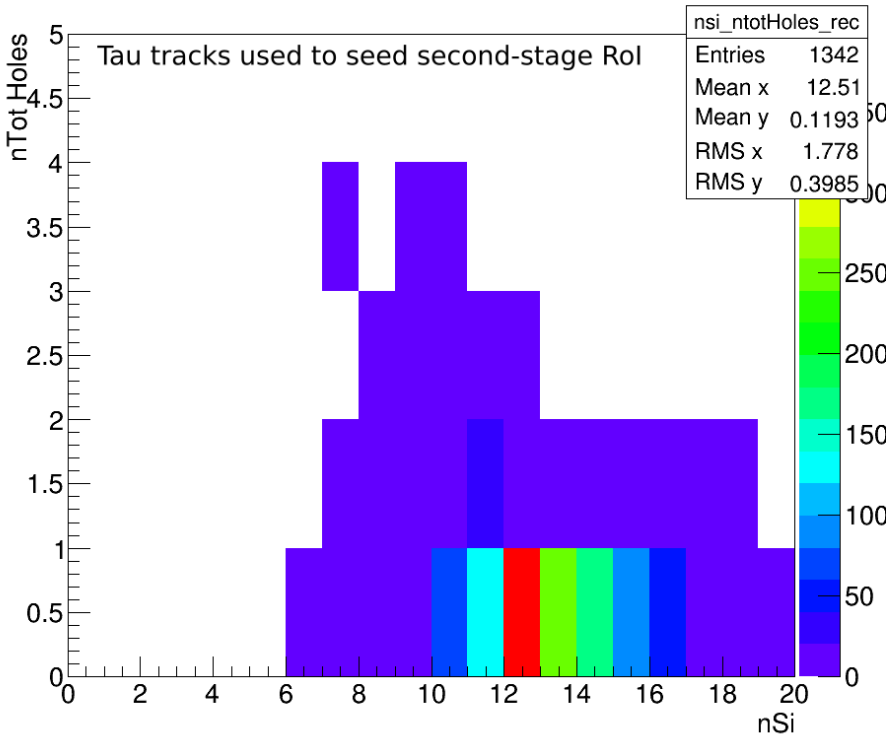
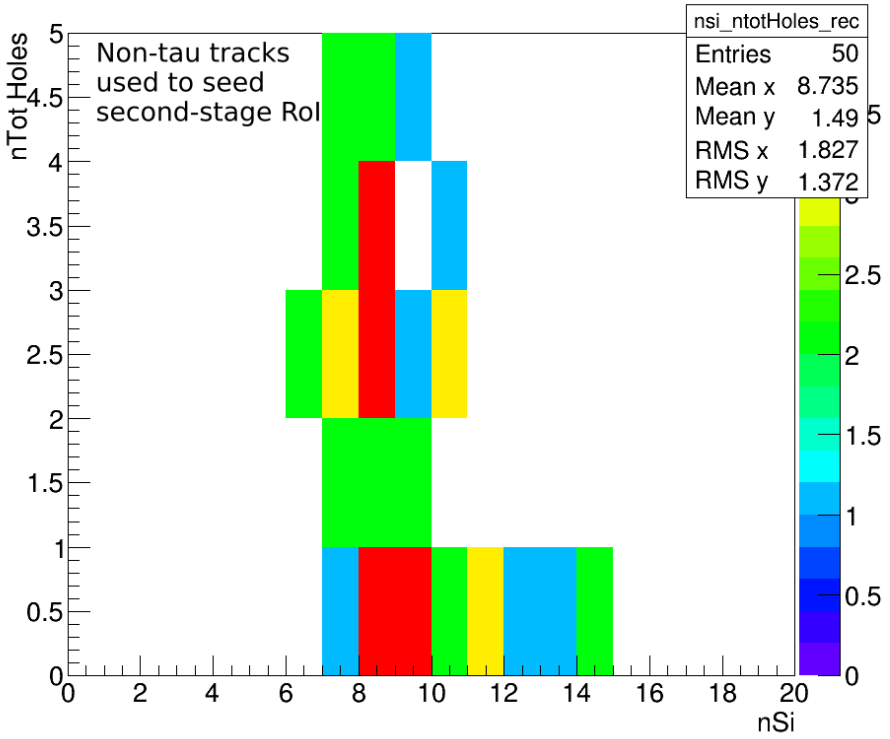


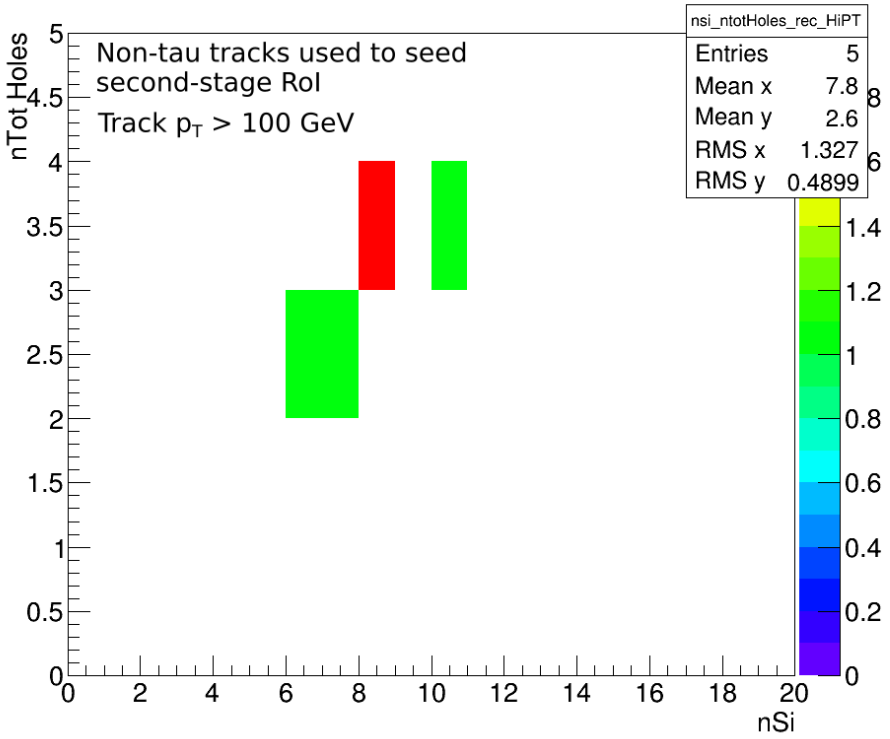
FIGURE 5.2: Number of tracks as a function of total number of track holes, nTotHoles, and total number of track hits, nSi, for tau tracks which are used to position second-stage RoIs.

or non-tau, track properties for each case were compared to identify any with good separating power between tracks from taus and tracks from non-taus. In particular, because the focus was on spurious high p_{T} tracks, comparisons were also made between tracks from taus and non-tau tracks above a high p_{T} threshold of 100 GeV. This would ideally allow spurious high p_{T} tracks and genuine non-tau tracks to be considered separately. Example results for such comparisons of multiplicities of the total number of track hits and total number of track holes, where a hole is a missing expected hit in an active detector layer, are shown for tau tracks which seed second-stage RoIs in Figure 5.2 and for non-tau tracks which seed second-stage RoIs in Figure 5.3.

It was seen that multiplicities of track hits and track holes had good separation between tau tracks and non-tau tracks. The high p_{T} non-tau tracks had a low number of track hits and relatively high number of track holes which reflected the expected spurious nature of their reconstruction. Other variables such as the χ^2 of track fits were checked however track hit and track hole multiplicities showed the best separation. As well it was



a)



b)

FIGURE 5.3: Number of tracks as a function of total number of track holes, nTotHoles, and total number of track hits, nSi, for non-tau tracks which are used to position second-stage RoIs. Results are shown for all such tracks in *a*) and for only tracks with $p_T > 100$ GeV in *b*).

seen that track hit and track hole multiplicities in only the pixel sub-detector or multiplicities for the combination of pixel and SCT sub-detectors, referred to as total multiplicities, showed good discrimination while multiplicities in only the SCT sub-detector were less discriminating. As such, track quality selections were trialled based upon selections on track hit and track hole multiplicities in the pixel sub-detector, $N_{\text{hits}}^{\text{pix}}$ and $N_{\text{holes}}^{\text{pix}}$ respectively, and in the combination of pixel and SCT sub-detectors, $N_{\text{hits}}^{\text{total}}$ and $N_{\text{holes}}^{\text{total}}$.

Selections were implemented by applying the relevant multiplicity requirements on the tracks chosen to position the second-stage RoI which were previously identified as tau tracks or non-tau tracks. The changes in the efficiency of correct second-stage RoI placement could then be calculated under the assumptions that removing a chosen tau track will result in an incorrectly placed second-stage RoI, while removing a non-tau track will result in a correctly placed second-stage RoI. The effects of three trialled selections on efficiency are shown in Table 5.1. Although the results are produced for a relatively small sample of RoIs (1294 first-stage RoIs), they are taken to be representative of the ability of quality criteria to remove non-tau tracks without a significant rejection of tau tracks that were previously chosen to seed second-stage RoIs – rejection of non-tau tracks is seen, and it is assumed that the statistical size of the dataset is large enough that rejection of tau tracks remains small within statistical uncertainties. In addition, because the same set of events is studied for each choice of quality criteria, any changes observed are purely systematic effects. Similar selections with tighter requirements or p_{T} dependent thresholds were also trialled however they were seen to give worse performance than those shown in Table 5.1.

It is seen that the selection of $N_{\text{hits}}^{\text{pix}} \geq 2$ AND $N_{\text{holes}}^{\text{tot}} \leq 2$ has the largest gain in efficiency due to the rejection of non-tau tracks while having only a small reduction in efficiency due to loss of tau tracks. As such this selection was chosen and was implemented in the online running of tau triggers within the inner detector trigger.

It is expected that the residual incorrect placement of second-stage RoIs due to non-tau tracks would dominantly be due to genuine non-tau tracks. A potential method to reject genuine non-tau tracks would be the spatial matching of the chosen track to the tau calorimeter cluster, however this was not attempted because the focus of this study was on what could be achieved from applying a track quality selection.

TABLE 5.1: Effects of track quality criteria on the efficiency of choosing a tau track when selecting a track to seed the second-stage RoI from the first-stage RoI track collection. ϵ is the percentage of first-stage RoIs where a tau track is present and is chosen, relative to the number of first-stage RoIs where a tau track is present. $\Delta\epsilon$ is the actual change in ϵ when the track quality criteria are applied. The 'lost' ('recovered') subscript refers to the changes seen when the quality critera are applied to tau tracks (non-tau tracks) which were previously chosen to seed the second-stage RoI. $(\Delta\epsilon)_{\text{combined}}$ is defined as $(\Delta\epsilon)_{\text{lost}} + (\Delta\epsilon)_{\text{recovered}}$. The efficiency after the track quality criteria are applied is shown in the final column.

Selection	$(\Delta\epsilon)_{\text{lost}}$	$(\Delta\epsilon)_{\text{recovered}}$	$(\Delta\epsilon)_{\text{combined}}$	ϵ
No selection	-	-	-	96.99%
$N_{\text{hits}}^{\text{tot}} \geq 8$ AND $N_{\text{holes}}^{\text{tot}} \leq 2$	-0.39%	1.47%	1.08%	98.07%
$N_{\text{hits}}^{\text{pix}} \geq 2$	0.00%	1.78%	1.78%	98.76%
$N_{\text{hits}}^{\text{pix}} \geq 2$ AND $N_{\text{holes}}^{\text{tot}} \leq 2$	-0.08%	2.24%	2.16%	99.15%

5.4 Understanding Reduced Track Finding Efficiency at Low Track p_{T} in First-Stage Regions of Interest

It was noticed in histograms which monitored the track finding efficiency relative to tracks found by the offline tracking algorithms that efficiency decreased at low track p_{T} for first-stage RoIs. The same behaviour was not observed for tracking efficiencies in either the fast track finder or precision tracking stages of second-stage RoIs. The behaviour can be seen in Figure 5.4 for efficiency defined with respect to all offline tracks within first-stage RoIs of the `tau25_idperf` trigger in 2015 LHC collision data [82]. This efficiency was defined using an online track to offline track matching criteria of $\sqrt{(\eta_{\text{online}} - \eta_{\text{offline}})^2 + (\phi_{\text{online}} - \phi_{\text{offline}})^2} < 0.05$. It is seen that first-stage tracking efficiency starts to drop between 10 GeV and 20 GeV and decreases increasingly rapidly as p_{T} decreases. The same effect was also observed when calculating efficiency with respect to tracks matched to truth tau tracks in simulated data. The loss of tau tracks suggested that the trigger efficiency to accept events containing taus could be affected if the second-stage RoI was positioned incorrectly due to the lack of one or more tau tracks at low p_{T}.

Due to the narrow $\Delta\phi$ width of first-stage RoIs it could be possible for low momentum tracks (which are significantly bent in the *r*-ϕ plane by the solenoidal magnetic field) which start within the RoI volume to curve out of the RoI instead of the full trajectory of the track being fully contained. As a result the track may not be reconstructed. To

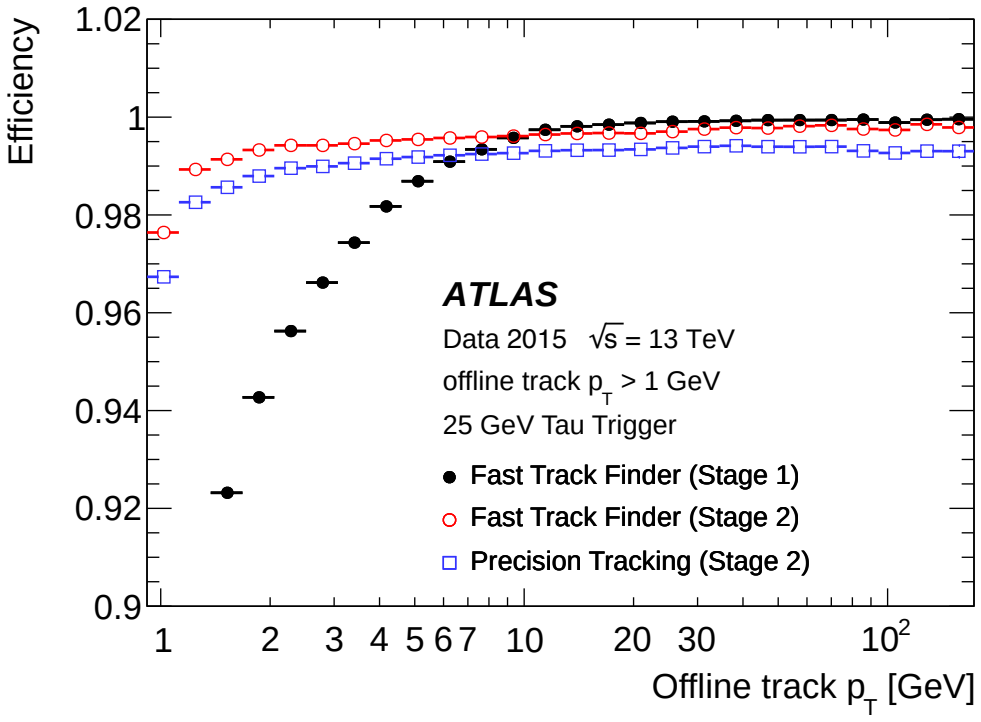
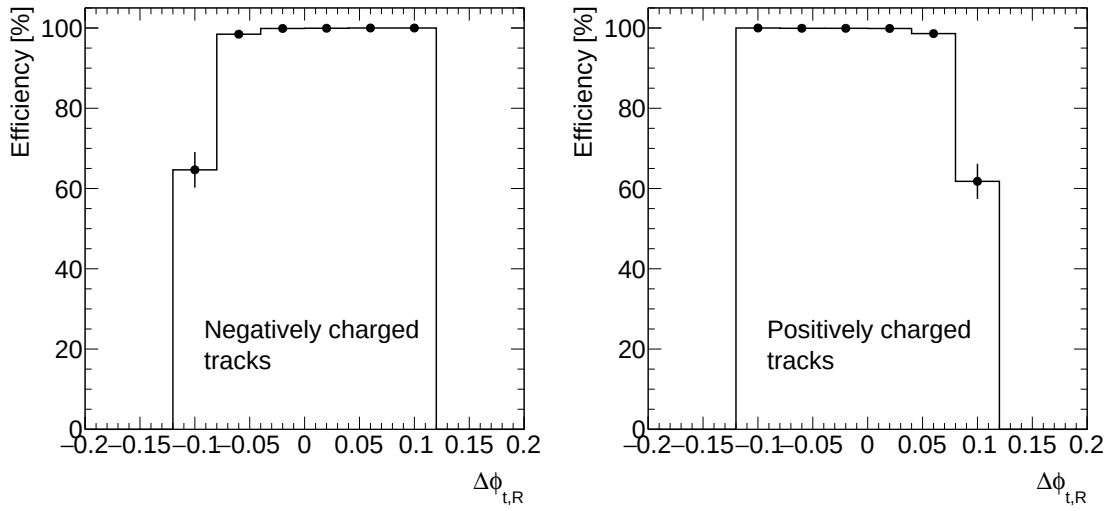


FIGURE 5.4: Track finding efficiency relative to tracks reconstructed by offline tracking algorithms in 2015 data as a function of offline track p_{T}. Bayesian statistical uncertainties are calculated but are too small to be seen. [82]

check for an indication that this hypothesis was correct the track finding efficiency as a function of $\Delta\phi_{\mathrm{t,R}} = \phi_{\mathrm{track}} - \phi_{\mathrm{RoI}}$ was observed separately for positively charged tracks and negatively charged tracks. This was done under the assumption that if tracks are bent out of the RoIs then tracks which start near the edge of an RoI $\Delta\phi$ width will either be bent into the RoI volume and will be more likely to be contained and reconstructed or will be bent out of the RoI almost immediately and not be reconstructed. This effect will be dependent upon the direction in which tracks are bent and as such will depend on the sign of track charge. The efficiencies were derived using a simulated dataset of 262,000 $Z \to \tau^{+}\tau^{-}$ events produced using the POWHEG matrix element generator [85–88] interfaced with the PYTHIA 8 parton shower generator [84], and simulated with the full ATLAS detector simulation. This dataset is used throughout the rest of this section. Tracks were studied if they were matched to an offline track using a matching criteria $\sqrt{(\eta_{\mathrm{online}} - \eta_{\mathrm{offline}})^2 + (\phi_{\mathrm{online}} - \phi_{\mathrm{offline}})^2} < 0.05$. The resulting efficiency histograms are shown in Figure 5.5. It is seen that significant decreases in tracking efficiency only occur on one side of the $\Delta\phi_{\mathrm{t,R}}$ distribution and that the distribution is approximately mirrored with the sign of track charge. This is consistent with the hypothesis that track loss at low



a) *b*)

FIGURE 5.5: Track finding efficiency in first-stage RoIs relative to tracks reconstructed by offline tracking algorithms as a function of $\Delta\phi$ from the centre of first-stage RoIs. The efficiency for negatively charged tracks is shown in *a*) and the efficiency for positively charged tracks is shown in *b*). Bayesian statistical uncertainties are calculated but are too small to be seen in most bins.

p_{T} is occurring due to tracks being bent out of the RoI.

A solution to the loss of tracks due to significant track curvature would be to increase $\Delta\phi$ to ensure that tracks are fully contained within RoI volumes. This was tested by running the `tau25_idperf` trigger on the simulated $Z \to \tau^+\tau^-$ dataset with a number of different $\Delta\phi$ values and observing the resulting track finding efficiency. To ensure the efficiencies could be compared between different choices of $\Delta\phi$ the offline tracks used in the denominator of the efficiency calculations were selected from within a fixed width of $\Delta\phi = \pm 0.4$ from the centre of first-stage RoIs rather than using the $\Delta\phi$ chosen for a given test. This width corresponds to the $\Delta\phi$ width of second-stage RoIs and therefore the upper limit of $\Delta\phi$ that would make sense to use for first-stage RoIs. An effect of this choice is that reference tracks outside of the $\Delta\phi$ width which is chosen for a given test will not be matched to an online track and this will contribute to an apparent reduction of tracking efficiency. The widening of $\Delta\phi$ therefore actually includes two effects: the recovery of tracks which are being bent out of the RoI and the inclusion of tracks which start outside of narrower $\Delta\phi$ values.

The track finding efficiency, relative to offline tracks which were matched to tau truth tracks using equivalent matching criteria to what was done for online to offline matching, as a function of the $\Delta\phi$ width used in first-stage RoIs is shown in Figure 5.6. Offline

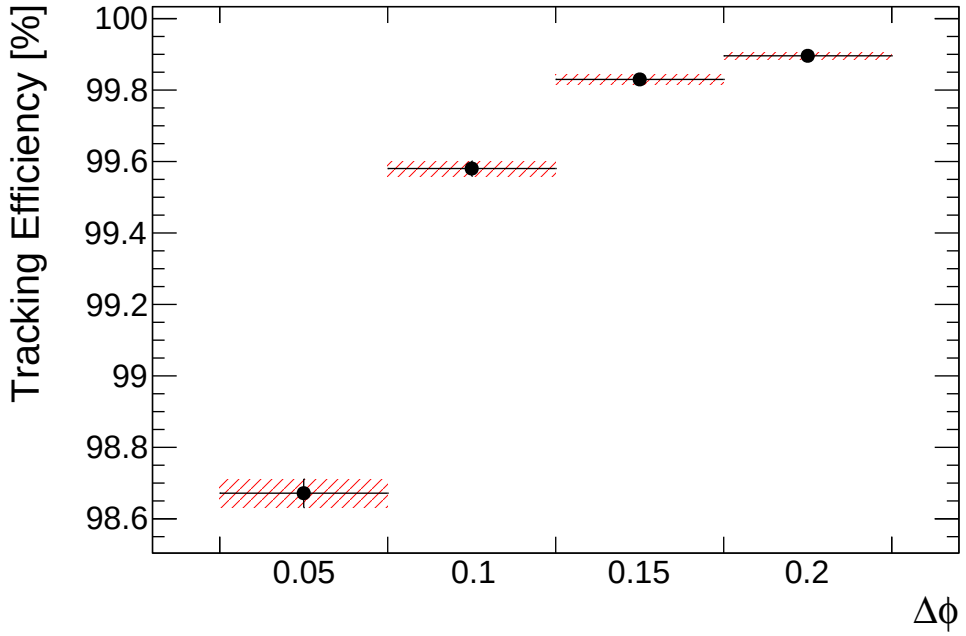
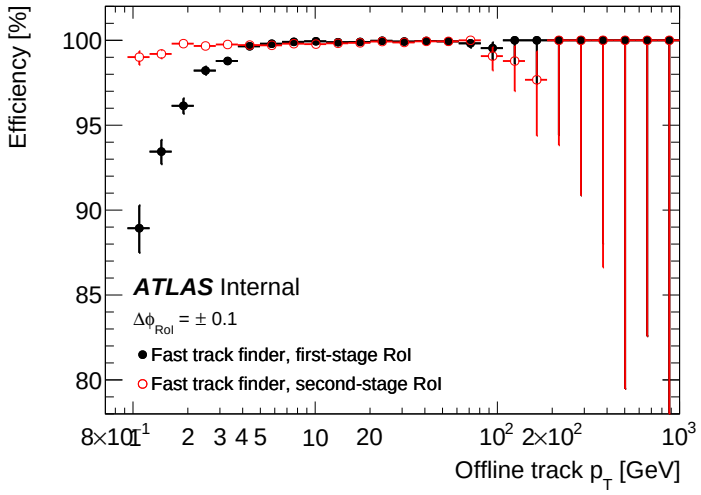


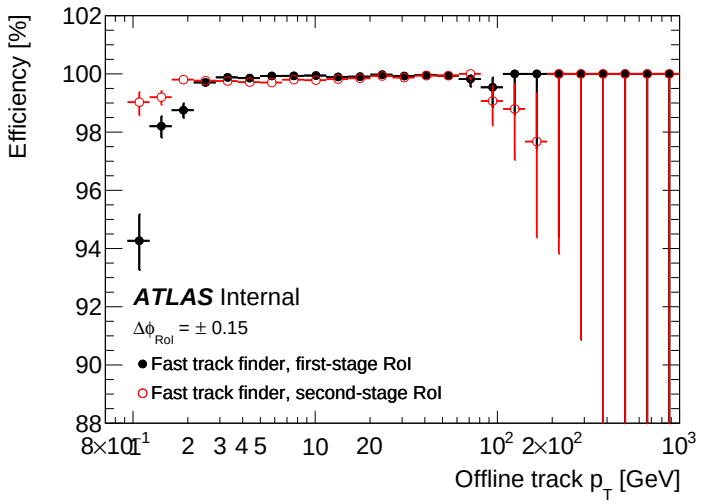
FIGURE 5.6: Track finding efficiency in first-stage RoIs relative to tracks reconstructed by offline tracking algorithms which are matched to truth tau tracks as a function of the $\Delta\phi$ half-width of the first-stage RoIs. Offline tracks used for the denominator of the efficiency calculation are selected in a width $\Delta\phi = \pm 0.4$ from the RoI centre to ensure that each efficiency is calculated with the same denominator. Bayesian statistical uncertainties are shown by the red hashed areas.

tracks were matched to truth tracks to ensure that only changes in efficiency with respect to tracks of interest were considered. It is seen that tracking efficiency increases as $\Delta\phi$ is increased beyond the initial value of $\Delta\phi = \pm 0.1$ and decreases relatively significantly for a narrower choice of $\Delta\phi = \pm 0.05$. Widths greater than $\Delta\phi = \pm 0.2$ were tested however no additional gain in tracking efficiency was seen. The tracking efficiency as a function of p_{T} is also shown for different $\Delta\phi$ choices in Figure 5.7. It is seen that the drop in efficiency below approximately 15 GeV improves as $\Delta\phi$ is increased. For $\Delta\phi = \pm 0.2$ the performance is equivalent to that of the fast track finder in the second-stage RoIs. This is consistent with the hypothesis that tracks are lost due to being bent out of RoI volumes. It is also observed in Figure 5.6 that tracking efficiencies are very high (greater than 99%) for $\Delta\phi \geq \pm 0.1$. As such the reduction in efficiency seen at low p_{T} only affects a very small fraction of the totality of tracks reconstructed by a 25 GeV tau trigger.

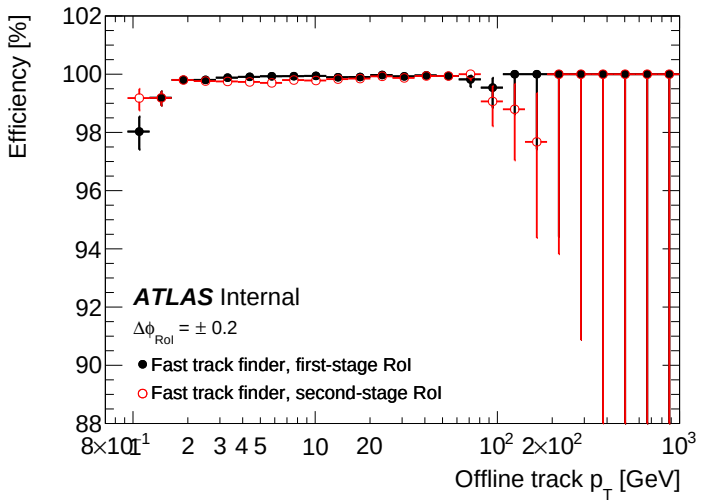
To understand the effect of the reduction in tracking efficiency at low p_{T} on the trigger efficiency to accept events containing taus, as well as how this effect changed as $\Delta\phi$ was varied, the efficiency of correctly positioning a second-stage RoI was calculated in a similar manner to what was done in Section 5.3. However the definition of the efficiency



a)



b)



c)

FIGURE 5.7: Track finding efficiency in first-stage RoIs relative to tracks reconstructed by offline tracking algorithms which are matched to truth tau tracks as a function p_{T} of the offline tracks, for different choices of RoI $\Delta\phi$ width. The efficiency for RoIs with $\Delta\phi = \pm 0.1$ is shown in *a*), the efficiency for RoIs with $\Delta\phi = \pm 0.15$ is shown in *b*), and the efficiency for RoIs with $\Delta\phi = \pm 0.2$ is shown in *c*). Offline tracks used for the denominator of the efficiency calculation are selected in a width $\Delta\phi = \pm 0.4$ from the RoI centre to ensure that each efficiency is calculated with the same denominator. Uncertainties shown are Bayesian statistical uncertainties.

was modified to consider only cases where a second-stage RoI would be mispositioned due to the lack of an online track which is associated with a truth tau track e.g. due to tracks being bent out of an RoI and not being reconstructed. In this way cases where a second-stage RoI would be mispositioned due to a non-tau track being chosen to seed the position of the RoI would be neglected. This was implemented by only considering first-stage RoIs where a tau truth track had the highest p_{T} (regardless of whether it was matched to an online track) or was matched to an online track which has the highest p_{T}. Online track to offline track matching was performed with the same requirement that is defined in Equation 5.2. Offline track to truth track matching was performed with a similar requirement to what is defined in Equation 5.3, except with the value used in the denominator of the η and ϕ terms changed to 0.1. In principle the track quality selection defined in Section 5.3 should have also been used when selecting the highest p_{T} track, and so the efficiencies may be slightly under or overestimated.

To understand the sources and magnitudes of inefficiencies better, RoIs were also categorised firstly by whether the truth tau decayed into one track or decayed into three tracks and secondly by whether the highest p_{T} tau track was above or below a low p_{T} threshold. The former categorisation reflects the difference in behaviour when lacking reconstructed tracks associated with truth tau tracks for one-track and three-track tau decays. If the single tau track in a one-track tau decay is not reconstructed then the second-stage RoI will clearly be positioned by a non-tau track and the event will likely not pass the tau trigger. However if the highest p_{T} tau track of a three-track tau decay is not reconstructed the second-stage RoI could still be correctly positioned as long as one of the other two reconstructed tau tracks is chosen. The latter categorisation defines whether any inefficiency observed originates only from tracks which are not reconstructed at low p_{T}. This categorisation also allows the observation of how significant any inefficiency in correctly positioning second-stage RoIs is in the low p_{T} range. The low p_{T} threshold was chosen as 5 GeV to allow observation of these effects in the region of significant track loss.

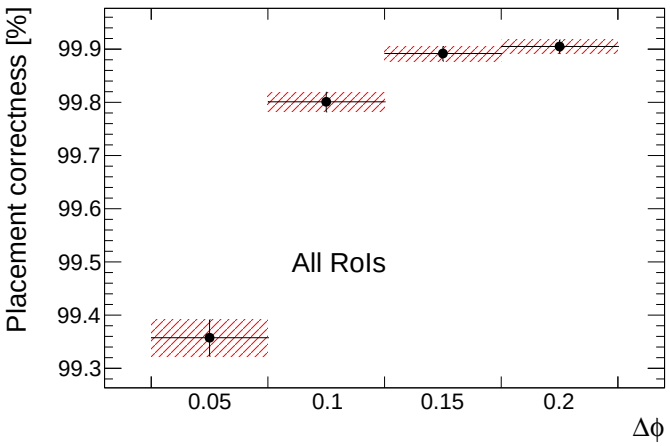
The efficiency to correctly position second-stage RoIs as a function of the $\Delta\phi$ width used in first-stage RoIs is shown in Figure 5.8 for all relevant RoIs, and also for the categorisation of RoIs based on the low p_{T} threshold. It is seen that the behaviour for all relevant RoIs is similar to the behaviour of tracking efficiency in Figure 5.6. However it is

also seen that the efficiency is very high at $\Delta\phi = \pm 0.1$ (approximately 99.8%) suggesting that a lack of reconstructed tracks associated with truth tau tracks does not occur often enough to significantly affect the correct placement of second-stage RoIs. It is seen from the efficiencies for RoIs categorised by the low p_{T} threshold that inefficiency is only significantly seen for RoIs in the category of tau tracks below the p_{T} threshold. The behaviour as a function of $\Delta\phi$ is also similar to what is seen for tracking efficiency in Figure 5.6. This suggests that second-stage RoIs are only incorrectly positioned due to low p_{T} tau tracks that are bent out of first-stage RoIs. Efficiencies for RoIs categorised by one-track and three-track tau decays were also observed and it was seen that the efficiencies for three-track decays were 100% for all $\Delta\phi$ widths. This showed that a reconstructed track associated with a truth tau track was always chosen to position the second-stage RoIs for three-track tau decays even if the highest p_{T} tau track was not reconstructed; equivalently this showed that efficiency reduction only occurred for one-track tau decays.

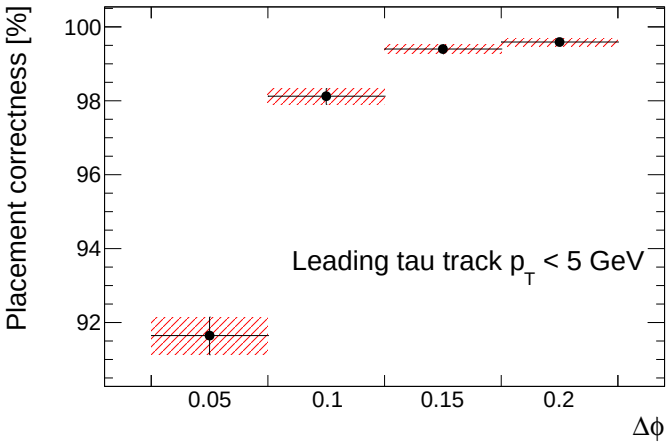
Computational times for the fast track finder in first-stage RoIs were also measured as a function of $\Delta\phi$. It was observed that mean execution time depended approximately linearly on $\Delta\phi$. Mean execution time for $\Delta\phi = \pm 0.1$ ($\Delta\phi = \pm 0.2$) was approximately 17 ms (28 ms). Given the high efficiency shown in Figure 5.6 for the default value of $\Delta\phi = \pm 0.1$ and the associated increase in execution time as $\Delta\phi$ is increased it was decided not to change the width of first-stage RoIs. However an update to the selection used to select tracks when measuring track properties and track finding efficiencies in the inner detector trigger as part of continuous performance monitoring was implemented: all tracks are required to have their full trajectory contained inside the volume of the associated trigger RoI. In this way the reduction in tracking efficiency from tracks bending out of first-stage RoIs would not mask other sources of track loss that may occur. This ensured that any changes or not-understood effects would be visible in continuous performance monitoring.

5.5 Parametrising Track Finding Methods Based On Track Seed Resolutions

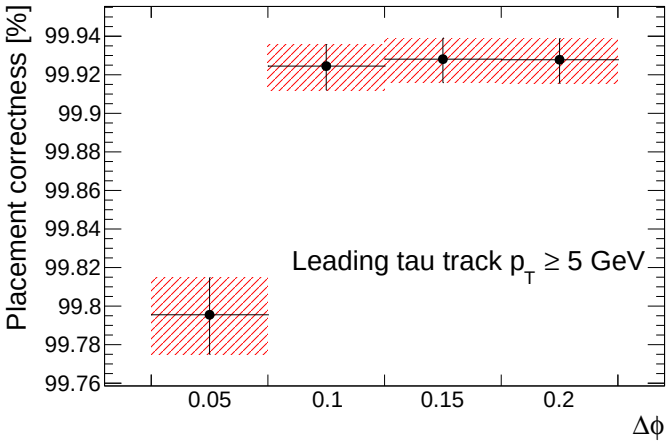
As previously mentioned the first step of track reconstruction prior to running the pattern recognition is the formation of track seeds. These seeds are formed of three space-points



a)



b)



c)

FIGURE 5.8: Efficiency to choose a tau track when selecting a track from the first-stage RoI track collection to seed the second-stage RoI, referred to as "placement correctness". The efficiency is defined relative to the number of first-stage RoIs where a tau truth track has the highest p_{T} or is matched to an online track which has the highest p_{T}. The efficiency for all relevant first-stage RoIs is shown in *a*), the efficiency for relevant RoIs with a leading tau truth track with $p_{\mathrm{T}} < 5$ GeV is shown in *b*), and the efficiency for relevant RoIs with a leading tau truth track with $p_{\mathrm{T}} \geq 5$ GeV is shown in *c*). Bayesian statistical uncertainties are shown by the red hashed areas.

and are referred to as 'triplets'. Triplets define the initial track guess which is then extrapolated into other detector layers to include additional hits and finally form an initial track candidate. As such it is important to ensure that track seeds which result in tracks of interest are successfully reconstructed. If they are not then track reconstruction cannot take place. To form the track seed-triplets an approach is used which forms combinations of a chosen detector hit with hits in other detector layers as 'doublets'. These doublets are categorised as either 'inner' or 'outer' type doublets depending on whether the second hit in the doublet is in a closer-in or further-out detector layer, respectively, than the first hit. For a chosen hit an inner-type and outer-type doublet are combined into a triplet with the chosen hit as the middle of the three hits.

An important consideration for the inner detector trigger is that the resolution of track seed-triplets and seed-doublets can be significantly worse than track resolutions. As such choice of the RoI sizes must be large enough to cover the spatial resolutions of seeds from tracks of interest as well as the spatial resolutions of the tracks themselves. The Δz width for the second-stage RoIs of $\Delta z = \pm 7$ mm was originally chosen to ensure that the z resolutions of tracks and track seeds for tracks near the centre of the RoI were covered. However this value was a rough estimate and also did not include consideration of how the spatial resolutions can vary, instead using a single inclusive width. If the z resolutions of track seeds are known then it is expected that the Δz width could be reduced while still including all previously covered track seeds for tracks of interest. As well, because it was expected that z resolution degrades as a function of pseudorapidity, the Δz widths could be parametrised as a function of $|\eta|$ such that narrower widths could be used for central pseudorapidities. In this way it was expected that computation time could be reduced compared to using a fixed width of ± 7 mm.

The Δz width is used within track seed creation to filter the collection of doublets formed from detector hits within an RoI. The direction of each doublet is extrapolated back to the beamline to determine a z for each doublet. If the z of a doublet is outside of the range $z_{\text{RoI}} \pm \Delta z$ then it is expected that the doublet will not result in a track that lies in the RoI and the doublet is rejected. Because this check is performed for each doublet it turn it is therefore possible to extend the parameterisation of the Δz widths to be done per-doublet as opposed to per-RoI. If z resolutions are known for hits in different layers of the detector then the Δz width can be defined based on the detector layers associated

with each doublet.

Measuring z resolutions of track seed-doublets

To measure the z resolutions of track seed-doublets, tracks were studied from the same 262,000 event POWHEG+PYTHIA 8 $Z \rightarrow \tau^+\tau^-$ dataset that was used in Section 5.4. The tracks which seeded the position of the second-stage RoI and therefore would be central in the second-stage RoI were selected. Doublets which had both hits also present in this track could then be selected during track seed creation. By measuring $\Delta z_{\mathrm{d,R}} = z_{\mathrm{doublet}} - z_{\mathrm{RoI}}$ for all such doublets it was possible to measure the z resolution of the track seed-doublets. Before the calculation of the z resolutions the collection of selected doublets was filtered further to ensure that doublets would not be considered if they were formed of pairings of track hits which would not be used to form the triplet that seeds the resulting track. This was done by looking at the ensemble of doublets formed for a chosen hit (i.e. the first hit in each doublet) and selecting only the inner and outer-type doublets which had the smallest $\Delta z_{\mathrm{d,R}}$. The remaining collection of doublets was then categorised based on inner- and outer-type doublets as well as which detector layer the first hit of the doublet was in. Each categorisation of doublets was then binned as a function of $|\eta|$ of the doublets and root mean square errors were calculated for the distribution of $\Delta z_{\mathrm{d,R}}$ in each bin such that the error values covered 95% of the data. These error values were taken to be the z resolution of the doublets. The resolutions could then be plotted as a function of $|\eta|$ and the resulting distributions could be fitted to produce a continuous parametrisation of doublet resolutions. Two fits were performed: a fit using a second-order polynomial function of $a|\eta|^2 + c$, and a fit using a fourth-order polynomial of $a|\eta|^2 + b|\eta|^4 + c$, where a, b, and c are parameters to be fit. These fit functions were chosen as simple guesses with parabolic dependence on $|\eta|$; if poor fit agreement was seen, more sophisticated fit functions could be employed.

Example resulting distributions and fits are shown in Figure 5.9 for doublets with the first hit in pixel sub-detector layers. Doublets with the first hit in SCT sub-detector layers are neglected for the remainder of this section because track seed-triplet creation only uses doublets with the first hit in pixel layers. It is seen that z resolutions are generally less than 1 mm with a small dependence on η. However the outer-type doublets for the

first endcap layer of the pixel sub-detector have resolutions of between 1 mm and 2 mm and show a larger increase as $|\eta|$ increases. Similar behaviour is seen for inner and outer-type doublets in other barrel and endcap layers, although the resolutions for outer-type doublets in the outer-most barrel layers rise to between 0.5 mm and 1 mm at $|\eta| \approx 2$. For some detector layers little difference was seen between the second-order and fourth-order polynomial fits. However in other layers, such as the second barrel layer of the pixel sub-detector for inner-type doublets seen in Figure 5.9, increases at high $|\eta|$ are described better for fourth-order polynomial fits. Because of this fourth-order polynomial fits were chosen when calculating resolutions for the parametrised track finding methods.

It can also be seen that the fits for outer type doublet-seeds in the second pixel barrel layer have poor agreement at very forward $|\eta|$ values. It was decided not to change the fit function because large multiples of the measured resolutions were used when implementing the parametrised methods, as described below. As a result, the calculated resolutions should fall in the tails of measured z residual distributions even if the fitted resolution is underestimated compared to the measured resolution. If the poor fit agreement was seen to have an impact on the performance of the parametrised methods, then polynomial fits could be produced in sub-ranges of $|\eta|$; for example fitting a second-order polynomial fit in the range $|\eta| < 2$ and a fourth-order polynomial fit in the range $2 < |\eta| < 2.5$.

Definitions and tests of parametrised track finding methods

Based upon these measured track seed-doublet resolutions two parametrised track finding methods were defined. The first, labelled as the 'parametrised doublet-seed filter' method, calculates per-doublet Δz widths to use when filtering doublets from the initial doublet collection. The Δz widths are calculated from the $|\eta|$ of the doublet and the detector layer that the first hit of the doublet is in. The second method, labelled as the 'parametrised second-stage RoI' method, parametrises the Δz of the second-stage RoIs based upon the widest expected doublet resolutions. The Δz widths are instead calculated as a function of $|\eta|$ of the second-stage RoI. This method also differed because the detector hits used to form doublets are selected based on their containment in RoIs, and therefore a smaller collection of hits will be used to form doublets compared to other

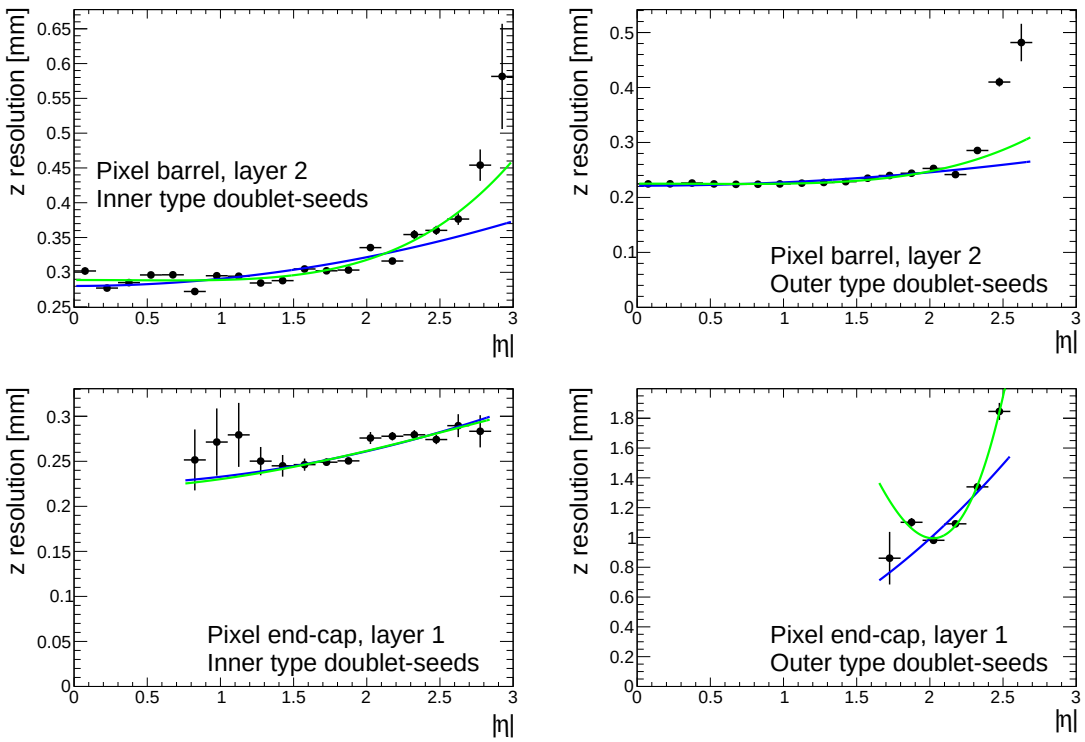


FIGURE 5.9: Fits to residuals in z between tracks and track doublet-seeds as a function of absolute pseudorapidity. Blue (green) lines are second-(fourth-)order polynomial fits. In the top row fits for the second barrel layer of the pixel sub-detector are shown. In the bottom row fits for the first endcap layer of the pixel sub-detector are shown. In the left (right) column fits for inner (outer) type doublet-seeds are shown. The $|\eta|$ ranges of data points follows from the $|\eta|$ coverage of combinations of hits in different layers.

methods. In both methods the parametrised Δz widths were limited as $\Delta z \leq 7$ mm such that any calculated widths which were larger than 7 mm were simply set to be 7 mm instead.

Tests were first performed using the parametrised doublet-seed filter method to check if the measured doublet resolutions were large enough to include doublets from all tracks from taus. These tests, as well as tests described below, unless stated otherwise, used the same POWHEG+PYTHIA 8 $Z \rightarrow \tau^+\tau^-$ dataset that was used to measure the doublet resolutions. The parametrised doublet-seed filter method was implemented using a range of multiples of the measured resolutions. The track finding efficiency relative to offline tracks which were matched to tau truth tracks was then compared for the use of the default fixed 7 mm width track finding and the use of the parametrised doublet-seed filter method with different multiples of the measured resolutions. Online tracks were matched to an offline track using a matching criteria $\sqrt{(\eta_{\text{online}} - \eta_{\text{offline}})^2 + (\phi_{\text{online}} - \phi_{\text{offline}})^2} < 0.05$, and offline tracks were matched to a truth track using the same criteria. This criteria is used for track matching in the remainder of this section. It was initially observed that for even large multiples of the resolutions, such as nine times the resolutions, track finding was noticeably lower by 1% to 2% in a number of bins for track parameter distributions. This effect was also more severe when calculating track finding efficiencies relative to all offline tracks. It was thought that due to the multiples of the resolutions ideally being large enough to cover all doublets from tracks which were central in the RoIs then this track loss instead came from tracks with a significant shift in z away from z_{RoI}. Because the doublet resolutions had been calculated assuming that tracks of interest were central in the RoIs then tracks which were away from z_{RoI} would naturally fall outside of the calculated Δz widths and would not be reconstructed.

Because of this, the evaluation of track reconstruction performance was also performed with the requirement that the offline tracks which were used as reference tracks were within $|\Delta z_{\text{t,R}}| \leq 2$ mm. This width represented the largest resolutions seen between online track z and associated offline track z, as well as containing most of the $\Delta z_{\text{t,R}}$ distribution except for long tails. Following this the track finding efficiency relative to offline tracks which were matched to tau truth tracks was again compared for the use of the

default fixed 7 mm width track finding and the use of the parametrised doublet-seed filter method with different multiples of the measured resolutions, however in this case it was done for only offline tracks within $|\Delta z_{t,R}| \leq 2$ mm. It was seen that for lower multiples of the resolutions, such as three times the resolutions, there was still a noticeable decrease in efficiency. However this effect reduced as increasingly larger multiples were used and performance was approximately the same between the non-parametrised and parametrised track finding methods when Δz widths of seven times the measured resolutions were used. No significant change was seen for larger multiples of the resolutions beyond this. As such the parametrised seed-doublet filter method was defined with Δz widths of seven times the measured resolutions.

It should be noted that although a large multiple of measured resolutions is used, the resulting Δz widths are in general still significantly smaller than the fixed $\Delta z = \pm 7$ mm width used the non-parametrised tracking. For inner-type doublets the measured resolutions are significantly less than 1 mm. For outer-type doublets, a multiple of seven times the measured resolutions can result in Δz widths that are similar in size or identical to the fixed $\Delta z = \pm 7$ mm width (bearing in mind that any calculated widths > 7 mm will simply be set to 7 mm); however, due to the $|\eta|$ dependence of the parametrised methods, widths smaller than 7 mm will still be calculated for central $|\eta|$, resulting in an improvement over the fixed Δz width approach.

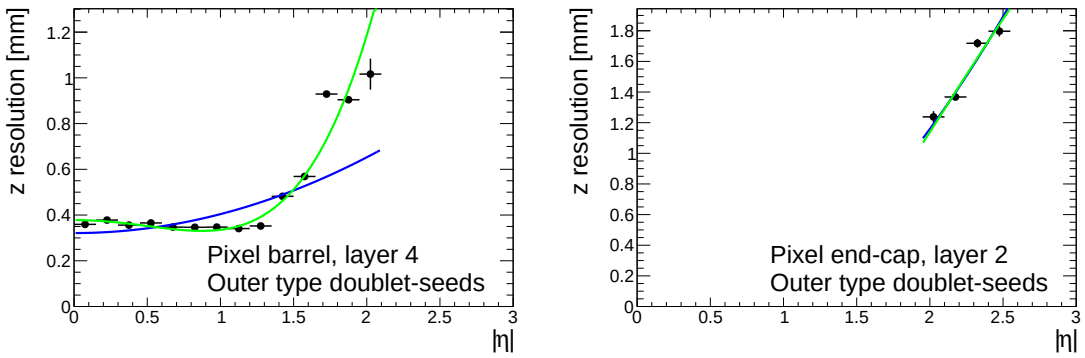
The need to use seven times the measured resolutions could arise from assumptions in the way seed-doublets were selected when measuring resolutions:

- Firstly, it was assumed that tau tracks were central in second-stage RoIs, and therefore $\Delta z_{d,R}$ would originate from the resolution of seed-doublets. Seed-doublets were therefore selected from the central track in second-stage RoIs. However, tau tracks are found with a non-negligible shift away from the centre of the RoIs, as can be seen from the need to require $|\Delta z_{t,R}| \leq 2$ mm when estimating tracking efficiency. The measured seed-doublet resolutions must therefore be inflated to cover the $\Delta z_{t,R}$ of tau tracks. This could originate from the second and third track in three-track tau decays being reconstructed away from the leading track according to detector spatial resolution; this could also originate from poor quality track reconstructions that result in large shifts away from the centre of the RoIs;

- Secondly, it was assumed that the seed-doublets with the smallest $\Delta z_{\mathrm{d,R}}$ would be used to construct the seed-triplets which actually seed tracks. Seed-doublets were therefore selected if they had the smallest $\Delta z_{\mathrm{d,R}}$ for hits in a given pair of detector layers. If track seed-triplets were instead formed from doublets with worse $\Delta z_{\mathrm{d,R}}$, then the measured resolutions would underestimate the actual resolutions.

Once the Δz widths were chosen for the parametrised seed-doublet filter method the widest expected widths could be observed as a function of $|\eta|$ and used to define the Δz widths used for the parametrised second-stage RoI method. By comparing the fits for inner- and outer-type doublets in different detector layers the combination with the largest resolutions could be chosen. It was seen that the largest resolutions occurred for outer-type doublets. Due to the limited $|\eta|$ ranges of the measured resolutions for outer-type doublets, the calculated largest expected resolutions were split between barrel and endcap regions and different layer resolutions were used to calculate the Δz widths for each region. The resolutions of outer-type doublets for the fourth barrel layer of the pixel sub-detector were chosen for $|\eta| < 2$ and the resolutions of outer-type doublets for the second endcap layer of the pixel sub-detector were chosen for $|\eta| \geq 2$; these resolutions were seen to be the largest in each $|\eta|$ region respectively. The fits associated with these choices are shown in Figure 5.10. The resolutions of outer-type doublets for the third endcap layer of the pixel sub-detector were not used due to a decrease in measured resolution for the final bin, resulting in an over-fitted fourth-order polynomial fit. The size of the resolutions were otherwise similar between the second and third endcap layer and so fits for the second layer were used instead.

The track finding efficiencies as a function of η, relative to offline tracks which were matched to tau truth tracks, are compared for the different track finding methods in Figure 5.11. These results were produced for a 4000 event subset of the POWHEG+PYTHIA 8 $Z \to \tau^+\tau^-$ dataset; this was done to allow for comparison and validation with standard internal ATLAS performance monitoring that used an identical collection of events. Results are shown for all offline tracks matched to tau tracks as well as only those which have $|\Delta z_{\mathrm{t,R}}| \leq 2$ mm. It is seen that for both methods increases in efficiency are seen in some bins while decreases are seen in others and the magnitude of the changes are similar between the two methods. For the efficiencies calculated with respect to all offline



a) *b*)

FIGURE 5.10: Fits to residuals in z between tracks and track doublet-seeds as a function of absolute pseudorapidity for the detector layers chosen for the parametrised second-stage RoI method. Blue (green) lines are second-(fourth-)order polynomial fits. The fits to outer-type doublets are shown for *a*) the fourth barrel layer of the pixel sub-detector and *b*) the second endcap layer of the pixel sub-detector. The $|\eta|$ ranges of data points follows from the $|\eta|$ coverage of combinations of hits in different layers.

tracks which are matched to tau truth tracks, the parametrised seed-doublet filter method is seen to reduce tracking efficiency at high $|\eta|$ values. The performance of the methods are much closer for efficiencies calculated with respect to tracks within $|\Delta z_{t,R}| \leq 2$ mm.

Changes to computation time when using parametrised track finding methods

To study the effects of using the parametrised track finding methods on computational time, the `tau25_idperf` trigger was run for a dataset of 3000 simulated $t\bar{t}$ events produced using POWHEG+PYTHIA 6 simulation packages and the full ATLAS detector simulation. Due to the hadronic nature of the targeted tau decays and the high rate of hadronic activity in p-p collisions, the dominant component of events processed by the tau triggers will be hadronic measurements from non-taus. As such the use of $t\bar{t}$ events which contain a number of high p_{T} hadronic jets as well as the presence of additional radiation allows the testing of the computational cost of track finding under strenuous conditions that are more similar to typical LHC collision events. As well, due to the increase in information to be processed as collision pile-up increases tests were performed using datasets with a range of pile-up values. Tests were performed by running computing jobs on similar computers within the Royal Holloway physics department's server farm. The computers in the server farm are constructed in a few different configurations:

- AMD Opteron 4284 CPU (8 cores, 3 GHz), 16 GB memory

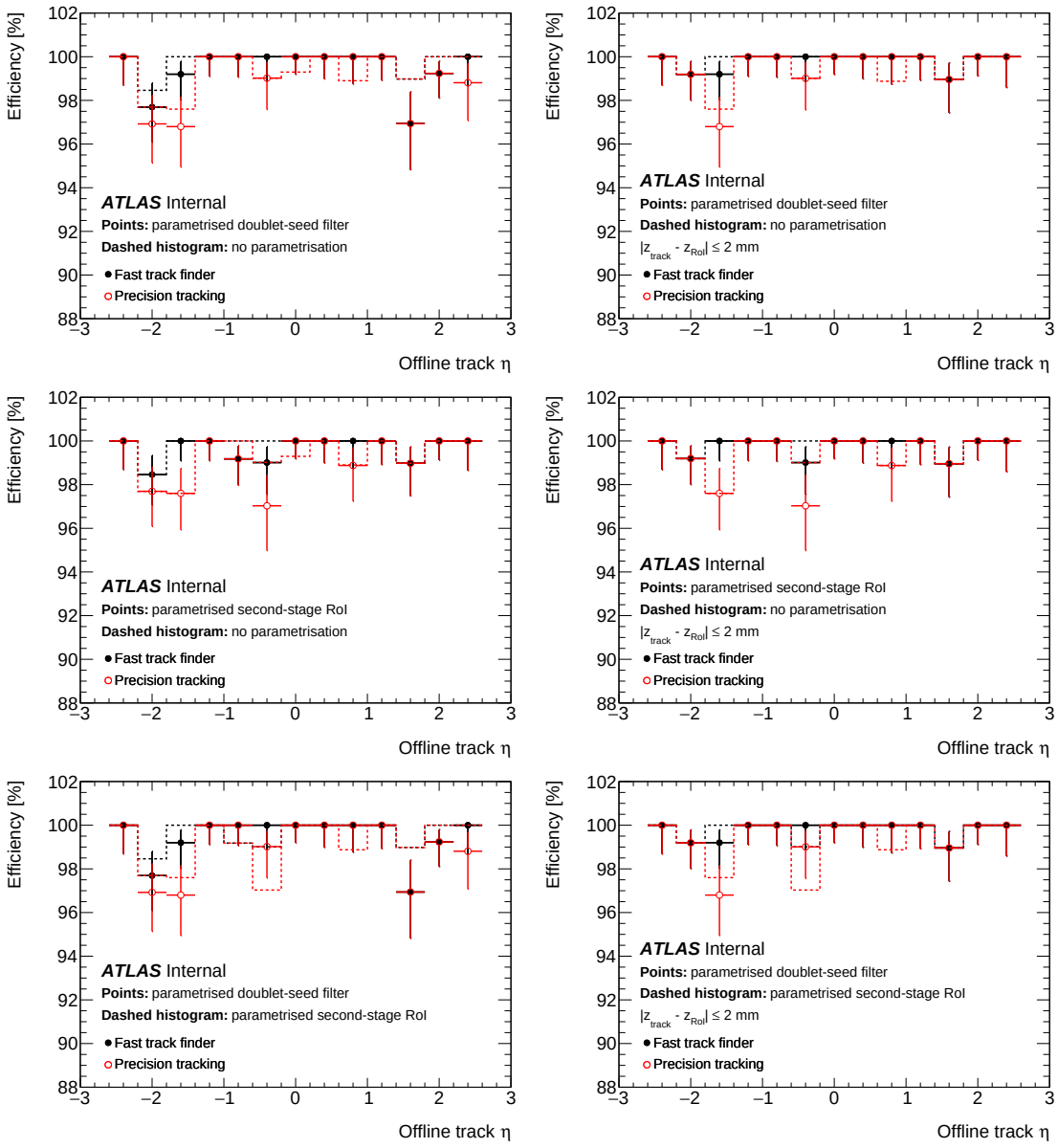


FIGURE 5.11: Track finding efficiency in second-stage RoIs relative to tracks reconstructed by offline tracking algorithms which are matched to truth tau tracks as a function of pseudorapidity of the offline tracks. In the left (right) column all relevant offline tracks (offline tracks within $|\Delta z| < 2$ mm of the RoI centre) are used when defining tracking efficiency. In the top row: points show results from the parametrised doublet-seed filter method, and dashed histograms show results from using no parametrisation. In the middle row: points show results from the parametrised second-stage RoI method, and dashed histograms show results from using no parametrisation. In the bottom row: points show results from the parametrised doublet-seed filter method, and dashed histograms show results from the parametrised second-stage RoI method. Uncertainties shown are Bayesian statistical uncertainties.

- AMD Phenom II X4 955 CPU (4 cores, 3.2 GHz), 8 GB memory
- AMD Phenom II X6 1090T CPU (6 cores, 3.2 GHz), 8 GB memory
- Intel Xenon E5345 CPU (8 cores, 2.33 GHz), 16 GB memory

Attempts were made to run jobs under similar computational loads where possible. To try to ensure that resulting measured execution times were comparable, the mean total execution times of the fast track finder in first-stage RoIs were scaled to the same value (chosen as the execution time for using no parametrisation). Due to the parametrised track finding methods only changing track finding in second-stage RoIs, the execution times for the fast track finder in first-stage RoIs should not vary.

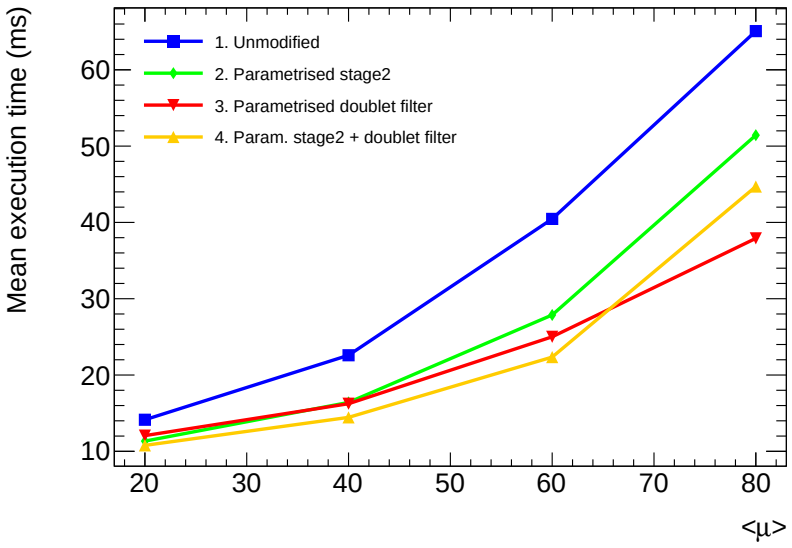
The resulting mean total execution times per-RoI for the fast track finder and for precision tracking in second-stage RoIs are shown in Figure 5.12 as a function of mean pile-up in the $t\bar{t}$ datasets. Execution times were also tested for the combination of using the parametrised seed-doublet filter method as well as the parametrised second-stage RoI method. The track finding efficiency was seen to be similar to the use of the individual parametrised methods. It is seen that significant reductions in the mean fast track finder execution time are present for parametrised methods compared to using no parametrisation. For an average pile-up of 60 the mean execution time is reduced by approximately 35% to 50%, depending on the parametrisation method used. The parametrised seed-doublet filter method is seen to give a significantly faster mean execution time at an average pile-up of 80. Because the combination of the parametrised seed-doublet filter method with the parametrised second-stage RoI method represents a reduction in the amount of information to process compared to just the parametrised seed-doublet filter method alone this is not expected. The reason for this could be a difference in the computational load from unrelated processes, or due to jobs running on different CPUs compared to other timing results. To avoid these possibilities, it would have been desirable to re-run jobs on servers that were selected to have the same architecture, and ensure that no other processes were run on these computers at the same time. With this unexpected feature noted, it is seen that differences between fast track finder mean execution times are smaller or negligible when comparing different parametrisation methods. For an average pile-up of 60, differences between methods are approximately 3 ms to 6 ms.

Due to an error in running the computing jobs the execution times for the precision tracking were considered for both the older single-stage tracking method and for second-stage RoIs in the two-stage tracking method. As a result the mean execution times were calculated for the combination of the distributions from each method. This arose due to how the precision tracking algorithm is technically treated when defining triggers in the computing jobs. For the fast track finder algorithm individual instances are created for each type of RoI to be run in (e.g. separately for first-stage and second-stage RoIs); however a single precision tracking algorithm instance is created and then used for all triggers requiring the running of precision tracking. Because the computing jobs also included a `tau25_idperf` trigger which used the older single-stage tracking method, the final recorded execution times included times from both the single-stage RoIs and the second-stage of two-stage RoIs. Because both versions of the `tau25_idperf` trigger should be seeded from the same L1 information, the number of RoIs, and therefore contribution to the timing distribution, should be equal.

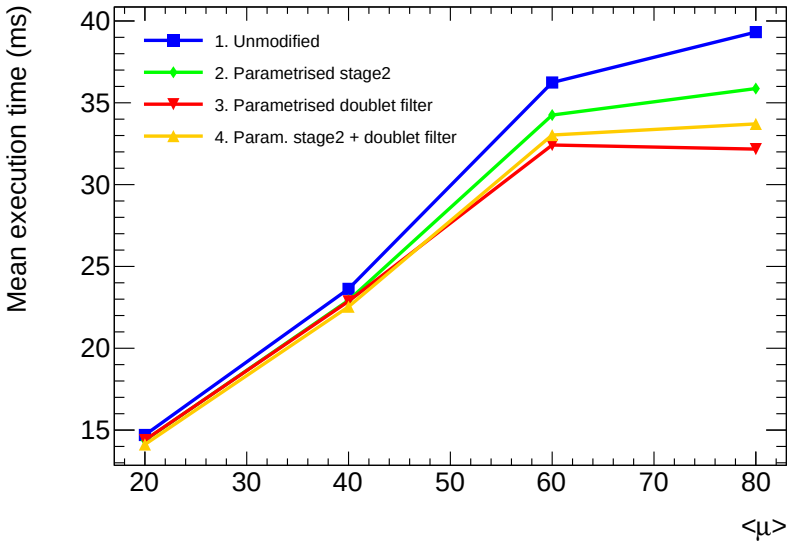
With this caveat noted the differences between different methods in Figure 5.12 are still indicative of changes in precision tracking execution time from the choice of method due to the fact that the single-stage tracking method will be unaffected by changes to the two-stage tracking method. It can be seen that the relative changes between using no parametrisation and using parametrised methods are significantly smaller than for the fast track finder mean execution times.

Final tests of parametrised track finding methods

Following from these tests in track finding efficiency and computation time, the parametrised methods were implemented in common ATLAS software and results were checked for standard internal ATLAS ID trigger performance monitoring tests. The parametrised seed-doublet filter method was chosen as the new baseline method intended for use in ATLAS data taking. The combination of the parametrised seed-doublet filter method with the parametrised second-stage RoI method was not tested further due to the lack of significant reduction in computational timing as well as ensuring that tracking efficiency would not be lost prior to the filtering of reconstructed doublets. The standard ATLAS tests performed similar tests of track finding efficiency and computation



a)



b)

FIGURE 5.12: Mean execution time per-RoI of track finding algorithms as a function of pile-up in simulated $t\bar{t}$ datasets. Execution times are shown for different track seed parametrisation methods. "Parametrised stage2" labels the parametrised second-stage RoI method. "Param. stage2 + doublet filter" labels the combined use of the parametrised second-stage RoI method and parametrised doublet-seed filter method. Computation times are shown for *a*) the fast track finder and *b*) precision tracking. Due to an error in processing the mean precision tracking times were calculated for the combined distribution of execution times for the single-stage tracking method and for second-stage RoIs in the two-stage tracking method.

time using the same `tau25_idperf` trigger for similar simulated datasets. Similar track finding efficiency performance was seen as in Figure 5.11.

The parametrised track finding methods were ultimately judged against an original goal of developing methods based on measurements of track seeds: that Δz widths could be reduced while still including all previously covered track seeds for tracks of interest, and therefore reconstructing all previously reconstructed tau tracks. Because decreases in tracking efficiency for tau tracks were seen when employing the developed parametrised methods, particularly for tracks with significant Δz shifts away from the centre of second-stage RoIs, it was judged that this goal had not been met. Furthermore, at the time these methods were developed there was not a specific need to reduce computational timing that could have merited the reduction in tracking efficiency. As a result the parametrised methods were not included in ATLAS data taking. However the methods were fully implemented in common ATLAS software and can be used in future if collision conditions become such as to warrant an improvement in computation time for a small tradeoff in track finding efficiency.

Chapter 6

Search for the Standard Model Higgs Boson Produced in Association With Dileptonically Decaying Top Quarks and Decaying Into a $b\bar{b}$ Pair

6.1 Introduction

The ATLAS Collaboration performs a number of searches for the $t\bar{t}H$ production mechanism, where each search is distinguished by the final state particles of the decaying Higgs boson. Of these final states, the decay of a Higgs boson into a $b\bar{b}$ pair has a significantly higher branching ratio than other Higgs decays. Assuming a Higgs mass of $m_H = 125$ GeV the branching ratio for $H \to b\bar{b}$ is approximately 58%, while the second highest branching ratio is approximately 21% for the $H \to W^+W^-$ decay mode [31]. Higgs branching ratios are shown in Figure 6.1. When considered along with the low cross-section of $t\bar{t}H$ production, this makes the $t\bar{t}H$ ($H \to b\bar{b}$) process, labelled as $t\bar{t}H(b\bar{b})$, an attractive target for searches. The ATLAS analysis searching for this process is distinguished further by the decay products of the $t\bar{t}$ pair. Top quarks dominantly decay as $t \to W^+ b$; the W boson can then decay either hadronically, $W^+ \to q\bar{q}$, with a branching fraction of approximately 67%; or leptonically, $W^+ \to l^+\nu_l$, with a branching fraction of approximately 33% [18]. This therefore defines three possible final states for the $t\bar{t}$ pair: all-hadronic, $t\bar{t} \to q\bar{q}\,q\bar{q}$; semi-leptonic, $t\bar{t} \to l^+\nu_l\,q\bar{q}$ (and the charge conjugate of this process); and dileptonic, $t\bar{t} \to l^+\nu_l\,l^-\bar{\nu}_l$. The all-hadronic final state provides the highest

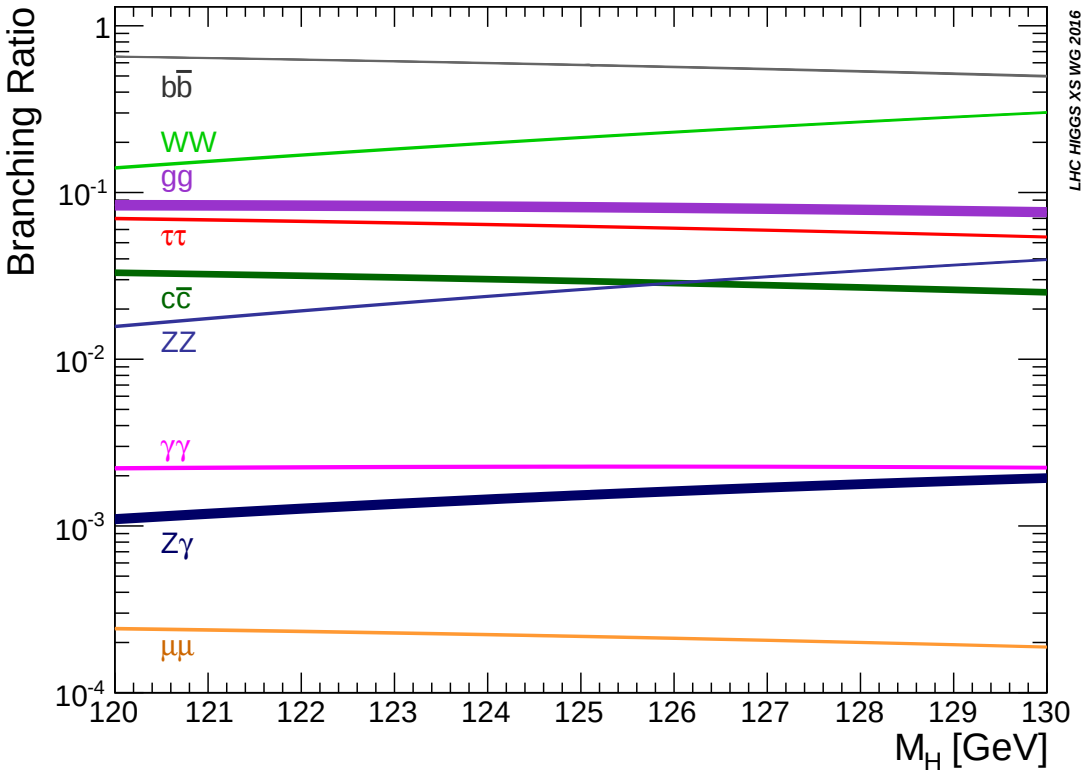


FIGURE 6.1: Higgs boson decay branching ratios as a function of Higgs boson mass [31].

branching ratio; however when combined with the $b\bar{b}$ pair of the Higgs decay, it is very difficult to distinguish the all-hadronic $t\bar{t}H(b\bar{b})$ process from QCD background processes, which are difficult to model and many orders of magnitude larger than the $t\bar{t}H$ signal. As such, the ATLAS analysis featured in this thesis focuses on the more distinct leptonic final states, defining the single-lepton and dilepton analysis channels; here 'lepton' refers only to electrons and muons (this definition includes electrons and muons that originate from the decay of a tau which was produced from the decay of a W boson). For convenience the analysis is referred to simply as the ATLAS $t\bar{t}H(b\bar{b})$ analysis even though the all-hadronic channel is not included. The authors work documented in this thesis largely contributes to the dilepton analysis channel, and so focus is placed here except for comparisons or combinations with the single-lepton channel. The dilepton channel has a lower branching ratio than the single-lepton channel, but has negligibly sensitivity to QCD backgrounds. An example Feynman diagram for the dileptonic $t\bar{t}H(b\bar{b})$ process is shown in Figure 6.2.

The ATLAS $t\bar{t}H(b\bar{b})$ analysis was performed using events from $\sqrt{s} = 13$ TeV LHC

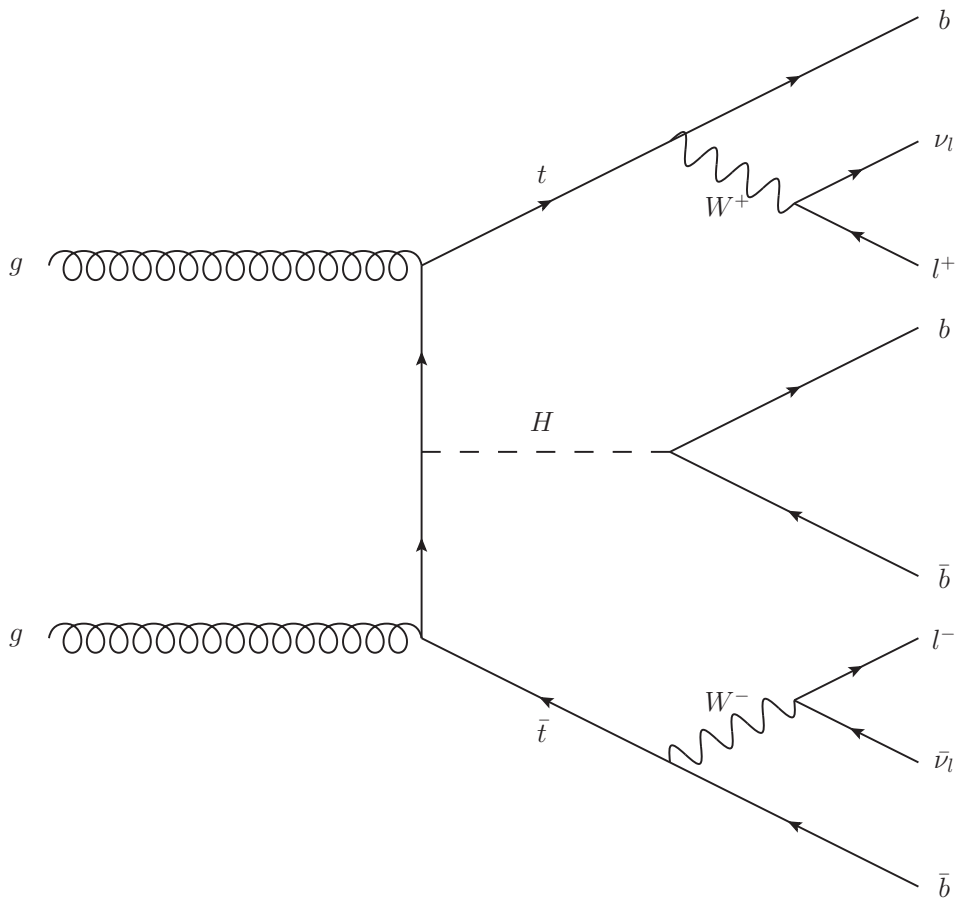


FIGURE 6.2: Example Feynman diagram for the dileptonically decaying $t\bar{t}H(b\bar{b})$ process. l indicates either an electron or a muon.

p-p collisions during 2015 and 2016 data taking. The work presented in this thesis contributed to analyses using 13.2 fb^{-1} [89] and 36.1 fb^{-1} [90] integrated luminosity of 13 TeV data, where both datasets use data from both 2015 and 2016. However final results are presented only from the 36.1 fb^{-1} analysis. After preliminary public release of the 13.2 fb^{-1} analysis an error was found by the $t\bar{t}H(b\bar{b})$ analysis group which affected the final measured $t\bar{t}H$ signal strength in the dilepton channel. This error occurred in the evaluation of the machine-learning-based classification BDT algorithm described below in Section 6.3. This algorithm is intended to be trained to separate signal and background processes by learning features of these processes in half of the total number of simulated collision events before being used to identify the remainder of simulated events as signal-like or background-like. Optimisation of the classification BDT using this split into a 'train' and 'test' dataset ensures that the algorithm is not overfitted to the exact details of the training dataset. However when the output of the classification BDT was evaluated as part of software developed by the ATLAS $t\bar{t}H(b\bar{b})$ analysis group, an error in computer file naming meant that instead a version of the algorithm which was trained on the entire collection of simulated events was used. The result was event classification scores that were overfitted to the simulated data being tested, resulting in a biased score that appeared to improve signal and background separation.

When the classification BDT with the correct file-name was used the sensitivity to $t\bar{t}H$ signal strength was reduced by approximately 20%. Due to the presence of this error the publicly released 13.2 fb^{-1} results were retracted. The 36.1 fb^{-1} $t\bar{t}H(b\bar{b})$ analysis supersedes the 13.2 fb^{-1} analysis, with many similarities in the definitions and methods used in each analysis. As well, most of the work presented in this thesis contributed to the 36.1 fb^{-1} analysis. Focus in this chapter is therefore placed on the 36.1 fb^{-1} analysis. Descriptions of the $t\bar{t}H(b\bar{b})$ analysis and the author's contributions can be assumed to be for the 36.1 fb^{-1} analysis unless otherwise stated.

It should be noted that the error in the classification BDT is also present in results presented in this chapter which make use of classification BDTs. The same use of an incorrect file-name was also present in the 36.1 fb^{-1} analysis until this error was noticed during editorial review of the analysis prior to publication. The issue was corrected for in the final published results however this occurred after the work presented in this chapter had been completed. While it would be possible to re-run the studies described here with

the corrected BDT, a main aim of Section 6.7 is to present the developmental process of refining, improving, and understanding the statistical fit model. Therefore, results are presented according to this chronological sequence (i.e. prior to the correction of the classification BDT), so as to for example motivate the reasoning behind changes to the fit model. It is expected that using the corrected BDT would dominantly affect sensitivity to signal, while overall agreement between predicted yields from simulation and measured data yields may also show some smaller changes.

6.2 Analysis Challenges

The search for the $t\bar{t}H(b\bar{b})$ process involves a number of highly challenging aspects that require detailed consideration and advanced analysis methods to overcome. The ability of the ATLAS $t\bar{t}H(b\bar{b})$ analysis to do so is a testament to the effectiveness of the analysis methods employed, the continual improvement of predictions from theory and simulation, and the dedication of the physicists who drive the development and improvement of the analysis as part of the $t\bar{t}H(b\bar{b})$ analysis group. A summary of the challenges faces is given here to motivate the need for, and highlight the effectiveness of, the strategies described in Section 6.3.

Many of these challenges involve the dominant background process of $t\bar{t}$+jets. An example Feynman diagram for this process is shown in Figure 6.3. This process is in many ways similar to $t\bar{t}H(b\bar{b})$; in essence $t\bar{t}H(b\bar{b})$ is $t\bar{t}$ pair production with the addition of a radiated Higgs decaying into a $b\bar{b}$ pair, while $t\bar{t}$+jets is $t\bar{t}$ pair production with the addition of QCD or electroweak radiation producing additional final state quarks. In particular the $t\bar{t}$+$b\bar{b}$ process is an irreducible background due to producing the same final state as $t\bar{t}H(b\bar{b})$. This therefore makes separating $t\bar{t}H(b\bar{b})$ events from $t\bar{t}$+jets events within the analysis very challenging, and indeed no single variable has strong separating power with respect to $t\bar{t}$+$b\bar{b}$. On top of this the cross-section for producing $t\bar{t}$+jets is much larger than for producing $t\bar{t}H(b\bar{b})$, with the cross-section for $t\bar{t}$+$b\bar{b}$ of the order of 10 times larger than $t\bar{t}H(b\bar{b})$. Along with this the theoretical modelling of $t\bar{t}$ with additional jets is challenging, particularly when considering additional c- or b-jets. For convenience $t\bar{t}$+jets processes containing additional heavy-flavour jets are labelled $t\bar{t}$+HF, while $t\bar{t}$+light refers to processes with no heavy-flavour jets. This difficulty in modelling

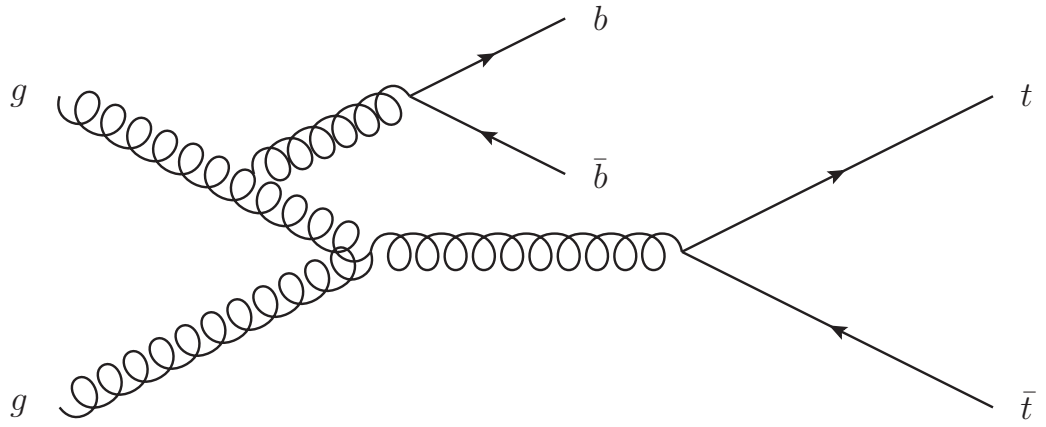


FIGURE 6.3: Example Feynman diagram for the $t\bar{t}$+jets process. In this example the additional jets are produced by a $b\bar{b}$ pair, an example of the $t\bar{t}$+$b\bar{b}$ process.

can result in large differences between predictions from event simulation packages which can be considered as large $t\bar{t}$+jets modelling uncertainties. As well the limited nature of modelling $t\bar{t}$+HF results in a limited ability to calculate the correct total cross-section and therefore the overall yield of these processes.

To add to these challenges stemming from theory, challenges also arise in the accurate and precise reconstruction of b-jets. Accurately calibrating simulated data to match real data for b-tagging efficiencies as well as rates of mis-identification of light- and c-jets as b-tags is a highly challenging process (performed by a dedicated ATLAS working group as opposed to performed within this analysis), particularly if utilising pseudo-continuous b-tagging. This brings associated b-tagging uncertainties which can play a large role due to the presence of a number of b quarks in the signal final state. The result of the much larger background contribution, difficulty separating signal from background, large uncertainties on the dominant background, and b-tagging uncertainties for the presence of a number of jets is that *a priori* the signal yield in an individual distribution is entirely covered by uncertainties on the background yield. As a result the analysis is considered to be systematically rather than statistically limited. Difficulties also arise simply from the number of final-state objects; the efficiency to reconstruct all final-state objects can be low, and it is hard to correctly choose the combinatorial groupings of reconstructed objects to match their parent particle. To further add to these theoretical and experimental challenges the analysis also has large computational needs. Due to the phase space requirements of producing additional high p_{T} heavy-flavour jets as well as requiring $t\bar{t}$+jets

events to decay leptonically, generating large enough simulated event samples given limited computing budgets is difficult. In particular the ATLAS $t\bar{t}H(b\bar{b})$ analysis makes use of a number of alternative simulated datasets for theoretical uncertainty estimates and validation of the analysis, further adding to the computational resources needed. The result of these computational challenges is that simulated datasets are in some cases limited by the number of events that have been generated, which results in additional uncertainties due to statistical uncertainty arising from the limited simulated sample size.

6.3 Analysis Strategy

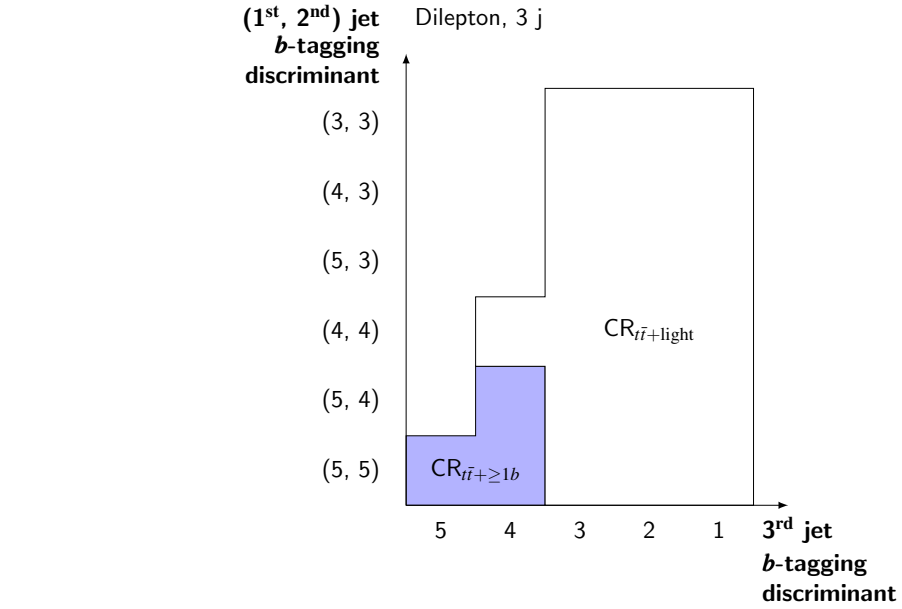
The single-lepton and dilepton channels are independently analysed using orthogonal event selections. Similar strategies are used in the analyses targeted at each of these channels, although they are developed and optimised individually. In both cases, the analysis first applies appropriate event selection criteria before categorising events into analysis regions based on the number of jets in each event as well as the number of b-tags and their quality. For the 13.2 fb^{-1} analysis these regions were defined in terms of the jet multiplicity and the multiplicity of b-tags at a fixed b-tagging working point; for example defining one analysis region with the requirement $N_{\text{jets}} \geq 4$ AND $N_{b\text{-tags}}^{\text{tight}} = 3$. This method is effective at categorising events into high or low signal purity regions, referred to as signal and control regions respectively, while at the same time categorising background $t\bar{t}$+jets events based on the flavour of the additional jets. The 36.1 fb^{-1} analysis goes a step further by using b-tag multiplicities for b-tags which are not all at the same working point (pseudo-continuous b-tagging). This allows more detailed control of region definitions with respect to background composition, specifically with respect to the flavour of additional jets in $t\bar{t}$+jets events. The definitions for the regions used in the dilepton analysis are shown in Figure 6.4. The numeric labels in this figure represent the different b-tagging working points of the MV2c10 b-tagger defined in Table 4.1 from Section 4.2.5: the range "5" to "2" labels the range of "very tight" to "loose" working points respectively, and "1" labels jets that do not pass any b-tagging working point. It can be seen that the control regions are labelled with a subscript which represents the dominant composition of additional jets in $t\bar{t}$+jets events in these regions. The relative proportions

of background processes in these regions are shown in Figure 6.5 and the signal purity and signal significance in these regions are shown in Figure 6.6.

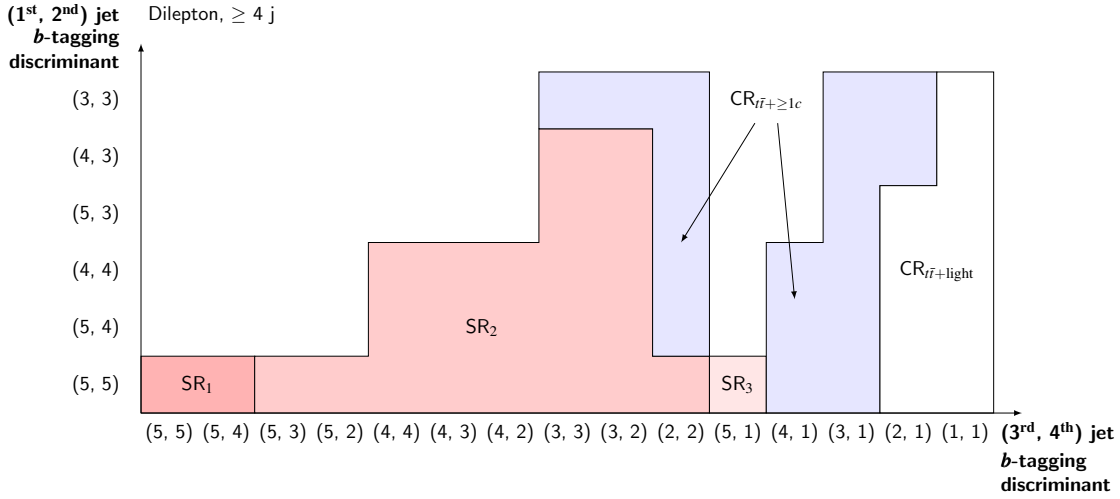
Boosted Decision Trees (BDTs) are employed in signal regions to further separate signal events from background events [91] (and references therein). BDTs are a class of multivariate analysis (MVA) algorithms which define decision trees consisting of sequential cuts on relevant variables, where the output of the decision tree is a classification score which defines how similar an input is to a category of interest; for example whether the input is similar to a signal process. Boosting is a learning process where many decision trees are iteratively defined, where each iteration checks the classification accuracy of the previous iteration, and then updates tree structures to minimise incorrect classification and tree complexity. The ensemble of trees from all iterations are then summed to provide a more accurate classification score.

In the $t\bar{t}H(b\bar{b})$ analysis BDTs are used in two ways: classification BDTs aim to separate events as being signal- or background-like, and reconstruction BDTs aim to identify the correct combination of jets and leptons into Higgs boson and top quark candidates. The Higgs boson candidate is formed from a pair of jets, and in the dilepton channel the top quark candidates are formed from the pairing of a lepton with a jet. The jet and lepton combinations identified by the reconstruction BDTs are used to define input variables for the classification BDT which themselves have good signal to background separation. The invariant mass of the jet pair that makes up the Higgs boson candidate is an example of a variable with good separation that depends on which specific jet combination has been chosen. In general the classification BDT is formed of input variables that consider the kinematic properties of pairs of objects (including the candidate object systems defined by the reconstruction BDT), as well as some event-level variables such as the number of jets with $p_{\mathrm{T}} > 40$ GeV. An example of an object pair variable that is highly ranked in the list of inputs is the minimum invariant mass of a pair b-tags. The output of the classification BDT is then used as the final discriminant of how signal-like an event is in the signal regions. The full list of input variables used in each analysis region can be seen in Appendix B of the 36.1 fb^{-1} $t\bar{t}H(b\bar{b})$ paper [90].

A statistical analysis is then performed: a binned profile likelihood fit is constructed considering yields in all analysis regions, and encoding systematic uncertainties as nuisance parameters. The classification BDT outputs are used as the distributions which are



a)



b)

FIGURE 6.4: Analysis region definitions in the dilepton channel for *a)* exactly 3 jets and *b)* greater than or equal to 4 jets. The *y*-axis labels the *b*-tag working points for the first two jets, while the *x*-axis labels the *b*-tag working points for *a)* the third jet or *b)* the third and fourth jet. The numeric labels represent the different *b*-tagging working points of the MV2c10 *b*-tagger defined in Table 4.1 from Section 4.2.5: the range "5" to "2" labels the range of "very tight" to "loose" working points respectively, and "1" labels jets that do not pass any *b*-tagging working point. The jets are ordered by *b*-tagging discriminant score in descending order. Regions shaded in red are signal enriched regions, regions shaded in blue are $t\bar{t}$+HF enriched regions, and regions shaded in white are $t\bar{t}$+light enriched regions. $t\bar{t}+1b$ refers to events containing a single additional *b*-jet, and $t\bar{t}+\geq 1c$ refers to events containing at least one *c*-jet and no additional *b*-jets. [90]

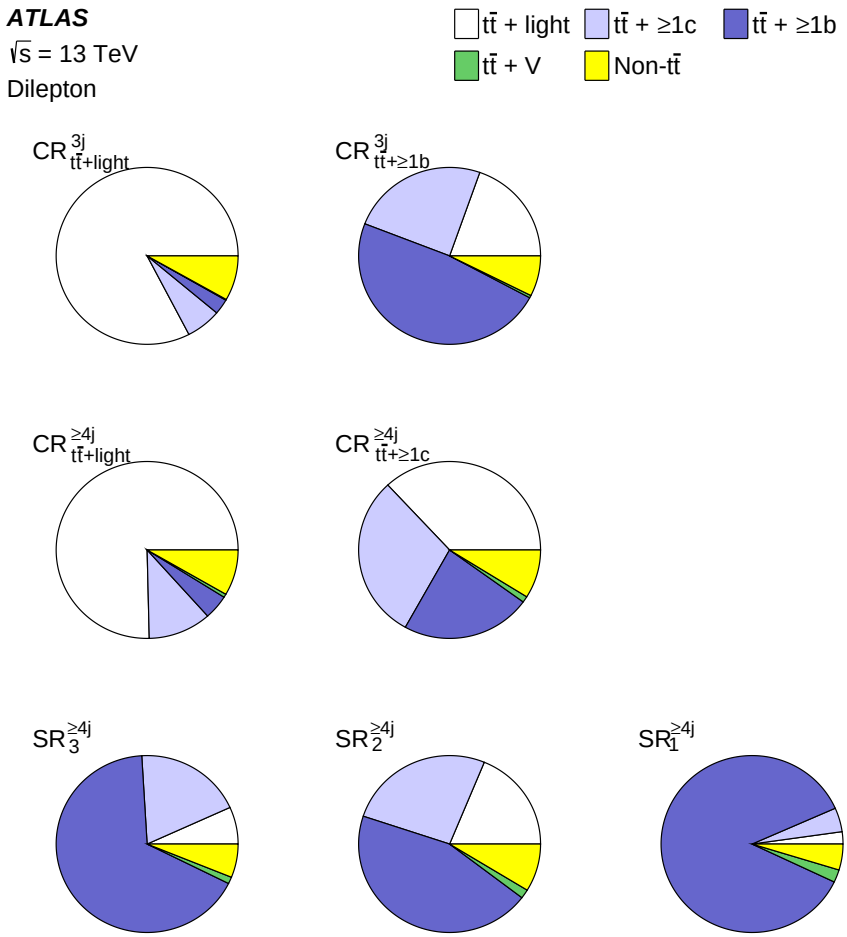


FIGURE 6.5: The relative contribution of different processes to the background in each dilepton analysis region. $t\bar{t}+\geq 1b$ refers to events containing at least one additional b-jet, $t\bar{t}+\geq 1c$ refers to events containing at least one c-jet and no additional b-jets, and $t\bar{t}+V$ refers to events containing an additional W or Z. [90]

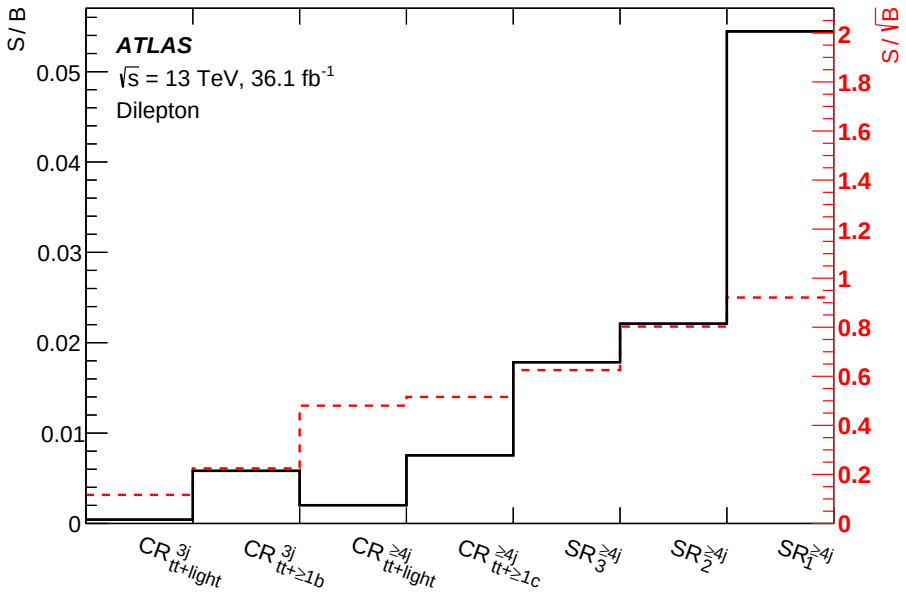


FIGURE 6.6: The signal to background ratio, S / B, and signal strength, S / $\sqrt{B}$, in each dilepton analysis region. The solid black line shows S / B values and relates to the vertical axis on the left, while the dashed red line shows S / $\sqrt{B}$ and relates to the vertical axis on the right. [90]

fitted in signal regions. The fit is performed simultaneously in all analysis regions by considering bins from each region during the fit. The fit allows for a measured value of the $t\bar{t}H$ signal strength relative to the Standard Model prediction to be extracted, as well as limits on the value of the signal strength, whilst simultaneously constraining and correcting background processes.

6.4 Object Definitions and Event Selection

Physics objects are defined using the methods described in Section 4.2. The primary vertex of the event is defined as the vertex with the largest sum of $(p_{\mathrm{T}})^2$ of associated tracks. Electrons are defined using the loose likelihood identification criteria and are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.47$, and must pass the loose lepton isolation working point. Electrons which fall into the region between the barrel and endcap electromagnetic calorimeter modules, defined as $1.37 < |\eta| < 1.52$, are excluded. Similarly muons are defined using the loose identification criteria and the loose lepton isolation working point, and are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$. Tracks for both electrons and muons are required to match the primary vertex of the event by requiring $|\mathrm{IP}_z| < 0.5$

mm, and requiring $|\mathrm{IP}_{r\phi}|/\sigma_{\mathrm{IP}_{r\phi}} < 5(3)$ for electrons (muons). Here, $|\mathrm{IP}_z|$ is the longitudinal impact parameter, defining the distance of closest approach between the lepton track and the primary vertex in z, and $|\mathrm{IP}_{r\phi}|/\sigma_{\mathrm{IP}_{r\phi}}$ is the significance for the transverse impact parameter, defining the analogous quantity in $r\phi$ space. Jets are defined using the anti-k_t jet algorithm with a radius of 0.4 and are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. Quality criteria, as discussed in Section 4.2.4, are imposed on jets to identify jets from noncollision sources or detector noise [72], and any event containing jets failing these criteria is removed. To reduce jets from pile-up interactions, jets with $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$ are required to pass an appropriate threshold on the value of the jet vertex tagger, as discussed in Section 4.2.4. b-tagged jets are identified using the MV2c10 algorithm. Hadronic taus are required to pass the medium tau-identification working point, and are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. As mentioned in Section 6.1, electrons and muons which originate from the decay of a tau which itself originates from a decaying W boson are considered in the same way as other electrons and muons used in the analysis.

An overlap removal procedure is employed to ensure that detector measurements are not used in the definition of more than one reconstructed object, which would result in double counting measurements. This procedure removes objects from an event if they are considered to overlap with other objects, according to a defined order of preference for which object should be kept. Overlaps are considered between: electrons and jets, muons and jets, and hadronically decaying taus and electrons or muons. For electrons and jets the closest jet within $\Delta R_y = \sqrt{\Delta y^2 + \Delta\phi^2} = 0.2$ of a chosen electron is removed, where y is the rapidity of an object. If there is a jet within $\Delta R_y < 0.4$ of the electron after the first selection the electron is itself removed. For muons and jets overlap is considered if a muon and jet are within $\Delta R_y < 0.4$ of each other. If the jet has three or more tracks the muon is removed; this scenario is taken to be indicative of muons produced from heavy-flavour decays within jets, and therefore this overlap removal reduces the presence of such muons. If the jet has less than three tracks the jet is removed instead, which avoids muons undergoing significant energy loss being identified as a jet. A hadronic tau is removed if it is within $\Delta R_y < 0.2$ of an electron or muon.

After reconstructed objects have been defined and overlap removal has been performed, an event selection is applied to remove events which are not relevant to the analysis or which are poorly modelled. For the dilepton analysis channel this selection

is split into 3 sub-channels based on the flavour of the two leptons: ee, $\mu\mu$, and $e\mu$. The $e\mu$ channel comprises two sub-channels: the $e\mu$ channel, where the leading p_{T} lepton is an electron, and the μe channel, where the leading p_{T} lepton is a muon; however these sub-channels are not considered in most analysis definitions or studies. Events are first selected under the requirement that all relevant detector subsytems are operational and performing well; this is detailed by a Good Runs List (GRL) which labels blocks of data that fulfil this condition. Events are also required to have at least one vertex with at least two tracks that have $p_{\mathrm{T}} > 0.4$ GeV. Events are then selected based on whether they fire relevant triggers. In both the dilepton and single-lepton channels triggers require a single lepton, labelled as single-lepton triggers, and in some cases additional requirements on lepton identification and isolation are applied. A range of triggers are combined through a simple logical OR of different triggers: some triggers have lower p_{T} requirements but have tighter identification and isolation requirements, while other triggers have higher p_{T} requirements but are looser with respect to identification and isolation. In this way selection efficiency is maximised by recovering higher p_{T} leptons which do not pass the stringent identification and isolation requirements of lower p_{T} triggers. There are also different triggers employed for 2015 data and 2016 data, where the 2015 triggers feature lower p_{T} requirements and in some cases looser identification and isolation requirements, but are otherwise similar triggers to those used in 2016 data. Due to the lower instantaneous luminosity and pile-up of 2015 data-taking, lower trigger thresholds were able to be used while keeping trigger rates acceptable. The triggers used in 2015 and 2016 and their requirements are shown in Table 6.1.

For the ee channel 3 single-electron triggers are used: low, medium, and high p_{T}. Identification requirements loosen as the p_{T} threshold increases, and the low p_{T} trigger employs loose variable isolation in 2016 data taking. No medium p_{T} trigger is used in 2015 data taking because the identification and isolation requirements for the low p_{T} trigger are already relatively loose. For the $\mu\mu$ channel 2 single-muon triggers are used: low and medium p_{T}. Loose isolation is required for the low p_{T} 2015 trigger, whereas medium variable isolation is used for the low p_{T} 2016 trigger. The medium p_{T} trigger has no identification or isolation requirements. Finally for the $e\mu$ channel the logical OR of all the triggers employed in the ee channel with the triggers employed in the $\mu\mu$ channel is used. These triggers are used for the 32.1 fb^{-1} analysis. The 13.2 fb^{-1} analysis uses a set

TABLE 6.1: Triggers employed in the analysis event selection split by data taking year, and their associated requirements on lepton p_{T}, lepton identification, and lepton isolation. The name column specifies a name to conveniently refer to these triggers. Dashes are used to indicate table entries that are not applicable.

Name	2015			2016		
	p_{T} [GeV]	ID	Isolation	p_{T} [GeV]	ID	Isolation
Single-electron triggers						
*e*_lowPt	24	medium	none	26	tight	variable loose
*e*_medPt	-	-	-	60	medium	none
*e*_highPt	120	loose	none	140	loose	none
Single-muon triggers						
μ_lowPt	20	-	loose	26	-	variable medium
μ_medPt	50	-	none	50	-	none

of triggers with 24 GeV p_{T} thresholds on the *e*_lowPt and μ_lowPt triggers in 2016. The set of triggers are otherwise identical between the two analyses. When applying these trigger selections to simulated data, events are randomly assigned to either the 2015 or 2016 selection, where the proportion of events in each matches the proportion of events in real data samples.

Further selection criteria are then placed on the events which pass these triggers. All dilepton sub-channels are required to contain exactly two oppositely charged leptons. These leptons are also required to pass tighter identification than the initial lepton object definitions; this improves the purity of the event selection. Electrons are required to pass the tight likelihood identification working point and Gradient isolation criteria. Muons are required to pass the medium identification working point and are also required to pass Gradient isolation criteria. p_{T} requirements are then placed on these selected leptons. The leading lepton is required to have $p_{\mathrm{T}} > 27$ GeV in all channels for both 2015 and 2016 events. For the 13.2 fb^{-1} analysis this leading lepton requirement is instead $p_{\mathrm{T}} > 25$ GeV. This requirement is imposed by the single-lepton triggers which are used: to ensure that trigger rates in simulated data match what is seen in real data, lepton thresholds are required to be 1 GeV greater than what is required by the lowest p_{T} threshold trigger. For 2016 the lowest trigger threshold in both single-electron and single-muon triggers is 26 GeV, which therefore defines a 27 GeV analysis selection threshold, which is also consistently applied to 2015 events. This leading lepton is also required to

be matched to a lepton reconstructed in the trigger which has the same flavour and is within $\Delta R < 0.15$. The subleading lepton is required to have $p_{\mathrm{T}} > 10$ GeV in the $\mu\mu$ and $e\mu$ channels, and $p_{\mathrm{T}} > 15$ GeV in the ee channel. For the ee and $\mu\mu$ channels, selections are then applied to the dilepton pair. The dilepton invariant mass is required to be greater than 15 GeV to reject low mass resonances and must be outside of the range between 83 – 99 GeV to reject events from Z boson decays; this is the so-called Z-window cut. Events are also rejected if they contain any hadronic tau candidates, to ensure orthogonality with other ATLAS $t\bar{t}H$ analyses searching for channels which include leptons and hadronic taus. Selections are then placed on jets. Events are required to contain at least three jets with $p_{\mathrm{T}} > 25$ GeV. Finally at least two jets are required to be b-tagged at the medium b-tagging working point.

6.4.1 Studies of the Choice of Triggers in 2016 Data

As described previously, the single-lepton and dilepton channels use the same single-lepton triggers to select events. However, there are additional triggers available which could allow inclusion of more events. Specifically, triggers which require two leptons, labelled dilepton triggers, and triggers which require a single lepton with a number of jets, labelled lepton+jets triggers. The lepton momentum and isolation thresholds of dilepton and lepton+jets triggers are looser than those of the single-lepton triggers due to the reduction in trigger rates brought by requiring additional physics objects; using dilepton and lepton+jets in the analysis could therefore allow inclusion of events containing leptons which fall below the thresholds of single-lepton triggers. In particular, the simulated datasets used in the analysis include events with a leading lepton which has $p_{\mathrm{T}} \geq 20$ GeV. Because the single-lepton triggers impose a leading lepton requirement of $p_{\mathrm{T}} > 27$ GeV, this leaves the range 20 GeV $\leq p_{\mathrm{T}} < 27$ GeV for leading leptons which can be included by employing dilepton and lepton+jets triggers.

The effects of including dilepton and lepton+jets triggers while lowering the leading lepton p_{T} requirement were studied using simulated $t\bar{t}H$ and $t\bar{t}$ datasets, to provide an estimate of the change in yields for signal and for the dominant background process, and therefore a rough estimate of the change in signal purity and signal significance. The $t\bar{t}$ simulated dataset used in this study was the POWHEG+PYTHIA 6 dataset (the newer

POWHEG+PYTHIA 8 dataset was not available at the time this study was performed). In addition the $t\bar{t}$ yield was split according to the presence, or not, of a non-prompt or fake lepton in the event according to the method described in Section 6.5.1. It is expected that decreasing the leading lepton p_{T} requirement when using additional triggers would also increase the proportion of events containing non-prompt or fake leptons because the proportion of non-prompt or fake leptons is expected to increase as lepton p_{T} decreases. The dominant source of non-prompt or fake leptons is from $t\bar{t}$ background events and so measuring the proportion of these leptons in $t\bar{t}$ processes allowed a rough estimate of the effect on the proportion of non-prompt or fake leptons in the analysis. The study was performed for events passing 2016 event selections; due to the lower p_{T} thresholds for single-lepton triggers in 2015 event selections as well as the lower integrated luminosity of 2015 data, the focus was on how yields in 2016 would be affected.

The dilepton and lepton+jets triggers which were available in analysis-level datasets at the time of this study are split into ee dilepton triggers, $\mu\mu$ dilepton triggers, $e\mu$ dilepton triggers, and μ+jets lepton+jets triggers. e+jets triggers were used during data taking, however these triggers were not available in the analysis-level datasets at the time of this study. The dilepton triggers are either symmetric triggers, which place the same requirements on both leptons, or asymmetric triggers, which place tighter requirements on one of the leptons. The μ+jets trigger used requires one muon plus three jets with equal p_{T} requirements on each jet. The additional triggers used and their requirements are shown in Table 6.2. As mentioned in Section 6.4, trigger requirements also imply analysis selections to ensure the good modelling of trigger rates in simulated data compared to real data. For requirements on leptons this is again taken to be an analysis selection which is 1 GeV higher than the p_{T} requirement on the trigger lepton. For the 20 GeV requirement on jets the recommended analysis selection is requiring a jet p_{T} of at least 80 GeV; this is due to the much longer 'trigger turn-on curve' compared to purely leptonic triggers, where the efficiency of the trigger has not reached a plateau value and which is difficult to model in simulated data. As such the μ+`3jets` trigger only contributes additional events where there are three jets with p_{T} of at least 80 GeV each.

The effect of utilising additional triggers was studied by comparing yields when using only the single-lepton triggers and a leading lepton p_{T} selection of 27 GeV to yields

obtained when using a logical OR of single-lepton triggers and additional triggers during the event selection described previously, while also reducing the leading lepton p_{T} threshold in this event selection to either 20 GeV, or 1 GeV above the leading lepton trigger threshold, depending on which was higher. The changes in yields were studied for using single-lepton triggers and dilepton triggers, single-lepton triggers and the lepton+jets trigger, and finally for using single-lepton triggers and the dilepton triggers and the lepton+jets trigger. The change in signal purity and signal significance is then estimated by calculating S/B and $S/\sqrt{B}$ for each case and comparing the values to those obtained for just using single-lepton triggers with the 27 GeV leading lepton selection, where S is the $t\bar{t}H$ yield and B is the total $t\bar{t}$ yield for events containing a prompt lepton and events containing a non-prompt or fake lepton. Signal purity and signal significance are both studied to allow for an estimate of the change in statistical power while observing any underlying change in signal purity. Results are studied in a few analysis regions based on the 13.2 fb^{-1} definitions, by placing a selection on number of jets and number of b-tags at the tight b-tagging working point. $3j, 2b$ is used as an example of a three jet region and an example of a control region, while $\geq 4j, \geq 4b$ is used to study results in a main signal region. The results of these studies are shown in Table 6.3. Uncertainties are not quoted but the statistical uncertainty from the number of simulated events is expected to be less than the one significant figure rounding quoted in the table.

It is seen that yield increases are generally modest for $t\bar{t}H$, and for $t\bar{t}$ containing prompt leptons, with increases at most of 5%. On the other hand the yield of $t\bar{t}$ events containing fake or non-prompt leptons increases more significantly as expected, with increases of up to 10% when using dilepton triggers and lepton+jets triggers. However the changes to S/B are almost negligible, and $S/\sqrt{B}$ increases by at most 3%. It must be noted that the change in yields for SL OR DIL OR L+J is not simply the addition of the change in yields for SL OR DIL and SL OR L+J due to overlap of events gained in each configuration. As well it must be noted that although the yield for $t\bar{t}$ events containing fake or non-prompt events has a relatively larger increase than other yields, the associated change to S/B and $S/\sqrt{B}$ is still small due to this category of events making up approximately 1% of total background events.

Due to the small change in yields and $S/\sqrt{B}$ it was decided not to use dilepton or lepton+jets triggers in the analysis and not to attempt to lower the leading lepton p_{T}

TABLE 6.2: Additional triggers considered for use in the analysis and their associated requirements on lepton or jet p_{T}, lepton identification, and lepton isolation. The name column specifies a name to conveniently refer to these triggers. If a single value is shown for a column entry then this value is applied to all objects within the row. The p_{T} values for the $\mu\mu$_asym trigger are comma separated to indicate that each lepton has a different p_{T} requirement. Dashes are used to indicate table entries that are not applicable..

Name	p_{T} [GeV]	ID	Isolation
ee dilepton triggers			
*ee*_sym	17	very loose	none
$\mu\mu$ dilepton triggers			
$\mu\mu$_sym	14	-	none
$\mu\mu$_asym	22, 8	-	none
$e\mu$ dilepton triggers			
$e\mu$_asym	e: 17	e: loose	e: none
	μ: 14	μ: -	μ: none
μe_asym	e: 7	e: none	e: none
	μ: 24	μ: -	μ: none
μ+jets lepton+jets triggers			
μ+3jets	μ: 14	μ: -	μ: variable loose
	3 jets: 20	3 jets: -	3 jets: -

selection.

6.5 Data and Simulated Event Datasets

As described in Section 6.1 the ATLAS $t\bar{t}H(b\bar{b})$ analysis uses p-p collision events recorded in 2015 and 2016 at $\sqrt{s} = 13$ TeV. The 13.2 fb^{-1} analysis uses 3.2 ± 0.1 fb^{-1} of 2015 data and 10.0 ± 0.4 fb^{-1} of 2016 data [92] (values quoted come from the removed publicly released 13.2 fb^{-1} analysis, and so internal ATLAS documentation is cited instead). The 36.1 fb^{-1} analysis uses the same 2015 dataset with a larger 32.9 ± 0.7 fb^{-1} 2016 dataset [90].

The analysis uses a large number of Monte Carlo simulated datasets, covering the range of background processes considered and the use of additional simulation configurations for systematic uncertainty estimations and analysis validation. As well, the Z+jets process and processes containing fake or non-prompt leptons require dedicated corrections derived from data which are applied prior to the statistical analysis to ensure they are well modelled. Simulated datasets include a number of features which are common across all samples and configurations, which will now be listed. The top quark mass is

TABLE 6.3: The relative change in yields and signal purity and significance when using expanded trigger combinations and a lowered leading lepton p_{T} selection, compared to the values obtained when using only single-lepton triggers and a 27 GeV leading lepton p_{T} selection. Values are obtained for analysis regions defined using simple cuts on numbers of jets and b-tags, where b-tags are defined using the tight working point. Fake refers to the presence of non-prompt or fake leptons. SL refers to single-lepton triggers, DIL refers to dilepton triggers, and L+J refers to lepton+jets triggers. Values are rounded to one significant figure.

Trigger combination	$\Delta\text{yield}_{t\bar{t}H}$	$\Delta\text{yield}^{\text{prompt}}_{t\bar{t}}$	$\Delta\text{yield}^{\text{fake}}_{t\bar{t}}$	$\Delta(\frac{S}{B})$	$\Delta(\frac{S}{\sqrt{B}})$
		$\geq 3j, \geq 2b$			
SL OR DIL	4%	4%	7%	−0.1 %	2 %
SL OR L+J	2%	1%	3%	1 %	2 %
SL OR DIL OR L+J	5%	5%	10%	0.2 %	3 %
		$3j, 2b$			
SL OR DIL	4%	4%	7%	−0.6 %	1 %
SL OR L+J	1%	1%	1%	−0.02%	0.5%
SL OR DIL OR L+J	4%	5%	9%	−0.6 %	2 %
		$\geq 4j, \geq 4b$			
SL OR DIL	3%	4%	8%	−0.3 %	1 %
SL OR L+J	2%	2%	8%	0.1 %	1 %
SL OR DIL OR L+J	4%	5%	8%	−0.2 %	2 %

taken to be $m_t = 172.5$ GeV. The effects of pile-up interactions are modelled by generating interactions using PYTHIA 8.186 [84] and then overlaying these interactions on the hard-scatter processes of the simulated dataset; the events are then reweighted to match the pile-up profile observed in data. Decays of b- and c-hadrons are simulated using the EVTGEN v1.2.0 package [93], except in datasets simulated using the SHERPA event generator which performs these decays itself. In most cases interaction with the ATLAS detector is simulated using the full GEANT4 simulation [52], however some additional samples used for uncertainty estimation employ the faster AFII simulation which uses a parametrisation for calorimeter response [53]. Simulated events are also reweighted to ensure object reconstruction efficiencies, energy scales, and energy resolutions are calibrated to those observed in data. Finally, all simulated events are reconstructed using the same reconstruction algorithms as for data events.

The main background processes for the analysis after the event selection is applied are the dominant $t\bar{t}$+jets process, Z+jets, $t\bar{t}V$, diboson processes, and production of single top quarks through Wt processes. As well, the production of Higgs bosons with a single top quark in tWH and $tHqb$ processes are considered and are treated as background events. Additional small backgrounds of $t\bar{t}t\bar{t}$, $t\bar{t}WW$, tZ, and tZW are also considered. A few

processes are considered which only contribute events containing fake or non-prompt leptons: W+jets, and s-channel and t-channel single-top production. The background processes, except for $t\bar{t}$+jets and $t\bar{t}V$, are collectively grouped as 'non-$t\bar{t}$' in figures and tables; this grouping also includes all background events containing fake or non-prompt leptons, including $t\bar{t}$+jets processes including these leptons. All processes are simulated by generating events at matrix element (ME) level using a dedicated MC generator, and then interfacing the results with a dedicated parton shower (PS) generator to simulate the decays of hadronic jets and other QCD effects. The configurations used for each process are listed in Table 6.4. The datasets for the 13.2 fb^{-1} analysis are in some cases different but are usually a roughly equivalent older configuration, for example the Z+jets dataset used SHERPA 2.1.1 instead of SHERPA 2.2.1. This is not specified in the table except for the POWHEG+PYTHIA 6 $t\bar{t}$+jets sample which is used in some studies in this thesis. The additional datasets used for uncertainty estimation are detailed in Section 6.6.

The signal $t\bar{t}H$ process is simulated using a Higgs mass of $m_H = 125$ GeV. All Higgs decays are included in the simulation, allowing $t\bar{t}H$ processes which are not $t\bar{t}H(b\bar{b})$ to be counted as signal if they pass the analysis event selection. The factorisation and renormalisation scales were set to $\mu_\mathrm{F} = \mu_\mathrm{R} = H_\mathrm{T}/2$ where H_T is the sum of transverse masses $\sqrt{p_\mathrm{T}^2 + m^2}$ of all final-state particles at the level of the matrix element calculation. Top quarks are decayed using MADSPIN [94] with all spin correlations preserved. The $t\bar{t}H$ cross-section is calculated to be 507^{+35}_{-50} fb at NLO with NLO electroweak corrections [31, 95–99]. The Higgs decay branching fractions are calculated using HDECAY [31, 100].

The $t\bar{t}$+jets processes for the 36.1 fb^{-1} search were simulated with the h_damp parameter in the POWHEG+PYTHIA 8 dataset to be 1.5 times the mass of the top quark. This parameter controls the p_T of the first gluon emission beyond the Born configuration, and has the effect of regulating high-p_T radiation. This value was seen to best describe data, and particularly relevant to this analysis is that this relatively high value improves normalisation agreement between data and simulated data for events containing one or two additional b-jets [101]. The $t\bar{t}$+jets processes for the 13.2 fb^{-1} search used a lower h_damp value of 1.0 times the top mass. The renormalisation and factorisation scales were set to the transverse mass of the top quark $\sqrt{m_t^2 + p_{\mathrm{T},t}^2}$ where $p_{\mathrm{T},t}$ is the transverse momentum of the top quark in the $t\bar{t}$ centre-of-mass reference frame. The cross-section was predicted to be 832^{+46}_{-51} pb using the Top++2.0 program [102].

TABLE 6.4: Processes considered in the ATLAS $t\bar{t}H(b\bar{b})$ analysis and the configurations used to create baseline simulated event datasets. Order and normalisation refer to the order of QCD processes considered, unless otherwise specified.

Process	Matrix Element Generator	Order	Normalisation	PDF	Parton Shower Generator	PS Tune
$t\bar{t}H$	MG5_aMC [103]	NLO	NLO+NLO EW [31, 95–99]	NNPDF3.0NLO [104]	PYTHIA 8.2	A14 [105]
$t\bar{t}$+jets (36.1 fb^{-1} search)	POWHEG v2 [85–88]	NLO	NNLO+NNLL [106–109]	NNPDF3.0NLO	PYTHIA 8.2	A14
$t\bar{t}$+jets (13.2 fb^{-1} search)	POWHEG v2	NLO	NNLO+NNLL	CT10 [110, 111]	PYTHIA 6.4 [112]	Perugia 2012 [113]
Z+jets	SHERPA 2.2.1 [114]	4 partons at LO 2 partons at NLO	NNLO [115]	NNPDF3.0NNLO	SHERPA [116]	Sherpa author's tune
W+jets	SHERPA 2.2.1	4 partons at LO 2 partons at NLO	NNLO	NNPDF3.0NNLO	SHERPA	Sherpa author's tune
$t\bar{t}V$	MG5_aMC	NLO	NLO	NNPDF3.0NLO	PYTHIA 8.2	A14
Single-top (s-channel, Wt)	POWHEG v1	NLO	NNLO [117–119]	CT10	PYTHIA 6.4	Perugia 2012
Single-top (t-channel)	POWHEG v1	NLO	NNLO	CT10 4F [110, 111]	PYTHIA 6.4	Perugia 2012
Diboson	SHERPA 2.1.1	See Ref. [120]	NLO	CT10	SHERPA	Sherpa author's tune
tWH	MG5_aMC	NLO	NLO	CTEQ6L1 [121]	HERWIG++ [122]	UE-EE5
$tHqb$	MG5_aMC	LO	LO	CT10 4F	PYTHIA 8	A14
$t\bar{t}t\bar{t}$	MG5_aMC	LO	LO	NNPDF3.0NLO	PYTHIA 8	A14
$t\bar{t}WW$	MG5_aMC	LO	LO	NNPDF3.0NLO	PYTHIA 8	A14
tZ	MG5_aMC	LO	LO	NNPDF3.0NLO	PYTHIA 6	Perugia 2012
tZW	MG5_aMC	NLO	NLO	NNPDF3.0NLO	PYTHIA 8	A14

As seen in the analysis region definition described in Section 6.3, $t\bar{t}$+jets events are categorised based on the flavour composition of the additional jets in the events. This categorisation is performed by defining generator-level jets of particles. These jets are distinct from reconstructed jets and b-tags due to the definitions using 'truth' information of the particles in the simulated events and hadronic decays, as opposed to reconstructed information from measurements made with the ATLAS event reconstruction. These generator jets are defined by considering all stable particles (mean lifetime greater than 3×10^{-11} seconds) and reconstructing them as jets using the anti-k_t algorithm with a radius of $R = 0.4$. These jets are required to have $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$. They are then categorised by jet flavour by counting the number of b- or c-hadrons within $\Delta R < 0.4$ of the jet axis. Jets containing one b-hadron with $p_{\mathrm{T}} > 5$ GeV are labelled single-b-jets, while jets containing 2 or more b-hadrons where one has $p_{\mathrm{T}} > 5$ GeV and the others have no p_{T} restriction are labelled B-jets. Single-c- and C-jets are defined analogously for jets which are not already categorised as single-b- or B-jets. These jets can then be used to categorise the events according to what is shown in Section 6.3, for example with the presence of additional single-c- or C-jets and no additional single-b- or B-jets defining $t\bar{t}+ \geq 1c$ events. As well more detailed categorisation is applied when assigning correction factors and estimating uncertainties for $t\bar{t}+ \geq 1b$ events, as follows:

- $\mathbf{t\bar{t} + b\bar{b}}$: exactly two additional single-b-jets;
- $\mathbf{t\bar{t} + b}$: exactly one additional single-b-jet;
- $\mathbf{t\bar{t} + B}$: exactly one additional B-jet;
- $\mathbf{t\bar{t}+ \geq 3b}$: events not falling into the above categories;
- $\mathbf{t\bar{t} + b(MPI/FSR)}$: events containing additional b-jets originating entirely from multiparton interactions (MPI) or final-state radiation (FSR).

b-jets in the $t\bar{t} + b$(MPI/FSR) category are calculated by the parton shower generator, which has been tuned to such processes from data. A description of this tuning process is given in the reference for the A14 tune used in PYTHIA 8 [105] (and references therein).

Given the limitations in modelling $t\bar{t}$+jets events containing additional heavy-flavour jets, the POWHEG+PYTHIA 8 dataset is corrected using a dedicated $t\bar{t}b\bar{b}$

dataset. Specifically, the relative contributions of the $t\bar{t}+ \geq 1b$ categories shown in the above list, except for $t\bar{t}+b$(MPI/FSR), are corrected to have the same fractions as in the dedicated $t\bar{t}b\bar{b}$ dataset. The $t\bar{t}b\bar{b}$ dataset is generated at NLO precision [123] using SHERPA+OPENLOOPS [124] with SHERPA 2.1.1 and the CT10 four flavour scheme (4F) PDF set [110, 111]. The renormalisation scale was set to the CMMPS value $\mu_{\text{CMMPS}} = \prod_{i=t,\bar{t},b,\bar{b}} E_{\text{T},i}^{1/4}$ [123], and the factorisation scale was set to $H_{\text{T}}/2 = \frac{1}{2}\sum_{i=t,\bar{t},b,\bar{b}} E_{\text{T},i}$. The resummation scale, which sets an upper limit on the hardness of parton shower emissions, was also set to $H_{\text{T}}/2$. This sample is referred to as 'SHERPA4F'. '4F' in this name, and in association with the CT10 4F PDF set, refers to the fact that b quarks are treated as massive as opposed to being treated as massless (along with u, d, s, c quarks) in the five flavour (5F) scheme which is used in most other samples. In the CT10 4F PDF set this is done by employing the fixed flavour number scheme (specifically for four active flavours), where b quarks are not considered as partons within the proton and are instead treated exclusively as massive final-state particles. With this consideration 'SHERPA4F' calculates the kinematics of the two additional b-jets with NLO precision in QCD. As a result this dataset is chosen as the most precise prediction for $t\bar{t}+ \geq 1b$ events. The comparison between relative predicted fractions for each considered $t\bar{t}+ \geq 1b$ category is shown in Figure 6.7. b-jets from $t\bar{t}+b$(MPI/FSR) are separate from the additional two b quarks in SHERPA4F and are in fact not present in this sample, and so as mentioned this category is not scaled. $t\bar{t}+b$(MPI/FSR) events make up 10% of POWHEG+PYTHIA 8 $t\bar{t}+ \geq 1b$ events.

For the Wt single-top dataset, interference effects from overlap between $t\bar{t}$ and Wt processes was handled by employing the 'diagram removal' scheme [125]. For the t-channel single-top dataset the top quarks were decayed with MADSPIN.

The Z+jets and W+jets datasets were generated at ME level using COMIX [126] and OPENLOOPS. The matrix elements are then merged with the SHERPA PS calculation using the ME+PS@NLO prescription [127]. As mentioned previously the Z+jets dataset has a correction applied which is derived from comparisons with data. This correction factor is derived by comparing data yields for events containing opposite-sign same-flavour lepton pairs (e^+e^- or $\mu^+\mu^-$) for an event selection where the Z-window cut is inverted to select events with dilepton invariant mass within the range 83 – 99 GeV. The result of

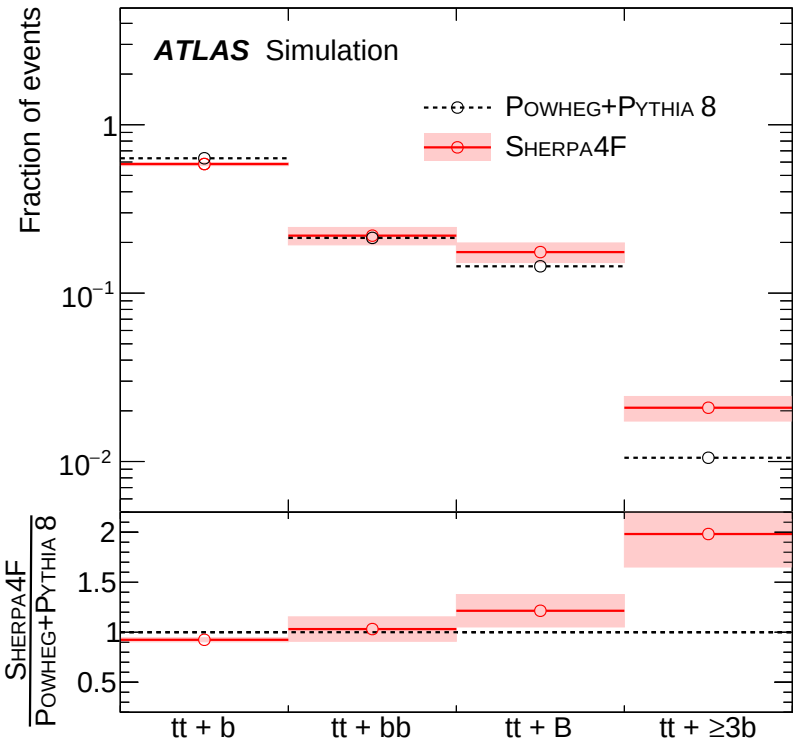


FIGURE 6.7: Relative fractions of each $t\bar{t}+ \geq 1b$ sub-category predicted by the SHERPA4F $t\bar{t}b\bar{b}$ dataset and compared with the POWHEG+PYTHIA 8 $t\bar{t}$ dataset. The fractions are shown before any event selection. The shaded region on SHERPA4F shows its associated uncertainties discussed in Section 6.6. Events from the $t\bar{t} + b$(MPI/FSR) category are not included in the normalisation of this figure. [90]

this is that Z+jets events containing heavy-flavour jets are scaled by 1.3 in event normalisation, while Z+jets events without heavy-flavour jets are not scaled.

6.5.1 Estimation of Yields and Distribution Shapes for Events Containing Fake or Non-Prompt Leptons

As mentioned previously there is a background category containing events from all background processes which contain fake or non-prompt leptons. Specifically, these events contain leptons passing the object definitions and analysis selection described in Section 6.4 where one or both of the leptons are either:

- measurements in the ATLAS detector which do not originate from a lepton but are mis-reconstructed as a lepton, labelled 'fake' leptons;
- or leptons produced from the decay of particles which were not directly produced by the hard scatter process, labelled 'non-prompt' leptons (for example leptons coming from the decays of b quarks which could themselves originate from the decaying $t\bar{t}$ pair).

For convenience both fake and non-prompt contributions are referred to as 'fake' leptons within this section, and events containing fake leptons are simply referred to as fake events (although of course these are still genuine simulated events which are otherwise equivalent to other events).

The origin and yields of fake leptons can be studied by utilising truth information in simulated event datasets, which can detail the origin and types of reconstructed leptons. In many analyses this information would be an input or reference for methods which estimate fake yields using relatively complicated data-driven techniques; however one significant benefit of targeting the $t\bar{t}H(b\bar{b})$ dilepton final state is that this information from simulation can be considered more accurate than in other channels and can therefore make use of simpler corrections from data. This is so because of the selection of two leptons passing the identification, isolation, and p_{T} requirements described in Section 6.4. Fake leptons can be produced by the highly numerous QCD background processes present in p-p collisions, however the chance to produce two fake leptons which pass these event selection requirements is effectively negligible. These QCD processes

are the main focus of the fake estimate methods used in other analyses. The processes are difficult to accurately or precisely model theoretically, leading to the need for estimates which are taken directly from data. Without the need to consider QCD processes in the $t\bar{t}H(b\bar{b})$ dilepton channel the predictions from simulation are expected to be much more reliable. In principle the more advanced data-driven methods could be used in the dilepton channel, for example the $t\bar{t}H(b\bar{b})$ single-lepton analysis uses a method called the matrix method [128]. However the work involved in using this method in the dilepton channel becomes even more technically demanding than the single-lepton case due to the need to consider matrices for combinations of two leptons as opposed to a single lepton, and any improvement to fake lepton estimates are expected to be small.

Fake electrons and fake muons originate from different sets of sources, reflecting the difference in how these physics objects are reconstructed within ATLAS. Fake electrons can originate from: photons which interact within detector material to produce real electrons or e^+e^- pairs; decays of mesons produced in hadronic jets into real electrons, as well as other non-prompt particles which decay into a real electron; overlap of a track and an EM calorimeter deposit from a photon; and mis-reconstruction of hadronic jets with few tracks and significant EM showering as an electron. Fake muons can originate from: decays of mesons produced in hadronic jets into real muons, as well as other non-prompt particles which decay into a real muon; hadrons which pass into the muon detection systems either without showering within the hadronic calorimeter, or 'punching-through' the hadronic calorimeter during showering to reach the muon detection modules; and light mesons which decay early within the tracking volume.

The method used to estimate fake events in the dilepton channel, referred to as the MC estimate method, is performed by first categorising selected leptons as either being prompt leptons or fake leptons originating from the sources described above by looking at the truth information associated with the leptons. All fake events from background process datasets can then be gathered into a fake background category, forming the fake background estimate from simulation. Fake events from signal $t\bar{t}H$ processes are not included in this category due to the event still counting as a signal Higgs process. Data is then introduced as a way of deriving correction factors to ensure the fake estimate models data well, as well as estimating systematic uncertainties based on the discrepancy between simulation and data. This is done by inverting the opposite-charge-sign (OS)

requirement on the two leptons in the event selection to define same-sign (SS) control regions which can be used for data against simulation comparisons. This is done based on the idea that there will be relatively more processes that can produce a fake lepton with the same sign as the other selected lepton than there are processes which can do the same but with two prompt leptons, meaning that the same-sign region is relatively fake-enriched. More precisely, this is based on the idea that there are fewer Standard Model processes that produce same-sign lepton pairs. The other assumption is that the accuracy of modelling of fake leptons will be similar for OS and SS events, meaning that correction factors can be relatively accurately extrapolated from SS events to the OS events used in the analysis. Modelling of fakes in SS regions is then studied by comparing fake-category event yields to the quantity $(N^{\text{data, SS}} - N^{\text{MC, prompt, SS}})$, where N is the yield in a particular histogram bin, the superscript 'MC, prompt' refers to prompt-category simulated events, and the superscript SS refers to the same-sign requirement. In this way fake-category simulated events can be compared to 'fake data' events, under the assumption that prompt-category simulated events are well modelled in SS regions.

The categorisation of selected leptons using truth information is done by spatially matching the reconstructed leptons to truth particles within the simulated event. This spatial matching is done by retrieving a truth track which is associated to the reconstructed tracks of the leptons. The truth particle itself is retrieved from the truth track, and a reconstructed lepton is considered to be matched to this truth information if it is within $\Delta R < 0.2$ of the truth particle. The origin of the truth particle can then be traced backwards through the chain of truth interactions to determine the origin of the truth particle. In this way prompt leptons are defined as:

- **Prompt isolated electrons**: truth electrons which are themselves promptly produced;
- **Photon-conversion-recovered electrons**: truth electrons which originate from a photon which itself originated[1] from a prompt electron produced by: a top quark, a W boson, a Z boson, or a Higgs boson;
- **Prompt isolated muons**: truth muons which are themselves promptly produced.

[1]This origin includes photons which are produced by a chain of electron and photon production via bremmstrahlung and photon conversions, as long as this chain originated from a prompt electron.

Leptons failing these truth selections are categorised as fake leptons, including reconstructed leptons which are not spatially matched to a truth particle.

After fake-category events are selected based on the presence of fake leptons, they are split into two channels when considering modelling of fakes and any corrections to be applied. These channels are defined based on the flavour of the subleading lepton in each event: the 'subleading-electron' channel comprises events from the ee and μe analysis channels, while the 'subleading-muon' channel comprises events from the $\mu\mu$ and $e\mu$ analysis channels. This channel definition is made under the assumption that fake leptons are more likely to be produced at lower p_{T} compared to prompt leptons. As such, fake events in the subleading-electron and subleading-muon channels should be dominantly comprised of fake leptons of the same flavour as the channel definition. Previous studies performed by the ATLAS $t\bar{t}H(b\bar{b})$ analysis group for the 13.2 fb^{-1} analysis showed that 98% of fake muons were in the subleading-muon channel, and 80% of fake electrons were in the subleading-electron channel [92] (these values were not released publicly, and so the internal ATLAS documentation is cited). Events in these channels are then studied using the same-sign selection in an inclusive analysis region by requiring $N_{\text{jets}} \geq 3$ AND $N_{b\text{-tags}}^{\text{tight}} \geq 2$. More exclusive region definitions are not used due to the low number of events passing the same-sign selection. Studies of same-sign events in this region were performed using 36.5 fb^{-1} of 2015+2016 data (which is roughly the same dataset as the 36.1 fb^{-1} dataset used in the final analysis aside from using an older luminosity calculation and having some additional data runs included or removed based on differing Good Runs Lists) and most simulated event datasets, except for tH processes. Older SHERPA 2.1 datasets were used for Z+jets and W+jets due to the unavailability of the newer SHERPA 2.2.1 samples at the time this study was performed. As well, Z+jets events had no Z+HF correction factor applied. Under the assumption that the contribution from events containing two fake leptons is effectively negligible, no all-hadronic final state datasets are used.

Histograms of distributions within this same-sign inclusive region comparing data to simulated data are shown for the subleading-electron channel in Figure 6.8 and for the subleading-muon channel in Figure 6.9. Comparisons are made between all simulated events and real data, and also between fake-category simulated events and real data with

prompt-category simulated events subtracted. Uncertainties are statistical only. Yield tables for fake-category simulated events and 'fake data' are shown in Table 6.5. It is seen that modelling is reasonably good, although the low yield of data events in the same-sign region restricts the statistical precision, as do lower numbers of simulated events. Significant shape mismodelling is not seen from any of the figures, however overall normalisation disagreements are seen in both channels in the yield table. As such no shape corrections are applied, but normalisation correction factors are applied using Equation 6.1:

$$N^{\text{fake, OS}} = c_{\text{fake}}\, N^{\text{MC, fake, OS}} \equiv \frac{(N^{\text{data, SS}} - N^{\text{MC, prompt, SS}})}{N^{\text{MC, fake, SS}}} N^{\text{MC, fake, OS}} \tag{6.1}$$

N denotes the total yield in a region, and the superscripts refer to the source of events: MC refers to events from simulation, prompt (fake) refers to prompt-category (fake-category) events, OS (SS) refers to events from opposite-sign (same-sign) regions. c_{fake} is the correction factor to be applied. Applying this equation to Table 6.5 the correction factor for the subleading-electron channel is $c_{\text{fake}}^{\text{sub-}e} = 1.18$ and the correction factor for the subleading-muon channel is $c_{\text{fake}}^{\text{sub-}\mu} = 1.31$. These factors are applied to corresponding opposite-sign events as event weights during event normalisation. After applying these correction factors, it is seen that fake-category background events make up approximately 1% of the total background in opposite-sign analysis regions, including individual exclusive analysis regions defined in Section 6.3. A normalisation uncertainty based on these correction factors is placed on the fake background category, and is discussed in Section 6.6.

6.6 Systematic Uncertainties

As mentioned in Section 6.2, the precision of results in the $t\bar{t}H(b\bar{b})$ analysis is limited due to the size of systematic uncertainties, which means that detailed estimation and appropriate handling of systematic uncertainties is important. As mentioned in Section 6.1, systematic uncertainties are in general encoded in the statistical analysis as additional nuisance parameters, the usage of which is discussed in Section 6.7. Two sources of uncertainty representing the normalisation of $t\bar{t}$+HF components are instead encoded as

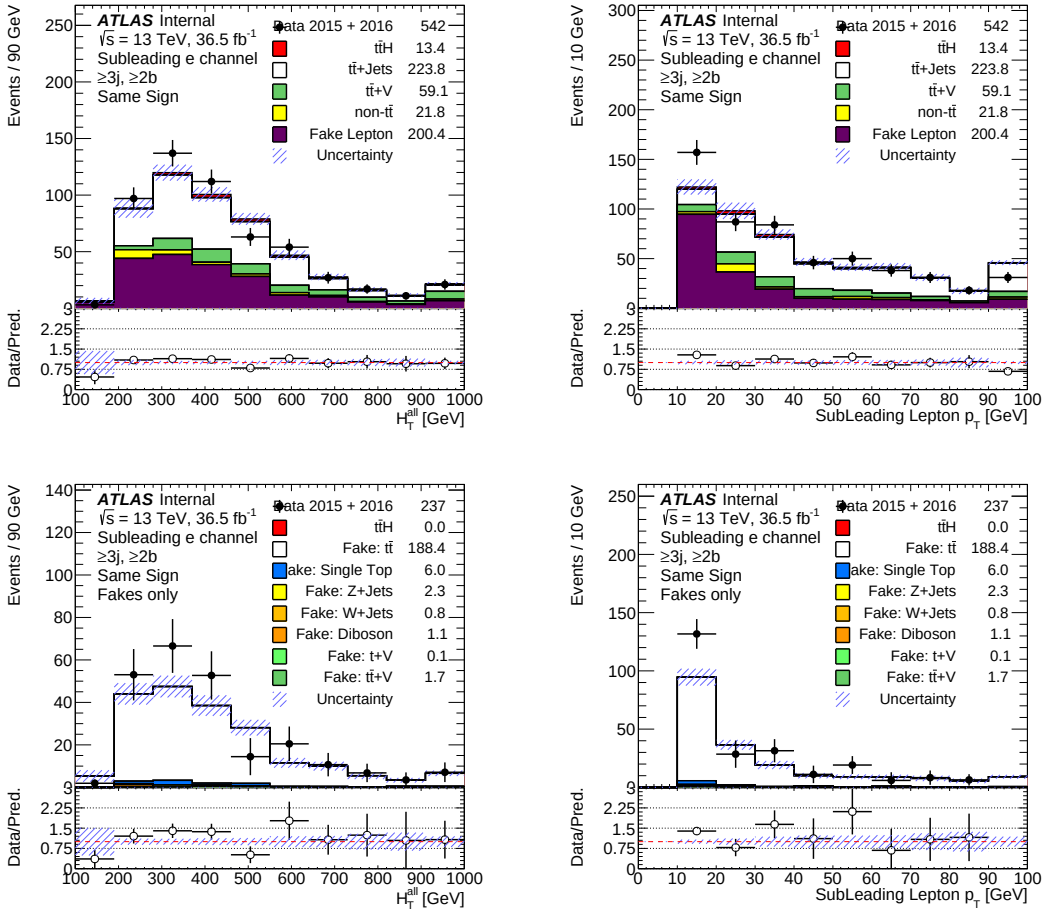


FIGURE 6.8: Comparison of data and simulated data for the subleading-electron channel in the same-sign region, using 36.5 fb^{-1} of 2015+2016 data. The distributions shown are the scalar sum of jet and lepton p_{T}, labelled $H_{\mathrm{T}}^{\mathrm{all}}$, and the p_{T} of the subleading electron. The direct comparison of data against simulation is shown in the top row of figures, while the comparison of fake-category simulated events against real data with prompt-category simulation subtracted (corresponding to the quantity $(N^{\mathrm{data,SS}} - N^{\mathrm{MC,prompt,SS}})$) is shown in the bottom row of figures. Uncertainties shown are statistical only, where the hashed area is the total statistical uncertainty for simulated data and the uncertainty bars are the statistical uncertainty on data.

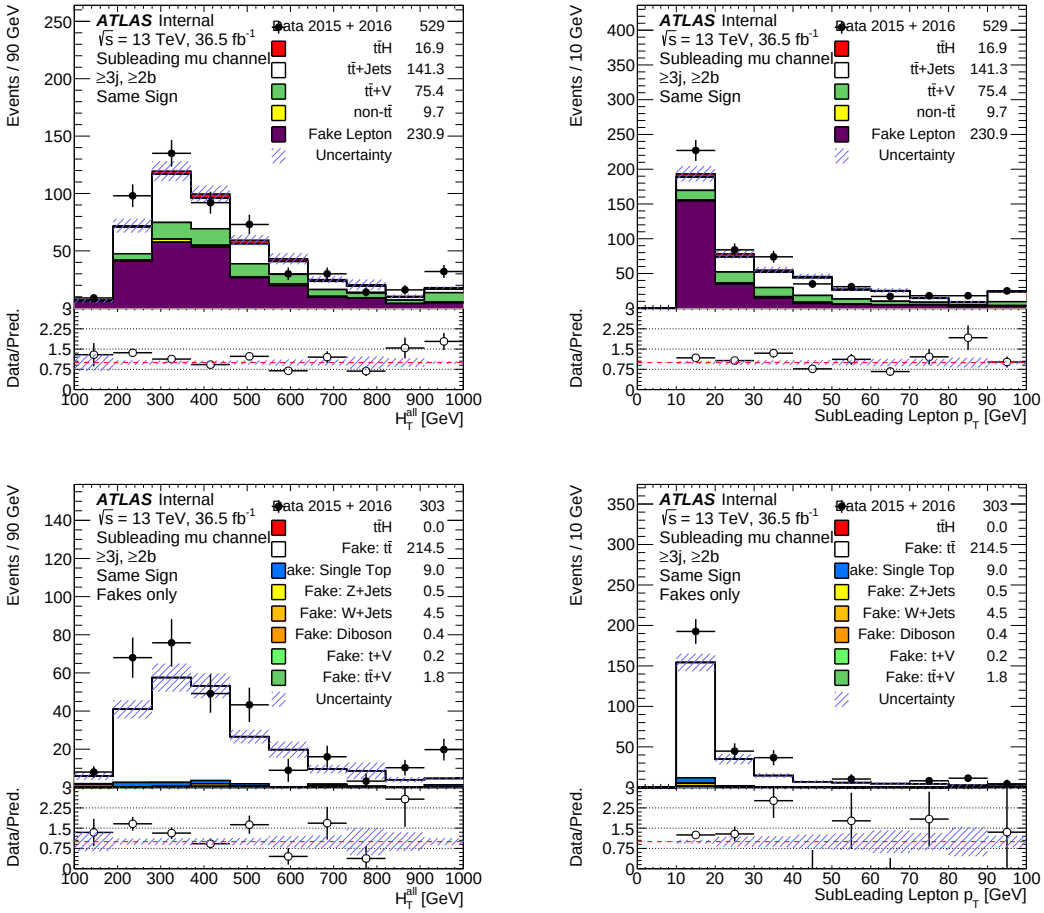


FIGURE 6.9: Comparison of data and simulated data for the subleading-muon channel in the same-sign region, using 36.5 fb^{-1} of 2015+2016 data. The distributions shown are the scalar sum of jet and lepton p_{T}, labelled $H_{\mathrm{T}}^{\mathrm{all}}$, and the p_{T} of the subleading muon. The direct comparison of data against simulation is shown in top row of figures, while the comparison of fake-category simulated events against real data with prompt-category simulation subtracted (corresponding to the quantity $(N^{\mathrm{data,SS}} - N^{\mathrm{MC,prompt,SS}})$) is shown in the bottom row of figures. Uncertainties shown are statistical only, where the hashed area is the total statistical uncertainty for simulated data and the uncertainty bars are the statistical uncertainty on data.

TABLE 6.5: Fake-category event yields for the same-sign region, split into subleading-electron and subleading-muon channels and using 36.5 fb^{-1} of 2015+2016 data. 'Data' refers to the quantity $(N^{\text{data, SS}} - N^{\text{MC, prompt, SS}})$, where N is the yield and the superscripts refer to the source of events. Processes other than data are from the appropriate simulated event datasets. 'Total' is the sum of all simulated events for the corresponding channel. Uncertainties are statistical only.

Process	Yield, subleading-e channel	Yield, subleading-μ channel
$t\bar{t}$+jets	188.4 $\pm$ 10.5	214.5 $\pm$ 13.2
Single-top	6.0 $\pm$ 1.1	9.0 $\pm$ 1.8
Z+jets	2.3 $\pm$ 1.0	0.5 $\pm$ 0.5
W+jets	0.8 $\pm$ 0.3	4.5 $\pm$ 2.1
$t\bar{t}$ +V	1.7 $\pm$ 0.2	1.8 $\pm$ 0.2
Diboson	1.1 $\pm$ 0.9	0.4 $\pm$ 0.2
t+V	0.13 $\pm$ 0.04	0.15 $\pm$ 0.03
Total	200.4 $\pm$ 10.7	230.9 $\pm$ 13.5
Data	237.3 $\pm$ 25.8	302.6 $\pm$ 24.4

normalisation factors applied as a direct modifier to the relevant yields. These normalisation factors are discussed below for the relevant uncertainties and in Section 6.7. As well, systematic uncertainties are defined to either affect both the normalisation of relevant processes and the shape of distributions analysed by the likelihood fit, or to only affect the normalisation of relevant processes. Systematic uncertainties are split into two categories: 'experimental uncertainties' consider uncertainties from reconstruction and measurement made by the ATLAS detector and reconstruction software, and are applied to all simulated processes. All experimental uncertainties except for the uncertainty on the measured luminosity affect both the normalisation of processes and the shape of their distributions. 'Modelling uncertainties' consider uncertainties from the theoretical modelling of different processes. Most modelling uncertainties affect both the normalisation and shape of distributions for the relevant processes, although cross-section and normalisation uncertainties affect only the normalisation. However, cross-section and normalisation uncertainties can affect the shape of the total prediction from simulation due to the fact that each simulated process will follow a different distribution shape. Changing the normalisation of a process will therefore change the contribution to the shape of the total prediction.

A number of uncertainties are defined by comparing results obtained with nominal

datasets to results obtained from relevant alternative datasets simulating the same process. In most of these cases, unless noted otherwise, the alternative sample used the faster AFII detector simulation as opposed to the full GEANT4 simulation. This is due to the tight computational budget of the shared ATLAS simulation resources. In these cases systematic uncertainties are defined by comparing a version of the nominal simulated dataset which instead used the AFII detector simulation to the relevant AFII alternative dataset. The difference in the shapes of distributions and normalisation between the datasets can then be taken as the uncertainty estimate and propagated to the distribution shape or yield of the nominal dataset which uses full detector simulation. In this way any differences resulting from using AFII instead of the full simulation should be avoided.

6.6.1 Experimental Uncertainties

Luminosity and Pile-up

The first source of experimental uncertainty is from the measurement of the total integrated luminosity of the recorded 2015+2016 data. This uncertainty is a single total normalisation uncertainty applied to all simulated processes. The size of this uncertainty is 2.1%. This value is derived using the method of 'van der Meer' scans, following a methodology similar to what is described in Reference [129]. In this method measurements of interaction rate are performed during dedicated LHC runs where the x-y separation of the two LHC beams is varied. By observing the variation in interaction rate the scale of luminosity measurements can be calibrated. The value for this uncertainty was derived from LHC runs in August 2015 and May 2016.

An uncertainty is also included associated with the pile-up reweighting correction factors applied to simulated events. When these weights are derived from a comparison of simulation to a measurement of data, there is an associated uncertainty in the ratio of simulation to data. This uncertainty is included by recalculating pile-up event weights with an upward and a downward variation of this ratio, according to its uncertainty. The 'up' and 'down' pile-up event weights can then be applied to simulation in place of the nominal pile-up weight, resulting in shape and normalisation uncertainties on all simulated processes. This variation is defined to cover the uncertainty in the ratio of

prediction to measurement for inelastic cross-sections in the fiducial volume defined by $M_X > 13$ GeV, where M_X is the mass of the hadronic system [130].

Jet Reconstruction

A number of uncertainties are associated with the calibration of jet energy scale and jet energy resolution, as well as the selection of jets based on the jet vertex tagger. The jet energy scale (JES) uncertainties are derived from test-beam data, LHC collision data, and simulation [70], and are divided into uncertainties associated with JES calibration methods and uncertainties from other sources. The uncertainties from calibration are originally defined as 75 uncertainties associated with event topology, MC simulation, dataset statistics, and propagated uncertainties of the electron, muon, and photon energy scales, and are split between the methods used to derive JES calibrations. However a simplified set of 8 uncertainties is instead used by the $t\bar{t}H(b\bar{b})$ analysis[2]. These uncertainties are defined by globally decomposing the covariance matrix for the combination of all uncertainties into eigenvector components. The 7 eigenvectors with the largest magnitude are then defined as the first 7 uncertainties, while the remaining eigenvectors are combined in quadrature to produce the 8th uncertainty. The reduction in correlation information when using this simplified uncertainty set is small.

The JES uncertainties from other sources are associated with: modelling of pile-up corrections, jet calibration in high eta regions, high-p_{T} jets, and punch-through of jets. As well, there are uncertainties associated with modelling of jet response for different flavours of jet, split into uncertainties for b-jets and for other jets. Finally, there is an uncertainty on the jet flavour composition of light-jets with regards to whether they are initiated by light-quarks or by gluons. JES uncertainties are not large for an individual jet, roughly 1% to 6% depending on jet p_{T}, however the uncertainties can be large overall due to the number of jets considered in the $t\bar{t}H(b\bar{b})$ analysis.

A combination of uncertainties on measured jet energy resolution are included as a single overall uncertainty. This uncertainty is then split into uncertainties for different analysis regions, resulting in two jet energy resolution uncertainties. The details of this definition are described in Section 6.7.4.

[2]The cited paper, [70], refers to 6 uncertainties, however the 36.1 fb^{-1} $t\bar{t}H(b\bar{b})$ analysis uses a recommendation which uses the same methodology but instead defines 8 uncertainties. This change in the number of uncertainties simply corresponds to a later derivation of JES uncertainties.

Finally an uncertainty is placed on the jet vertex tagger associated with the efficiency of a jet to pass the JVT requirement.

Jet Flavour Tagging

Uncertainties associated with b-tagging and related mis-tagging of c- and light-jets as b-tags are derived from the analyses used to calibrate the b-tag efficiencies and mis-tag rates [74, 76]. Similar to what is done for JES calibration uncertainties, the various sources of systematic uncertainty from these analyses are decomposed into eigenvariations components by building the covariance matrix with all uncertainties and then performing eigenvariation decomposition on this matrix. These uncertainties are defined separately for b-jet b-tags, c-jet mis-tags, and light-jet mis-tags. This results in 30 b-tag uncertainties, 20 c-jet mis-tag uncertainties, and 60 light-jet mis-tag uncertainties. The uncertainties are ordered by the magnitude of their corresponding eigenvariation. The size of uncertainties for b-tagging b-jets varies between 2% and 10% per jet; for uncertainties on mis-tagging c-jets the uncertainties vary between 5% and 20% per-jet; and for mis-tagging light-jets the uncertainties vary between 10% to 50% per-jet.

As well there is an additional uncertainty associated with the inclusion of hadronic tau jets in c-jet mis-tag rates. Hadronic tau jets are treated as c-jets when deriving and measuring c-jet mis-tag rate corrections and uncertainties. This uncertainty covers the extrapolation between c-jets and these τ-jets.

Lepton Reconstruction

Uncertainties associated with electrons and muons are taken from the appropriate comparisons between data and simulation [47, 60]. There are uncertainties associated with efficiencies of the lepton reconstruction, trigger, identification, and isolation. As well, there are uncertainties associated with lepton momentum scale and resolution. Muon uncertainties are split into more detailed components reflecting the combination of separate inner detector and muon spectrometer measurements, for example with separate energy resolution uncertainties for the ID measurement and MS measurement. The effects of lepton uncertainties are small, approximately 1% or less.

Missing Transverse Momentum

Uncertainties on lepton and jet energy scales and resolutions described above are propagated into the reconstruction of missing transverse momentum. Specific uncertainties on the soft term of the missing transverse momentum are also included. An uncertainty on the momentum scale of the soft term is included, and two uncertainties corresponding to the parallel and perpendicular components of the soft term momentum resolution relative to the direction of the missing transverse momentum vector calculated from all contributions except the soft term [79]. Missing transverse momentum uncertainties have only small effects on the analysis, reflecting the fact that missing transverse momentum is not used for event selection.

6.6.2 Modelling Uncertainties

$t\bar{t}H$

Two uncertainties are placed on the cross-section of the $t\bar{t}H$ process, originating from uncertainty on the QCD scales and from the uncertainty on PDF+α_S, labelled as XS(scale) and XS(PDF) respectively [31, 95–99]. The XS(scale) uncertainty is $^{+5.8\%}_{-9.2\%}$ and the XS(PDF) uncertainty is $\pm 3.6\%$. The effect of PDF variations on the shape of distributions as well as changes in analysis selection acceptance is studied in Section 6.6.3. Similar effects from QCD scale variations were checked by other members of the ATLAS analysis group using similar methodology and were found to be negligible, and therefore were not included in the analysis. Uncertainties on the branching ratio of the Higgs decay modes are also considered [31]. An uncertainty is also defined for the theoretical modelling of parton showering and hadronisation by comparing results from the nominal MG5_aMC+PYTHIA 8 dataset to an alternative $t\bar{t}H$ dataset which uses the same MG5_aMC matrix element generator, but is instead interfaced with the HERWIG++ parton shower package [122]. The MG5_aMC+HERWIG++ dataset is simulated using the full detector simulation and so the uncertainty is directly defined as the difference in distribution shapes and normalisation between the two datasets.

$t\bar{t}$+jets

As described in Section 6.2, theoretical modelling of $t\bar{t}$+HF processes is challenging, and associated uncertainties can completely obscure signal yield. As a result handling of $t\bar{t}$ modelling uncertainties, particularly with respect to $t\bar{t}$+HF, is crucial to allow for constraint of these uncertainties in the statistical analysis as well as correction of mismodelling. Before the systematic uncertainty model is described, a brief description is given of the challenges involved in simulating $t\bar{t}+ \geq 1b$ processes, so as to motivate the origins and magnitudes of uncertainties and mismodelling.

$t\bar{t}+ \geq 1b$ production is difficult to model due to the necessary modelling of higher order QCD proceses, the modelling of gluons splitting into b quarks and additional gluon production, and the involvement of the mass of modelled b quarks. Each of these features originates its own set of problems and challenges, either at the level of theoretical calculations or at the level of producing simulation packages using these calculations. An additional challenge is the small number of experimental measurements of $t\bar{t}$+HF processes.

Firstly, the fact that leading order diagrams for $t\bar{t}b\bar{b}$ processes involve α_S^4 QCD dependence and a number of final-state particles means that theoretical calculations are prone to large uncertainties from QCD scale variations (up to 80%), as well as being challenging calculations in their own right. To reduce these large scale uncertainties NLO calculations must be done, which reduce the scale uncertainties to a maximum of 20% if used. In dedicated $t\bar{t}b\bar{b}$ simulations, such as the SHERPA4F dataset, these processes are indeed simulated at NLO in the matrix element. However for inclusive $t\bar{t}$ datasets, such as the POWHEG+PYTHIA 8 dataset, the NLO matrix element calculations only cover up to leading order for $t\bar{t}b$ and $t\bar{t}g$ final states.

Next is the issue of modelling gluon splitting and additional[3] gluon production. $t\bar{t}b\bar{b}$ simulations include calculation of an additional gluon which splits into a $b\bar{b}$ pair, and so this is consistently handled within the matrix element calculation. However inclusive $t\bar{t}$ samples include matrix element calculations up to one additional gluon but not the decay of the gluon into b quarks. As a result the actual production of additional b quarks must instead be handled by the parton shower calculation. As well, events containing

[3]Here 'additional' refers to any gluons beyond leading order $t\bar{t}$ production. This includes the $g \to b\bar{b}$ gluon in leading order $t\bar{t}b\bar{b}$ matrix elements.

two additional gluons which produce collimated $b\bar{b}$ pairs that are reconstructed as two B-jets can produce a yield comparable to that of the $t\bar{t}H(b\bar{b})$ signal. The $t\bar{t}b\bar{b}$ simulations can calculate up to the second additional gluon within the matrix element, however the decay of the second gluon must be handled by the parton shower in a similar way to the first additional gluon in inclusive $t\bar{t}$ samples. On the other hand inclusive $t\bar{t}$ samples must calculate the emission and splitting of the second gluon completely within parton shower calculation. Indeed any additional gluons beyond the first (second) additional gluon must be handled entirely within the parton shower calculation for inclusive $t\bar{t}$ ($t\bar{t}b\bar{b}$) simulations. As a result both categories of simulation are sensitive to parton shower tuning and modelling, even for the first (second) additional gluon in inclusive $t\bar{t}$ ($t\bar{t}b\bar{b}$) simulations due to the parton shower modelling of the splitting of the gluon, and similarly are sensitive to the matching used to combine the matrix element and parton shower simulations within a dataset.

Lastly is the issue of the mass of modelled b quarks. This issue affects results from $t\bar{t}b\bar{b}$ simulations related to the choice of either 5FS or 4FS flavour schemes; because the $t\bar{t}H(b\bar{b})$ analysis only employs a 4FS $t\bar{t}b\bar{b}$ simulation this issue is not directly relevant; however, the importance of the 4FS choice is. When b quarks are treated as massless, as in 5FS simulations, regions of phase space with collinear b quark emissions must be cut away to avoid collinear singularities in the calculations, and instead the b quark emissions must be performed by the parton shower simulation. Including the b quark mass in the matrix element calculations, as in 4FS simulations, allows the full phase space to be modelled, including the production of a single B-jet from a collinear $g \rightarrow b\bar{b}$ splitting. However as described previously, splittings of additional gluons to b quarks must still be handled by the parton shower.

Cutting-edge discussions on the topic of simulating $t\bar{t}+ \geq 1b$ can be found in Reference [131], as well as a similar older paper describing the SHERPA4F dataset used in this analysis [123], and chapter I.6.8 of Reference [31].

The systematic uncertainties for $t\bar{t}$+jets are broadly split into three categories: overall normalisation uncertainties; modelling uncertainties which do not affect $t\bar{t}$ category and $t\bar{t}+ \geq 1b$ sub-category relative fractions; and modelling uncertainties which affect only the relative fractions of $t\bar{t}+ \geq 1b$ sub-categories. As well, some uncertainties in the second category are split into three uncertainties affecting only $t\bar{t}$+light, $t\bar{t}+ \geq 1c$, and

$t\bar{t}+\geq 1b$ events, respectively. This is done to reflect the specific modelling challenges of simulating $t\bar{t}$+HF processes, and therefore to allow these processes to be constrained and corrected individually.

The first overall normalisation uncertainty is a single inclusive $t\bar{t}$ cross-section uncertainty of $\pm 6\%$ [102]. This uncertainty is associated with the cross-section uncertainty of the NNLO+NNLL normalisation of the inclusive $t\bar{t}$ calculations, and covers effects from variation of the QCD scales, PDF choice, α_s value, and top-quark mass. There are then two additional normalisation uncertainties associated specifically with $t\bar{t}+\geq 1c$ and $t\bar{t}+\geq 1b$ events. Each category has its own overall normalisation uncertainty. However, the handling of these uncertainties is different to other cross-section or normalisation uncertainties in the $t\bar{t}H(b\bar{b})$ analysis. Due to large discrepancies between the normalisation of data and simulated data in analysis regions which are pure in $t\bar{t}$+HF, no prior uncertainty estimate is assigned to the $t\bar{t}+\geq 1c$ and $t\bar{t}+\geq 1b$ normalisation uncertainties, and they are instead allowed to be 'free-floating' in the likelihood fit. This means that the fit can scale the prior normalisation of $t\bar{t}+\geq 1c$ and $t\bar{t}+\geq 1b$ simulated events without any direct penalty, a technicality that is described in further detail in Section 6.7. This is done to ensure that the normalisation of simulated $t\bar{t}$+HF events can be corrected without using the uncertainties affecting $t\bar{t}$+HF which are described below, which could result in associated changes to distribution shapes.

For modelling uncertainties which do not affect the $t\bar{t}$ category and $t\bar{t}+\geq 1b$ sub-category relative fractions, all uncertainties are defined by comparisons of datasets simulated using different parameter choices or simulation packages. To ensure that $t\bar{t}$ category and $t\bar{t}+\geq 1b$ sub-category relative fractions do not change in these uncertainty estimates, all datasets are weighted to have the same $t\bar{t}$+light, $t\bar{t}+\geq 1c$, and $t\bar{t}+\geq 1b$ relative fractions as the nominal POWHEG+PYTHIA 8 $t\bar{t}$ dataset, as well as being weighted to have the same $t\bar{t}+\geq 1b$ sub-category relative fractions as in the SHERPA4F dataset. The exception to the SHERPA4F $t\bar{t}+\geq 1b$ sub-category weighting is again the $t\bar{t}+b$(MPI/FSR) category, meaning that the contribution of this category can vary due to these uncertainties. The additional $t\bar{t}$ datasets are described below:

- **SHERPA5F**: inclusive $t\bar{t}$ dataset simulated using SHERPA 2.2.1, where the matrix element calculation is merged with the SHERPA parton shower calculation using the

ME+PS@NLO setup, which is interfaced with OPENLOOPS. This gives NLO accuracy for one additional parton and LO accuracy for four additional partons. The NNPDF3.0NNLO PDF set was used, and renormalisation and factorisation scales were both set to $\sqrt{0.5 \times (m^2_{\mathrm{T},t} + m^2_{\mathrm{T},\bar{t}})}$, where $m_{\mathrm{T},t} = \sqrt{m_t^2 + p^2_{\mathrm{T},t}}$ and $p_{\mathrm{T},t}$ is the transverse momentum of the top quark in the $t\bar{t}$ centre-of-mass reference frame, and $m_{\mathrm{T},\bar{t}}$ is similarly defined for the anti-top quark. Note that this dataset is not a dedicated $t\bar{t}b\bar{b}$ simulation unlike the similarly named SHERPA4F dataset;

- **POWHEG+HERWIG 7**: inclusive $t\bar{t}$ dataset simulated using POWHEG interfaced with the HERWIG 7 parton shower and hadronisation generator [132]. The settings for POWHEG are identical to the POWHEG+PYTHIA 8 dataset, while HERWIG 7 uses the H7-UE-MMHT set of tuned parameters;

- **RadHi**: inclusive $t\bar{t}$ dataset which is identical to the POWHEG+PYTHIA 8 dataset except for the renormalisation and factorisation scales being halved, the h_{damp} parameter doubled, and the Var3c upward variation of the A14 tuned parameter set being used [133]. This dataset, along with the RadLo dataset, allows estimation of uncertainty associated with variations in the amount of predicted initial state and final state radiation;

- **RadLo**: inclusive $t\bar{t}$ dataset which is identical to the POWHEG+PYTHIA 8 dataset except for the renormalisation and factorisation scales being doubled, and the Var3c downward variation of the A14 tuned parameter set being used;

- **MG5_aMC+HERWIG++ 3F**: $t\bar{t}c\bar{c}$ dataset used in the $t\bar{t}+ \geq 1c$ uncertainty estimation, see Reference [134];

- **MG5_aMC+HERWIG++ 5F**: inclusive $t\bar{t}$ dataset used in the $t\bar{t}+ \geq 1c$ uncertainty estimation, see Reference [134].

The first uncertainty of this type is defined by comparing results from using the nominal $t\bar{t}$ dataset to results from using the SHERPA5F dataset. In this way uncertainties in the matrix element prediction and parton shower prediction are simultaneously estimated. The second uncertainty of this type is defined by comparing results from using the nominal $t\bar{t}$ dataset to results from using the POWHEG+HERWIG 7 dataset. In this

way an uncertainty on only the parton shower prediction is estimated. The third uncertainty of this type is an asymmetric uncertainty defined by comparing results from the nominal $t\bar{t}$ dataset to results from the RadHi and RadLo datasets, which respectively correspond to the 'upward' and 'downward' directions of this uncertainty estimation. These three uncertainties are decorrelated between $t\bar{t}$+light, $t\bar{t}+ \geq 1c$, and $t\bar{t}+ \geq 1b$, resulting in 9 separate uncertainties. There is then another uncertainty which is applied only to $t\bar{t}+ \geq 1c$ events. This uncertainty is defined by comparing results from using the MG5_aMC+HERWIG++ 3F dataset to results from using the MG5_aMC+HERWIG++ 5F dataset. The difference between these datasets is the $t\bar{t}c\bar{c}$ matrix element calculation performed in the former dataset, which includes a non-zero c quark mass and excludes c quarks and b quarks from consideration as partons in the PDFs, hence using the three flavour scheme in a similar manner to the four flavour scheme of the SHERPA4F dataset. The difference between these two datasets therefore shows the difference in modelling additional c-jets in the matrix element calculation as opposed to in the parton shower calculation. The difference between the two datasets is taken and then applied to $t\bar{t}+ \geq 1c$ events in the nominal $t\bar{t}$ dataset. Finally there is another uncertainty which is applied only to $t\bar{t}+ \geq 1b$ events. This uncertainty is defined by comparing results from the nominal $t\bar{t}$ dataset to results from the SHERPA4F dataset. In this way the difference between modelling additional b-jets in the matrix element calculation as opposed to in the parton shower calculation is estimated. This uncertainty is not applied to $t\bar{t} + b$(MPI/FSR) events.

For modelling uncertainties which only affect the $t\bar{t}+ \geq 1b$ sub-category relative fractions, all uncertainties are defined by variations of the SHERPA4F dataset. Three scale uncertainties are defined where the first varies the renormalisation scale up and down by a factor of two, the second changes the functional form of the resummation scale to μ_{CMMPS}, and the third sets a global scale choice of $\mu_{\text{Q}} = \mu_{\text{R}} = \mu_{\text{F}} = \mu_{\text{CMMPS}}$. The effects of PDF set choice are considered by comparing results using the nominal CT10 4F set againt results using the MSTW2008NLO set [135], and also against the NNPDF2.3NLO set [136]. Uncertainties are also defined by comparing against an alternative shower recoil scheme (the procedure used to handle conservation of momentum when the emission of a particle occurs in a parton shower calculation) and an alternative set of tuned parameters for the underlying event. These uncertainties are used to define the uncertainty band in

Figure 6.7 from Section 6.5. As well, as seen in Figure 6.7 there is a large discrepancy between the relative predictions of the POWHEG+PYTHIA 8 dataset and the SHERPA4F dataset for the $t\bar{t}+ \geq 3b$ category. This difference is not covered by the uncertainties described above, and so a normalisation uncertainty of 50% is placed on $t\bar{t}+ \geq 3b$ events.

Finally, a normalisation uncertainty of 50% is placed on events with contributions specifically from multiparton interactions, based on previous comparisons of underlying event sets of tuned parameters. As previously mentioned, the $t\bar{t} + b$(MPI/FSR) sub-category fraction as a whole can be varied by the matrix element simulation, parton shower simulation, and radiation variation uncertainties estimated from alternative datasets.

V+jets

A 35% normalisation uncertainty is placed on the Z+jets process. This uncertainty is decorrelated between 3 jet and $\geq$ 4 jet dilepton analysis regions, resulting in two uncorrelated normalisation uncertainties. This uncertainty covers the combination of two effects: firstly the size of the Z+HF correction factor applied to Z+jets events, and secondly QCD scale variations estimated within the nominal Z+jets dataset. These scale variations are estimated by separately scaling the renormalisation and factorisation scale values by 2 and $\frac{1}{2}$, where the effect is applied through recalculated event weights within the nominal sample. Scale variations are of the order $< 10\%$.

Similar normalisation uncertainties are used for the W+jets process in the single-lepton channel; however, in the dilepton channel because the W+jets process only contributes events which contain fake or non-prompt leptons, an overall normalisation uncertainty for such events from all processes is applied (described below), and no specific W+jets uncertainties are used in the dilepton analysis.

Single-top

The Wt single-top process has a cross-section uncertainty of $^{+5\%}_{-4\%}$ [117]. As well uncertainties associated with the theoretical modelling of parton shower and hadronisation and the modelling of initial- and final-state radiation are included using a method similar to what is done for the $t\bar{t}$+jets process, where results from the nominal dataset

are compared to results from appropriate alternative datasets. For the parton shower uncertainty the nominal POWHEG+PYTHIA 6 dataset is compared with an equivalent POWHEG+HERWIG++ dataset. For the radiation uncertainty the nominal dataset is compared with POWHEG+PYTHIA 6 datasets where the relevant factorisation and renormalisation scale variations are made to either increase or decrease additional radiation. Finally an uncertainty on the amount of interference between Wt and $t\bar{t}$ production is defined by comparing the nominal dataset, which uses the 'diagram removal' method, to a dataset which instead uses the 'diagram subtraction' method [125]. This dataset is produced using the full detector simulation, and so the uncertainty can be directly defined by comparing the nominal dataset to this alternative dataset.

s-channel and t-channel single-top processes also have a cross-section uncertainty of $^{+5\%}_{-4\%}$ [118, 119], and the t-channel process also has similar uncertainties on parton shower modelling and initial- and final-state radiation. However these uncertainties are not applied to the s-channel and t-channel events which contain fake or non-prompt leptons, and so are not included in the dilepton analysis.

$\mathbf{t\bar{t}+V}$

The $t\bar{t}Z$ process has cross-section uncertainties split into QCD scale and PDF+α_S uncertainties: the XS(scale) uncertainty is $^{+9.6\%}_{-11.3\%}$ and the XS(PDF) uncertainty is $\pm$4% [31, 137]. The $t\bar{t}W$ process has similarly split cross-section uncertainties: the XS(scale) uncertainty is $^{+12.9\%}_{-11.5\%}$ and the XS(PDF) uncertainty is $\pm$3.4% [31, 137]. Both processes also have an uncertainty associated with matrix element and parton shower modelling. This is defined by comparing results from the nominal MG5_aMC+PYTHIA 8 datasets to results from datasets generated using SHERPA [114]. Due to SHERPA employing its own parton showering estimation, this simultaneously compares variations in matrix element and parton shower predictions.

Events containing fake or non-prompt leptons

A single normalisation uncertainty of 25% is placed on dilepton channel events containing fake or non-prompt leptons. This value is chosen to cover the size of the correction factors described in Section 6.5.1, such that the original estimate from simulation is within

the size of this uncertainty[4]. No shape uncertainty is included due to the generally good shape agreement between data and simulated data. Modelling uncertainties described elsewhere in this section are not applied to the relevant processes within the fake or non-prompt lepton estimate. In principle they should be; however changes from doing so would be small due to the small size of the fake or non-prompt lepton background in all distribution bins.

The single-lepton channel fake or non-prompt lepton estimate uses a 50% normalisation uncertainty split between events containing an electron and events containing a muon, and also split between 5 jet and ≥ 6 jet regions.

Other processes

Diboson processes have a single 50% normalisation uncertainty, which covers uncertainties on the cross-section and additional jet production [120]. A single 50% normalisation uncertainty is also used for the $t\bar{t}t\bar{t}$ process. This is used due to the presence of $t\bar{t}t\bar{t}$ events in signal-rich analysis regions and distribution bins; 50% is taken as a conservative estimate to ensure that the uncertainty on $t\bar{t}t\bar{t}$ is not underestimated when measuring $t\bar{t}H$ signal. The tZ [103], $t\bar{t}WW$ [103], $tHjb$, and WtH processes have cross-section uncertainties split between QCD scale and PDF. The size of these uncertainties ranges from approximately 15% down to approximately 1%. The tZW processes has a single normalisation uncertainty of 50% (this is a preliminary estimate defined without reference to theoretical calculations).

6.6.3 Estimation of Acceptance and Shape Uncertainties Arising from PDF Choice

As described above a number of uncertainties on the cross-section of processes, as well as on the relative fractions of $t\bar{t}+ \geq 1b$ sub-categories, are defined based on the modelling of PDFs used in the simulated datasets. However, changes in modelling of PDFs can also result in differences in overall normalisation due to greater or fewer numbers of events falling into acceptance. Shapes of distributions can also be affected. These

[4]The uncertainty is defined relative to the new yield after the correction factors are applied, and so the uncertainty appears smaller than the roughly 30% increase on the original yield from application of the largest fake scale factor, despite actually being of an equivalent size.

acceptance and shape uncertainties can therefore also be estimated and included as separate uncertainties from the PDF uncertainties on cross-sections and relative fractions of $t\bar{t}+ \geq 1b$ sub-categories. This estimation was performed for the 13.2 fb^{-1} analysis, and therefore all work described in this section can be assumed to refer to the 13.2 fb^{-1} analysis unless stated otherwise. A few important differences exist between the 13.2 fb^{-1} analysis definitions and the 36.1 fb^{-1} analysis definitions used elsewhere in this thesis; these differences are highlighted when relevant. The use of shape and acceptance uncertainties arising from PDF choice in the 36.1 fb^{-1} analysis is discussed at the end of this section. Uncertainties are estimated for the $t\bar{t}H$ and $t\bar{t}$ processes. Results were obtained for both the dilepton and single-lepton analysis channels, however results in figures will be shown for the dilepton channel only.

As previously mentioned, parton distribution functions represent the probability density of finding a specified flavour of parton at values of the momentum fraction of total proton momentum, x, given an interaction scale, Q^2. This covers the valence uud quarks of the proton and the sea of quarks, anti-quarks, and gluons produced from the binding energy of the proton. Through the use of PDFs, cross-section calculations for proton–proton cross-sections can be performed by factorising the calculation in terms of the cross-sections of partonic interactions multiplied by the relevant PDFs, and then integrating over the x values of the partons and summing over parton flavours. This cross-section equation is shown in Equation 6.2. The labels 1 and 2 refer to the two colliding protons, i and j refer to the flavours of partons, f_i and f_j are the PDFs for the relevant partons, and $\sigma_{i,j}$ is the partonic cross-section.

$$\sigma_{pp\to X} = \sum_{i,j} \iint \mathrm{d}x_1 \mathrm{d}x_2 \; f_i(x_1, Q^2)\, f_j(x_2, Q^2)\, \sigma_{i,j} \tag{6.2}$$

The dependence of the PDFs on Q^2 is included in the theory of QCD using equations known as the DGLAP equations[5] [138–140], for a Q^2 which is hard enough to allow perturbative QCD calculations. However, the dependence of the PDFs on x is not known *a priori*, and instead must be determined from experimental measurements. PDFs are universal (for a given hadron, in this case for protons), which allows PDF measurements

[5]Named after the initials of the contributing authors

from many different experiments to be combined to produce fitted PDF sets. These measurements yield data points at particular x values for particular parton flavours, at a given Q^2 value; these measurements can then be extrapolated to other Q^2 values using the DGLAP equations. The PDF sets mentioned when previously discussing configurations of simulated datasets and systematic uncertainties, for example the NNPDF3.0NLO set, are sets of PDFs produced by different research teams and differ by the set of input data measurements (though many of same measurements are common across different PDF sets), input theoretical calculations, methods of parametrising the PDFs, and the details of the parametrisations. The PDF sets each have their own associated uncertainties, for example arising from systematic uncertainties of the input data measurements and from the choices of parametrisation used. However, the comparison of results from the different PDF sets is also of interest. Given the use of similar input data it can be expected that different PDF sets should result in similar nominal predictions and similar sizes of uncertainty; however it is seen that even with the latest available PDF sets there are still differences in these results. As such, PDF uncertainties from one set alone may be an underestimation of the true size of PDF uncertainties, and uncertainties should instead be estimated by considering multiple PDF sets.

To this end the PDF4LHC group compares multiple PDF sets and produces recommendations based on statistical combinations of these sets, with the most recent known as the 'PDF4LHC15' recommendations [141]. To produce these recommendations, the ensemble of the latest available PDF sets was examined and comparisons were made between the inputs and configurations used for each set. When forming the statistical combination of sets, only those that pass a set of compatibility requirements were considered:

- firstly, that the sets used a large number of datasets from a variety of experimental types;
- secondly, that cross-sections and evolution of α_s and PDFs were calculated at sufficient accuracy;
- thirdly, that all sets used the same central and variation values of $\alpha_S(m_Z^2)$;

- finally, that all known experimental and procedural sources of uncertainty were properly accounted for.

Three PDF sets passed these requirements, namely CT14 [142], MMHT2014 [143], and NNPDF3.0. These three PDF sets predict similar central values and sizes of uncertainties, however there are still some differences between them. The PDF4LHC15 group performed three statistical methods of combining these PDFs, where the first step of each combination was producing an equal number of Monte Carlo replicas of each PDF set. Of relevance to the studies in this section is the combination performed using the 'META–PDF' method [144], labelled '`PDF4LHC15_nlo_30`'. In this method a 'meta-parametrisation' functional form was used which can describe the parametrisations of each of the Monte Carlo replicas. This meta-parametrisation was then fitted individually to all Monte Carlo replicas, allowing the fit parameters to be averaged to produce a central 'meta-fit' PDF set and a covariance matrix to be constructed. This covariance matrix was then decomposed into an eigenvector representation and, similarly to what was previously discussed in JES and flavour-tagging uncertainties, the eigenvariations with largest variations could be selected as dominant PDF uncertainties. In this case the smaller eigenvariations are simply discarded, leaving 30 variations representing the 'minimal' set of uncertainties to accurately describe the combination of PDF sets. The two other combination methods provide more detailed uncertainty estimates, however the additional precision is not needed for estimation of acceptance uncertainties and speed is preferred in the statistical fit, meaning that the `PDF4LHC15_nlo_30` recommendation is the best choice. This recommendation provides one nominal PDF set plus 30 variation sets, where each is defined with a one sigma variation on one of the 30 eigenvariations.

To estimate the effect of PDF uncertainties on datasets in the $t\bar{t}H(b\bar{b})$ analysis, results must be obtained where the PDF sets used in the nominal datasets are replaced by the PDF sets from the `PDF4LHC15_nlo_30` recommendation. To avoid the computationally costly process of re-simulating datasets with the PDF changed, event weights are instead calculated using the LHAPDF software package [145]. These weights can then be applied to simulated events to mimic the effect of using a different PDF set, and comparisons can be made between nominal results and results with these weights applied. For the nominal MG5_aMC $t\bar{t}H$ dataset these weights can be calculated; however for the

POWHEG+PYTHIA 6 $t\bar{t}$ dataset used in the 13.2 fb^{-1} analysis the weights could not be calculated. POWHEG does not store the required parton information needed by the LHAPDF package to calculate the values of the different PDFs for the given partons. As such the PDF uncertainties must be estimated relative to a $t\bar{t}$ dataset which uses a different matrix element generator, and then the uncertainties can be placed around the nominal POWHEG predictions. This was chosen to be an MG5_aMC+HERWIG++ $t\bar{t}$ dataset. Both the nominal POWHEG dataset and the MG5_aMC dataset use the CT10 PDF set. The nominal MG5_aMC+PYTHIA 8 $t\bar{t}H$ dataset uses NNPDF3.0NLO.

Two types of PDF uncertainties can be defined: inter-PDF uncertainties, reflecting the choice of nominal PDF set, and intra-PDF uncertainties, reflecting the uncertainties within a chosen PDF set. To estimate the effect of inter-PDF uncertainties, nominal distributions can be compared to results reweighted to the `PDF4LHC15_nlo_30` central prediction, or central predictions for other specific PDF sets. In these studies the nominal results are compared to the `PDF4LHC15_nlo_30` central prediction as well as to the central predictions of the CT14, MMHT2014, and NNPDF3.0 sets. In this way the spread of inter-PDF predictions can also be observed. To estimate the effect of intra-PDF uncertainties the central `PDF4LHC15_nlo_30` prediction is compared to results obtained using the 30 variation PDF sets, and then the resulting uncertainties can be placed around the nominal dataset prediction. In these studies results from the individual variation sets are shown, as well as maximum and minimum envelopes representing the extrema of variations for the whole set of 30. In this way the magnitude of individual intra-PDF uncertainties could be estimated in an overall manner. When both types of uncertainties are estimated the original cross-section of the dataset in question is preserved. In this way any direct effects on the cross-section of processes are removed and are instead included in the previously mentioned cross-section systematic uncertainties. The effect of PDF uncertainties are estimated in the analysis regions used in the 13.2 fb^{-1} analysis which are defined using requirements on the number of jets and number of b-tags at the tight working point. One point of difference to studies described elsewhere which use similar region definitions is that the 13.2 fb^{-1} event selection requires ≥ 2 jets rather than ≥ 3 jets, which means that results obtained in the inclusive analysis region use this selection instead.

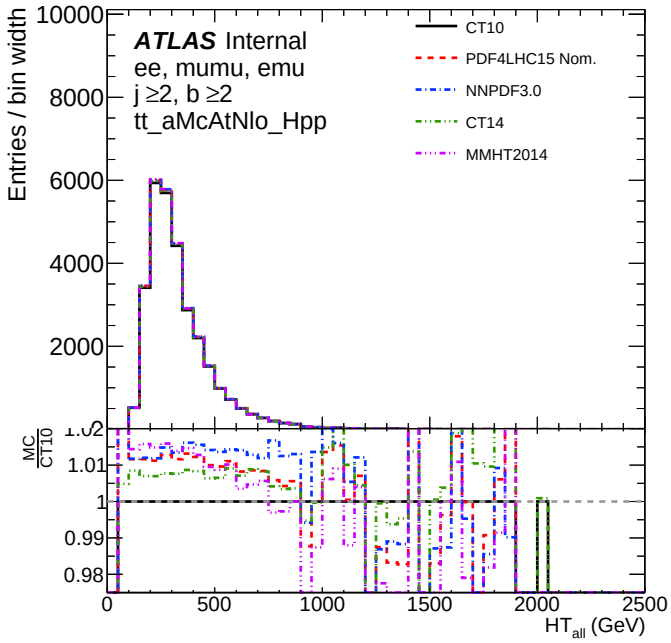


FIGURE 6.10: Comparison of the nominal distribution and distributions reweighted to inter-PDF variations for $t\bar{t}$. The distribution shown is the scalar sum of jet and lepton p_{T}, labelled $\mathrm{HT_{all}}$. The solid black line shows the nominal distribution and is labelled with the PDF set used in the nominal dataset, CT10. The other lines are predictions from the central values of other PDF sets. The ratio is defined with the nominal $t\bar{t}$ prediction in the denominator.

A histogram of a distribution reweighted to estimate the effect of inter-PDF uncertainties is shown in Figure 6.10 for the $t\bar{t}$ dataset, and in Figure 6.11 for the $t\bar{t}H$ dataset. Reweighting to the NNPDF3.0 PDF set is not performed in the $t\bar{t}H$ results due to the use of this PDF set in the nominal dataset. Histograms of distributions reweighted to estimate the effect of intra-PDF uncertainties are shown in Figure 6.12 for the $t\bar{t}$ dataset, and in Figure 6.13 for the $t\bar{t}H$ dataset. Results for intra-PDF uncertainties are shown for the max-min envelope of the 30 variations, and for the first five variations to demonstrate the shapes and magnitudes of individual variations. All distributions are shown in the inclusive $N_{\text{jets}} \geq 2$ AND $N_{b\text{-tags}}^{\text{tight}} \geq 2$ analysis region. Distributions are not normalised to a specific luminosity because the results of interest are the ratio of reweighted distributions from the same dataset, meaning that any scaling simply cancels out[6].

It is seen that inter-PDF variations for $t\bar{t}$ show a normalisation increase of approximately 1%, with small to negligible differences in distribution shapes. The three PDF sets which make up the PDF4LHC15 set are all within 1% of each other, which is the expected result. Other distributions related to event shape, such as pseudorapidity of

[6] As a result yields cannot be compared between figures for $t\bar{t}$ and $t\bar{t}H$, because the dataset cross-sections are not considered in the event weights.

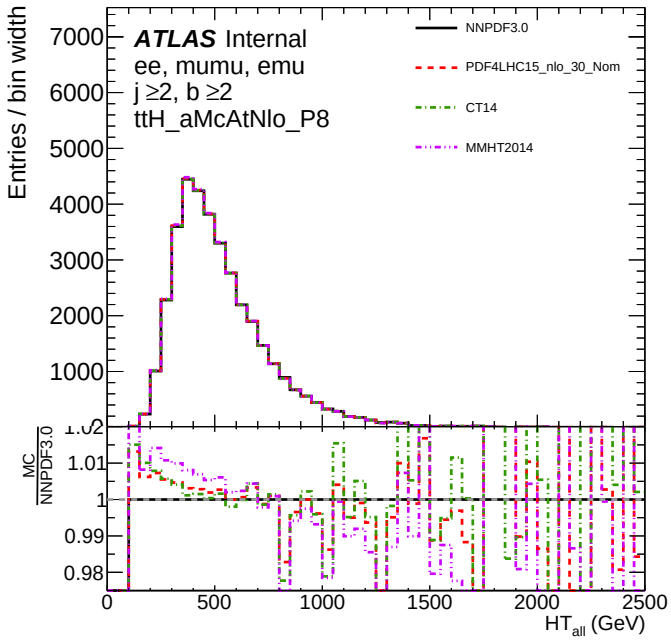
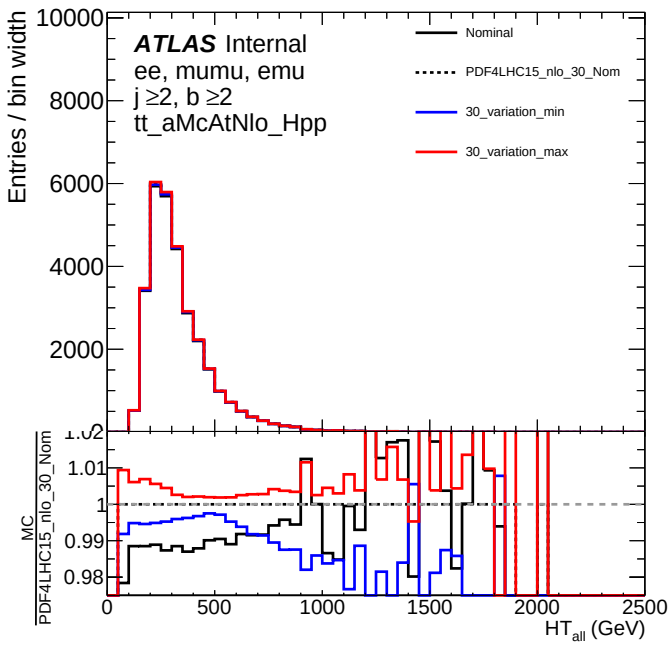


FIGURE 6.11: Comparison of the nominal distribution and distributions reweighted to inter-PDF variations for $t\bar{t}H$. The distribution shown is the scalar sum of jet and lepton p_{T}, labelled $\mathrm{HT_{all}}$. The solid black line shows the nominal distribution and is labelled with the PDF set used in the nominal dataset, CT10. The other lines are predictions from the central values of other PDF sets. The ratio is defined with the nominal $t\bar{t}H$ prediction in the denominator.

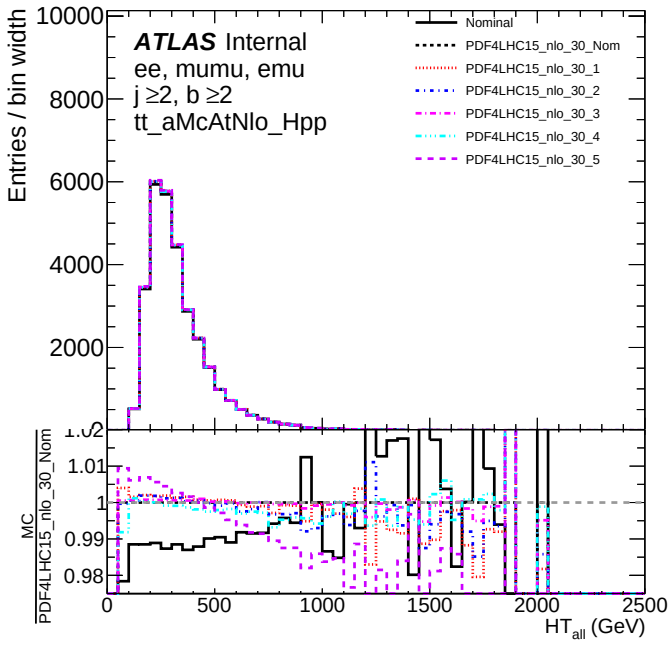
objects, show larger shape variations which reach 2% and rarely up to 4% in some histogram bins; distributions related to momentum and mass again show little shape difference. The inter-PDF variations for $t\bar{t}H$ are smaller than for $t\bar{t}$, which is expected due to the use of one of the PDF4LHC15 input PDF sets in the nominal $t\bar{t}H$ dataset. There is a small shape difference and a small normalisation increase but the effects are generally less than 1%. As well, it is not clear that a comparison of nominal $t\bar{t}H$ results with the results from using `PDF4LHC15_nlo_30` makes sense given that NNPDF3.0 PDF set is used in the process of defining results for both. The intra-PDF variations for both $t\bar{t}$ and $t\bar{t}H$ are much less than 1% in most of the 30 variations, with a few variations showing changes of up to 1% to 2% in the tails of distributions. The results for inter-PDF and intra-PDF variations were similar across the dilepton and single-lepton analysis channels, meaning that the same or similar approaches could be used for both when defining the systematic uncertainties used in the statistical fit.

Based on the observed PDF variations choices were made on how to define uncertainties in the statistical fit:

- firstly, it was decided that any equivalent uncertainties would be separated between

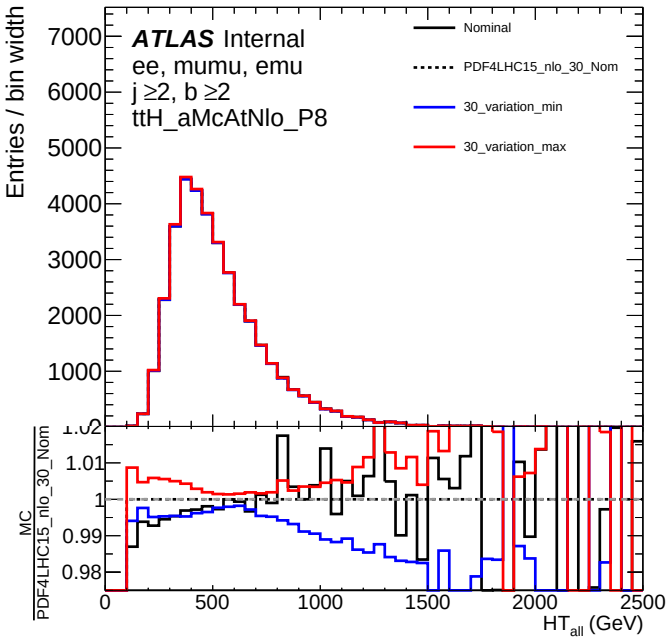


a)

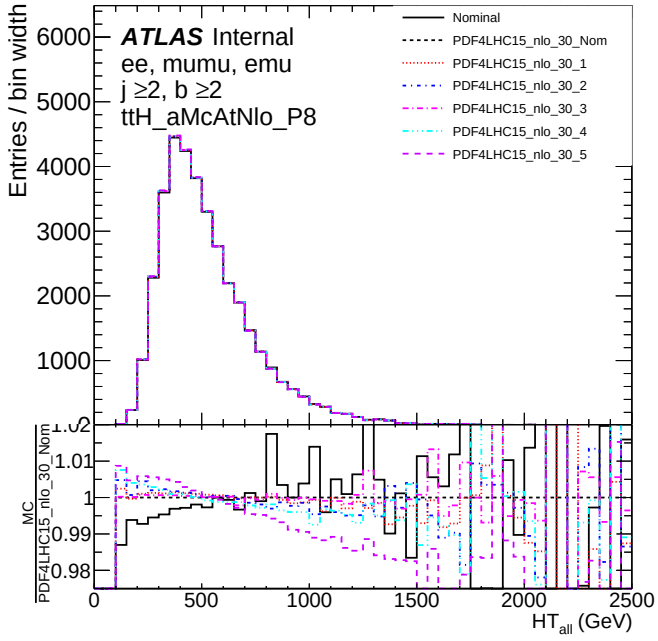


b)

FIGURE 6.12: Comparison of nominal distributions and distributions reweighted to intra-PDF variations for $t\bar{t}$. The distribution shown is the scalar sum of jet and lepton p_{T}, labelled $\mathrm{HT_{all}}$. The solid black line shows the nominal distribution for the dataset while the dashed black line shows the prediction from the central value of the `PDF4LHC15_nlo_30` set. Shown in *a*) are the maximum and minimum values per-bin of the 30 intra-PDF variations. Shown in *b*) are the first five intra-PDF variations. In both figures the ratio is defined with the nominal value for `PDF4LHC15_nlo_30` in the denominator.



a)



b)

FIGURE 6.13: Comparison of nominal distributions and distributions reweighted to intra-PDF variations for $t\bar{t}H$ in the dilepton channel. The distribution shown is the scalar sum of jet and lepton p_{T}, labelled $\mathrm{HT_{all}}$. The solid black line shows the nominal distribution for the dataset while the dashed black line shows the prediction from the central value of the `PDF4LHC15_nlo_30` set. Shown in *a*) are the maximum and minimum values per-bin of the 30 intra-PDF variations. Shown in *b*) are the first five intra-PDF variations. In both figures the ratio is defined with the nominal value for `PDF4LHC15_nlo_30` in the denominator.

$t\bar{t}$ and $t\bar{t}H$ because of the different simulation configurations used to model these processes (e.g. matrix element generator used), allowing PDF uncertainties to be separately correlated with other modelling uncertainties for the respective process;

- secondly, it was decided not to define an inter-PDF uncertainty for $t\bar{t}H$, which was due to the similarity of the inter-PDF variations and the related question of whether a comparison between NNPDF3.0 and PDF4LHC15 is meaningful;

- thirdly, it was decided to define a single inter-PDF uncertainty for $t\bar{t}$, which was defined by the difference between the nominal (CT10) prediction and the PDF4LHC15 prediction;

- finally, it was decided to define all 30 intra-PDF uncertainties individually, which allowed for the complete correlation of uncertainties; any negligible uncertainties could then be removed by the fitting software using an appropriate minimum-variation threshold. This pruning process is described in Section 6.7.

This defined 31 uncertainties for $t\bar{t}$ and 30 uncertainties for $t\bar{t}H$, which were all defined by applying the same PDF reweighting process to the distributions being fitted. This resulted in one-sided uncertainties that could then be symmetrised to form two-sided uncertainties. The symmetrisation process is described further in Section 6.7. Fits were then run by other analysers in the $t\bar{t}H(b\bar{b})$ analysis group with and without the PDF uncertainties included in individual fits to the dilepton and single-lepton channels. From these fits it was seen that only four of the 30 $t\bar{t}$ intra-PDF uncertainties were used, while 15 of the 30 $t\bar{t}H$ intra-PDF uncertainties were used. These surviving intra-PDF uncertainties and the inter-PDF uncertainty did not show any tightened constraints after the fit has been performed, and in most cases the fit kept the nominal dataset prediction without any correction. The few uncertainties which were used to pull the dataset predictions away from the nominal prediction did not have large pulls except for the fifth intra-PDF uncertainty for $t\bar{t}$. For this uncertainty a pull was observed equal to the magnitude of the upwards uncertainty. The effect of including PDF uncertainties on the measured value and uncertainty on the $t\bar{t}H$ signal strength was seen to be negligible and the effects on the free-floating $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ normalisation factors and other systematic uncertainties were small or negligible.

Due to the negligible effects of the PDF uncertainties on the parameter of interest and the other systematic uncertainties, the small magnitude of the PDF uncertainties, and the lack of constraints or significant corrections for most of the PDF uncertainties which were used by the statistical fit it was decided not to include PDF uncertainties in the 13.2 fb^{-1} $t\bar{t}H(b\bar{b})$ analysis. In principle they could have been included; however, the technical challenge of harmonising 61 additional uncertainties across the dilepton and single-lepton channel fits was large enough to warrant their exclusion given the small effects on the results of the fit from their inclusion. PDF uncertainties were again not included in the 36.1 fb^{-1} analysis (the studies were not repeated and PDF uncertainties were assumed to remain small, although in principle the studies should have been repeated too check).

6.7 Statistical Analysis Using a Binned Profile Likelihood Fit

To obtain a measurement of the presence of $t\bar{t}H$ signal in the observed data the ensemble of analysis regions, simulated signal and background processes, systematic uncertainties, and the chosen parameter of interest (representing $t\bar{t}H$ signal strength) are encoded in a likelihood function whose parameters can then be fit to give the best prediction of the observed data. This likelihood function is the combination of likelihoods for the observed data given the predicted signal and background yields, where the signal strength parameter is included as a scaling factor on the predicted signal yield, and likelihoods representing additional nuisance parameters which can affect the predicted yields, each of which have an uncertainty from an associated systematic uncertainty. Normalisation factors for some background processes are also included in a similar way to the signal strength parameter. While these normalisation factors are conceptually classified as nuisance parameters they are handled differently in the fit model, as discussed below. The use of a likelihood function which encodes systematic uncertainties in this way is referred to as a 'profile likelihood fit'. Using the method of maximum likelihood [146] parameter values can be obtained which give the highest probability to reproduce the observed data.

The method of maximum likelihood works by maximising a likelihood function, $L(\vec{\theta})$, as a function of its parameters to obtain estimated parameter values, $\hat{\theta}_i$, which provide

the most probable prediction of data. This is done by determining values for θ_i at maxima of $L(\vec{\theta})$ i.e. solving the series of equations:

$$\frac{\partial L}{\partial \theta_i} = 0, \quad \theta_i \in \vec{\theta} \tag{6.3}$$

Technically this maximisation is performed for $\ln L$ due to the convenience this brings of transforming multiplicative terms in the likelihood function into a series where each term can be differentiated individually. If the likelihood function is simple then $\ln L$ is approximately parabolic with respect to $\vec{\theta}$ and therefore has a single global maximum. However in practice likelihood functions such as the one used in the $t\bar{t}H(b\bar{b})$ analysis are high-dimensional problems with significant correlations between parameters and can therefore feature many local maxima. The best fit parameter estimators correspond to the true global maximum i.e. the largest value of $\ln L$. Maximisation of $\ln L$ is performed numerically using the MIGRAD routine of the MINUIT function fitting software package [147, 148]. Estimators of the variances (and covariances) of parameters can also be calculated from the likelihood at the position of best fit. This is done in two different ways in the MINUIT package for the $t\bar{t}H(b\bar{b})$ analysis [149]. The first type of variances, calculated using the HESSE routine, are defined using the knowledge that elements of the inverse covariance matrix can be calculated as the expectation values of the negative of the second derivatives of $\ln L$ with respect to a single parameter or pair of parameters[7]. To estimate the expectation values the inverse covariance matrix elements are instead calculated at the position of the best parameter estimates:

$$(\widehat{V^{-1}})_{ij} = - \left. \frac{\partial^2 L}{\partial \theta_i \partial \theta_j} \right|_{\vec{\theta}=\hat{\vec{\theta}}} \tag{6.4}$$

This matrix is numerically calculated by the HESSE routine by measuring finite differences at the position of $\hat{\vec{\theta}}$. The matrix is then inverted to arrive at the estimated covariance matrix. Implicit in this definition for variances is the assumption that $\ln L$ is parabolic about $\hat{\vec{\theta}}$ with respect to parameters θ_i and θ_j. In this way all correlations are included in the covariance matrix but any non-linearities of $\ln L$ are neglected. To calculate variances with any non-linear shape of $\ln L$ included, the MINOS routine is used. In this method

[7] This definition is taken from the Rao-Cramér-Frechet information inequality under the assumption that the estimator of a parameter is efficient (i.e. has minimal variance) and unbiased [146].

a chosen parameter is varied away from its estimator value and $\ln L$ is re-maximised at each variation with respect to all other parameters. The re-maximised likelihood is an example of a profile likelihood, defined as:

$$L_{\mathrm{p}}(\theta_i) = L(\theta_i, \hat{\hat{\theta}}(\theta_i)) \tag{6.5}$$

where θ_i is the chosen parameter and $\hat{\hat{\theta}}$ are the best estimated values of the remaining parameters at θ_i. Note that L_{p} only depends on the chosen parameter. Once the value of $\ln L_{\mathrm{p}}$ has decreased by a specified amount the difference between the varied value of the chosen parameter and its best estimated value is defined as the size of the variance of the parameter. In this way the shape of the maxima is directly followed to determine the parameter variance. This process is performed separately for positive and negative variations of the chosen parameter which can result in asymmetric final uncertainties on the parameter. Covariances with the other parameters are calculated as part of this process. MINOS uncertainties are calculated for the $t\bar{t}H$ signal strength parameter to ensure accurate calculation of the uncertainty on the measured signal strength. MINOS uncertainties are also used for the free-floating normalisation factors used for the $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$ processes.

Through the maximisation of the likelihood function nuisance parameters can be 'pulled' away from their prior estimate (according to the size of their associated systematic uncertainties) and agreement with data is improved. Uncertainties on the nuisance parameters can also be reduced according to the sensitivity of the likelihood function to variations of the nuisance parameters. These properties allow a best-fit value to be extracted for the $t\bar{t}H$ signal strength with reduced uncertainties compared to the prior estimates of systematic uncertainties. Within the $t\bar{t}H(b\bar{b})$ analysis the yields are arranged into binned distributions which either represent overall modelling of events (used in control regions) or attempt to separate background-like and signal-like events into different bins (used in signal regions). For control regions the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution is used. For signal regions the output of the classification BDTs trained for the relevant region is used. The bins are combined to form a binned likelihood function. The form of the binned likelihood function is:

$$L(\mu, \vec{k}, \vec{\theta}) = \prod_{\text{r}\in\text{regions}} \left[\prod_{\text{i}\in\text{bins}} \mathcal{P}(n_{r,i}|\nu_{r,i}(\mu, \vec{k}, \vec{\theta})) \right] \times \prod_{\text{j}\in\text{syst.}} f(\theta_{0\,j}|\theta_j) \tag{6.6}$$

where μ is the parameter of interest, $\vec{k}$ are normalisation factors, $\vec{\theta}$ are nuisance parameters, $n_{r,i}$ are the measured data yields, $\nu_{r,i}$ are predicted yields, $\theta_{0\,j}$ are prior estimates of nuisance parameters, $\mathcal{P}$ is a Poisson likelihood, and f are likelihoods used as constraint terms for nuisance parameters. Technically the double product over regions and bins is a single product over the ensemble of bins from all regions however the double product is shown to make explicit how different regions are included in the final likelihood function i.e. no distinction is made between control regions and signal regions.

$\nu_{r,i}$ are defined as:

$$\nu_{r,i}(\mu, \vec{k}, \vec{\theta}) = \mu S_{r,i}(\vec{\theta}) + B_{r,i}(\vec{k}, \vec{\theta}) = \mu S_{r,i}(\vec{\theta}) + \sum_{b\in\text{bkg.}} k_b B_{r,i,b}(\vec{\theta}) \tag{6.7}$$

where $S_{r,i}$ and $B_{r,i}$ are predicted signal and background yields respectively and the sum is over the ensemble of different background processes. Free-floating normalisation factors k_b are included for the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ processes as described in Section 6.6.2. For other background processes the respective k_b are fixed as $k_b = 1$ and yields are based wholly upon values of $\vec{\theta}$. The k_b variables for $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ are labelled as $k(t\bar{t}+ \geq 1b)$ and $k(t\bar{t}+ \geq 1c)$ respectively.

μ is defined as:

$$\mu = \frac{\sigma}{\sigma_{\text{SM}}} \tag{6.8}$$

where σ is the measured signal cross-section and σ_{SM} is the corresponding cross-section from Standard Model calculations. Prior values for μ as well as for $k(t\bar{t}+ \geq 1b)$ and $k(t\bar{t}+ \geq 1c)$ are set to 1.

The f functions in Equation 6.6 are included as a representation of auxiliary measurements of the nuisance parameters and their uncertainties [150, 151]. For nuisance parameters associated with the systematic uncertainties described in Section 6.6 these terms are defined using a Gaussian likelihood with the choice of $\theta_{0\,j} = 0$ as the Gaussian mean and a Gaussian width of $\sigma_{\theta_j} = 1$. The values $\theta_{0\,j} + \sigma_{\theta_j}$ and $\theta_{0\,j} - \sigma_{\theta_j}$ represent the estimated 'upward' and 'downward' variations for systematic uncertainty j as described

in Section 6.6.1. The relationship between a parameter x_j and its nuisance parameter is given by:

$$x_j = x_{0\,j} + \theta_j \sigma_{x_j} \tag{6.9}$$

where $x_{0\,j}$ is the prior estimate of x_j and σ_{x_j} is the prior uncertainty on x_j. Therefore if a nuisance parameter θ_j is pulled by the statistical fit when deriving the best fit model (with values of $\theta_j = \pm 1$ meaning a change in predicted yields of the size of the upward or downward systematic uncertainties) there will be an associated reduction to the likelihood function from the nuisance parameter's Gaussian likelihood. In this way the nuisance parameter likelihoods act as constraint terms which penalise the shift of nuisance parameters away from *a priori* knowledge, with the magnitude of the penalty corresponding to the size of the shift relative to the size of the estimated systematic uncertainty. Other f functions used are Poisson likelihoods which are used for nuisance parameters associated with statistical uncertainties on the predicted yields in each bin, referred to as 'gamma' parameters. These statistical nuisance parameters are defined so that they do not affect the predicted $t\bar{t}H$ yield, to ensure that the parameter of interest is not affected if these nuisance parameters are pulled.

The reduction in size of posterior nuisance parameter uncertainties compared to prior uncertainties arises due to the correlation of a given nuisance parameter with terms in the likelihood function other than the associated f likelihood. If a nuisance parameter has no correlation with other terms then variations in the total likelihood resulting from variations of the nuisance parameter will only occur for the relevant f term. The posterior nuisance parameter uncertainty will arise only from the $\pm\sigma_{\theta_j}$ variations of the parameter's f likelihood, and therefore the posterior uncertainty will be equal to the prior uncertainty and no constraint will be seen. However if variations of the nuisance parameter result in significant changes to the predicted yields, then the Poisson likelihoods for the predicted and measured yield in each bin will also significantly vary. The posterior nuisance parameter uncertainty will therefore arise from the combined variations in the calculated likelihood from the parameter's f likelihood and from the Poisson likelihoods. The rate of change for the total likelihood will be greater for this combination than for only variations of the f likelihood, and therefore the nuisance parameter's posterior uncertainty

will be reduced compared to the prior uncertainty. The comparison of predicted yields and measured yields has therefore constrained the uncertainty of the nuisance parameter. Finally, if the nuisance parameter has significant correlations with other nuisance parameters then variations of the nuisance parameter in question will result in variations to multiple f likelihood terms. This combination of changes will show a greater rate of change for the total likelihood than variations of a single f likelihood term, and therefore the nuisance parameter's posterior uncertainty will again be reduced compared to the prior uncertainty. The correlation of nuisance parameters (which is not included in prior uncertainties) has therefore constrained the uncertainty of the nuisance parameter.

To ensure that all systematic uncertainties are well defined and are not susceptible to statistical fluctuations, smoothing and symmetrisation is applied to the prior uncertainty estimates. Smoothing merges nearby bins with large relative inter-bin differences in the systematic uncertainty estimates (the bins are still separate in the likelihood calculation but the effect of the systematic uncertainty is defined based on merged bins); following this any remaining slopes in the inter-bin differences are smoothed to have a consistent shape across nearby bins. An example of smoothing for a systematic uncertainty can be seen in Figure 6.14. It can be seen that the large inter-bin differences in the initial uncertainty estimate have been transformed to a smooth slope, and the uncertainty estimates for the final two bins have been merged.

Symmetrisation is firstly defined for systematic uncertainties which have a single-sided prior estimate, for example comparisons of the nominal $t\bar{t}$ dataset with alternative datasets. The estimated uncertainty is taken as the upward uncertainty and the downward uncertainty is defined with the same relative difference to the nominal prediction but with the opposite sign. Symmetrisation is also defined for systematic uncertainties with separate upward and downward estimates. For these systematics the magnitude of the upward and downward variation is averaged and the average value is symmetrised about the nominal prediction. This avoids cases where the effects of upward and downward uncertainty estimates are in the same direction relative to the nominal prediction.

Nuisance parameters are also pruned prior to running the statistical fit to reduce the complexity of the model and to improve stability of the fit. Pruning compares the relative size of systematic uncertainties against specified thresholds and any nuisance parameters with uncertainties smaller than the thresholds will be pruned. The effects of a systematic

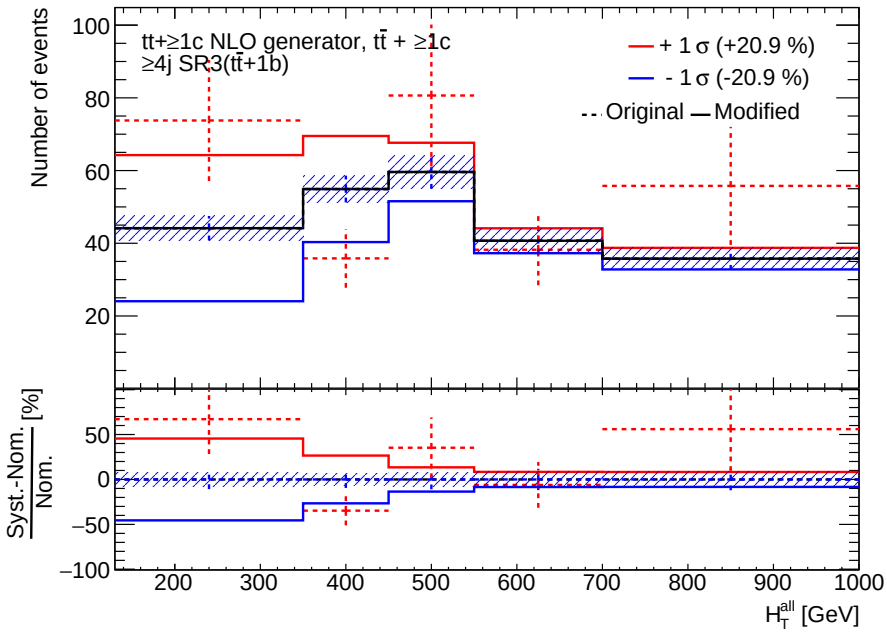


FIGURE 6.14: Example of smoothing applied to a systematic uncertainty. The uncertainty estimate prior to smoothing is shown as dashed red lines, and the uncertainty estimate after smoothing is shown as a solid red line. The black line shows the nominal distribution. The uncertainty shown is the comparison of the nominal $t\bar{t}+\geq 1c$ prediction with the prediction from the SHERPA5F dataset. The results are produced for the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in the $\geq 4j$ SR3 region. The effect of symmetrisation can also be seen from the presence of the solid blue line, which shows the downward uncertainty estimate.

uncertainty on normalisation and distribution shapes are considered separately which leads to three possibilities for the pruning of a nuisance parameter: the effect on normalisation is kept but the effect on distribution shapes is removed; the effect on distribution shapes is kept but the effect on normalisation is removed; or both effects are removed and the nuisance parameter is completely removed from the fit. The effects of pruning are considered individually in each analysis region meaning that the effects of a nuisance parameter can be completely pruned and removed from the fit in one region while still keeping the nuisance parameter in the fit model if it has a sufficient effect in another region. Minimum thresholds on the size of normalisation or shape effects were set as 1% for the $t\bar{t}H(b\bar{b})$ analysis.

Results from the author's work on the development and study of the $t\bar{t}H(b\bar{b})$ dilepton channel fit model are presented in this section. This work formed a continuous process during the development of the $t\bar{t}H(b\bar{b})$ analysis as a whole which included numerous updates, additions, and changes. A selection of significant developments and studies are presented here. Due to the continuous development of the fit model, changes are present

between sub-sections. Significant changes will be noted however smaller changes will not be mentioned for the sake of avoiding lists of minor details. Nuisance parameters may also be present which are not used in the final fit model and whose systematic uncertainties are not discussed in Section 6.6.

6.7.1 Development of Initial Fit Model

A first definition of the $t\bar{t}H(b\bar{b})$ statistical fit model was done using a simplified approach in which only processes and systematic uncertainties which were expected to be dominant were included. This included most processes except for the $tHjb$ and WtH processes and small backgrounds involving top quarks such as $t\bar{t}t\bar{t}$. Most systematic uncertainties were included except for electron and muon related uncertainties, and Wt related uncertainties. Cross-section uncertainties were defined with a single uncertainty instead of uncertainties split between the effects of QCD scale and PDF+α_S variations. At this stage of the development of the $t\bar{t}H(b\bar{b})$ analysis the systematic uncertainties defined through the comparison of nominal $t\bar{t}$ predictions to the SHERPA5F and SHERPA4F datasets were not used and instead uncertainties referred to as 'NLO generator' uncertainties were defined by comparing nominal $t\bar{t}$ predictions to predictions from datasets simulated using MG5_aMC+PYTHIA 8. The $t\bar{t}+ \geq 1b$ uncertainties defined using variations of the SHERPA4F dataset were also not included at this stage. The classification BDTs used as binned distributions in signal regions had not been trained at this stage in the development of the 36.1 fb^{-1} $t\bar{t}H(b\bar{b})$ analysis and so classification BDTs from the 13.2 fb^{-1} analysis were used instead as appropriate but unoptimised discriminants. This model is referred to as the 'simple' model.

Asimov fit for the simple fit model

To perform a first test of the fit model and observe the constraining power of the fit model on the uncertainties of the $t\bar{t}H$ signal strength and of the nuisance parameters an 'Asimov fit' was performed using the simple model [152]. In an Asimov fit the real data is instead replaced with an 'Asimov dataset'; this dataset is defined by the property that the estimated parameters, $\hat{\hat{\theta}}$ (where this refers to all parameters, including the signal strength parameter, μ), are equal to the true parameter values. To define this fit in

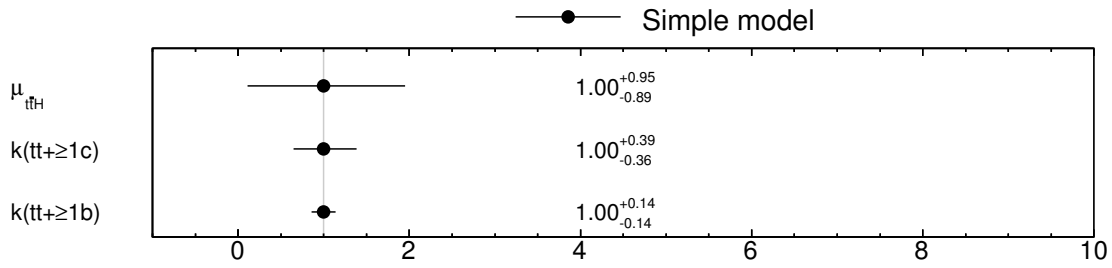


FIGURE 6.15: Fitted $t\bar{t}H$ signal strength and normalisation factors for the simple fit model for a fit to the Asimov dataset.

practice the predicted yields from simulated data are used instead of real data, in effect fitting the predicted yields to themselves. In this way the expected constraining power of the likelihood function can be determined including an estimate of the significance for a measurement of the presence of $t\bar{t}H$ signal at the Standard Model prediction. As well, the fit model is tested for biases by observing if any nuisance parameters are pulled away from their prior values (which by definition are the true values in the Asimov dataset). The results of the Asimov fit for the simple fit model are shown for the $t\bar{t}H$ signal strength and $k(t\bar{t}+\geq 1b)$ and $k(t\bar{t}+\geq 1c)$ normalisation factors in Figure 6.15 and for the nuisance parameters in Figure 6.16. The list of all unpruned nuisance parameters are shown on a scale defined by the values of $\theta_{0j} = 0$ and $\theta_{0j} \pm \sigma_{\theta_j} = \pm 1$ for nuisance parameter θ_j. As such, the posterior uncertainties on the nuisance parameters are equivalent to the prior systematic uncertainty estimates if the uncertainty bars equal the width of the green region.

It is firstly observed that no pulls in the parameters away from the prior values are seen, showing that the likelihood function is free of bias. Secondly, it is seen from the uncertainties on $k(t\bar{t}+\geq 1b)$ and $k(t\bar{t}+\geq 1c)$ that the likelihood function is rather sensitive to $k(t\bar{t}+\geq 1b)$ and therefore able to strongly constrain the normalisation for $t\bar{t}+\geq 1b$ while it is less sensitive to $k(t\bar{t}+\geq 1c)$ and therefore less able to constrain the normalisation of $t\bar{t}+\geq 1c$. Thirdly, it is seen that in this initial fit model the likelihood function is not highly sensitive to μ which results in uncertainties on μ of close to the size of the Standard Model $t\bar{t}H$ signal and would therefore result in a very low discovery significance from the dilepton channel alone if this signal is present in real data. Finally, it is seen that most experimental nuisance parameters are not significantly constrained except for leading b-tagging uncertainties and some jet energy scale uncertainties. On the other

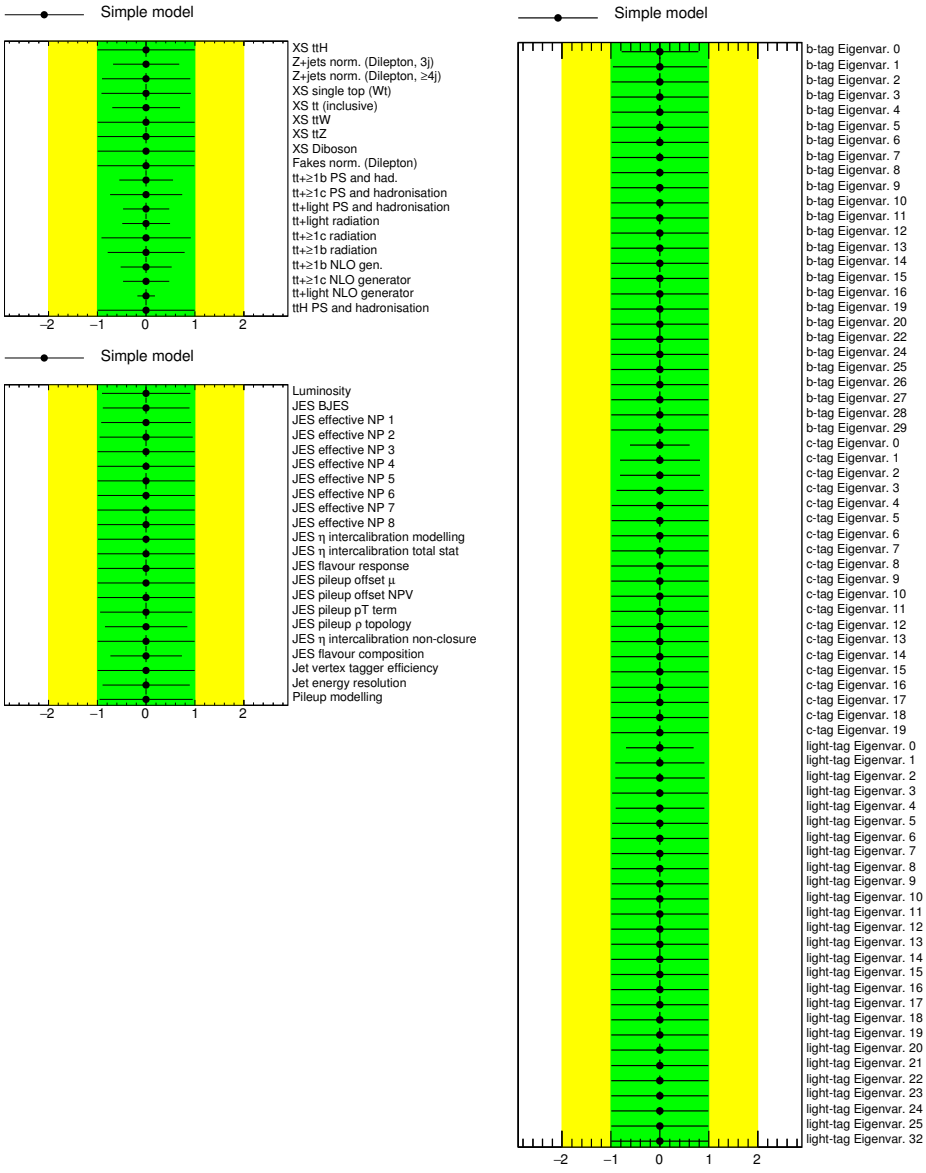


FIGURE 6.16: Fitted nuisance parameters for the simple fit model for a fit to the Asimov dataset. Points show the fitted values of the nuisance parameters and uncertainty bars show the posterior uncertainty of the nuisance parameters. The x-axis represents the pull of a nuisance parameter in units of the prior uncertainty. The green (yellow) region represents $\pm 1\sigma$ ($\pm 2\sigma$) uncertainties. In the top left plot theoretical nuisance parameters are shown, in the bottom left plot non-b-tagging instrumental nuisance parameters are shown, and in the right plot b-tagging nuisance parameters are shown.

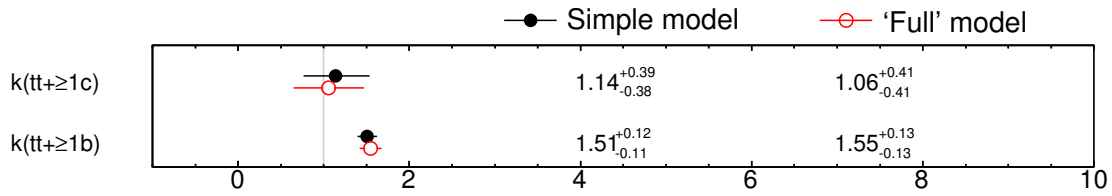


FIGURE 6.17: Fitted normalisation factors compared for the simple fit model and first full fit model for fits to data. The fits are performed under the background-only hypothesis and use the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in all regions.

hand, $t\bar{t}$ modelling uncertainties are significantly constrained. This reflects the limitations in $t\bar{t}$ modelling described in Section 6.6.2 that result in large prior uncertainty estimates for $t\bar{t}$ modelling. This also reflects the strong sensitivity of the likelihood function to $t\bar{t}$ processes due to the high proportion of $t\bar{t}$+jets in each analysis region relative to other background processes.

Background-only fit to data for initial fit models

The next test of the fit model was a fit to real data using a background-only fit model. A background-only fit model is defined by removing the $\mu S_{r,i}$ term in Equation 6.7; therefore only background processes are considered and μ is not fitted. $H_{\mathrm{T}}^{\mathrm{all}}$ distributions were used as binned distributions in signal regions instead of BDT outputs to test basic kinematic modelling in these regions. An expanded version of the fit model referred to as the first full fit model was also tested which included the processes and systematic uncertainties that were excluded from the simple model as described above. The results of the two fits could be compared to test the sensitivity of the fit to the additional processes and nuisance parameters. An additional systematic uncertainty which was not used in the final fit model was also added as part of the first full fit model. This systematic is labelled as '$t\bar{t}+ \geq 1b$ Sherpa OL bb categories' and is defined so that the upward uncertainty removes the effects of correcting the relative fractions of $t\bar{t}+ \geq 1b$ sub-categories to be the same as in the SHERPA4F dataset. The results of the background-only fit to data for both the simple fit model and the first full model are shown for the $k(t\bar{t}+ \geq 1b)$ and $k(t\bar{t}+ \geq 1c)$ normalisation factors in Figure 6.17 and for the nuisance parameters in Figure 6.18.

It is firstly observed that the normalisation factor $t\bar{t}+ \geq 1b$ is fitted to a value of approximately 1.5. This reflects the previously mentioned disagreement in total yields for

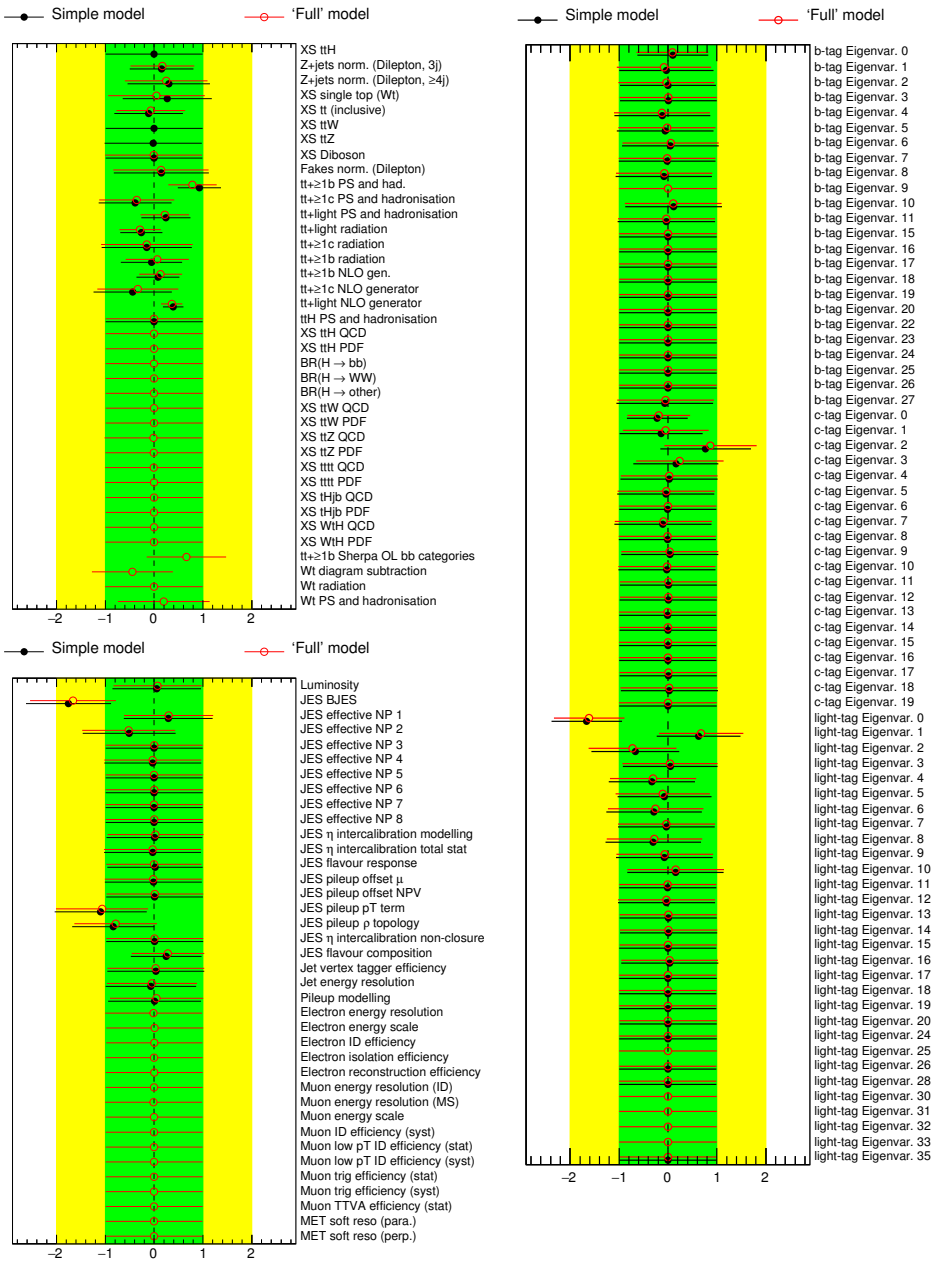


FIGURE 6.18: Fitted nuisance parameters compared for the simple fit model and for the first full fit model for fits to data. The fits are performed under the background-only hypothesis and use the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in all regions. Points show the fitted values of the nuisance parameters and uncertainty bars show the posterior uncertainty of the nuisance parameters. The x-axis represents the pull of a nuisance parameter in units of the prior uncertainty. The green (yellow) region represents $\pm 1\sigma$ ($\pm 2\sigma$) uncertainties. In the top left plot theoretical nuisance parameters are shown, in the bottom left plot non-b-tagging instrumental nuisance parameters are shown, and in the right plot b-tagging nuisance parameters are shown. In some cases nuisance parameters are only included in the first full fit model and only a red hollow point is shown.

$t\bar{t}+ \geq 1b$ between real data and simulated data in analysis regions which are pure in $t\bar{t}+ \geq 1b$. On the other hand the normalisation factor for $t\bar{t}+ \geq 1c$ is much closer to unity; however the uncertainty on $k(t\bar{t}+ \geq 1c)$ approximately covers both a value of one and a value close to that of $k(t\bar{t}+ \geq 1b)$. Secondly, a number of large nuisance parameter pulls, including some that are larger than the prior systematic uncertainty estimates, are present. Large instrumental pulls are seen for jet energy scale uncertainties and for leading light-jet and c-jet mis-tagging uncertainties. Pulls are seen for $t\bar{t}$ modelling parameters showing corrections to $t\bar{t}$ modelling and in particular the $t\bar{t}+ \geq 1b$ parton shower modelling uncertainty is pulled with a similar magnitude to the size of the prior systematic uncertainty estimate. Thirdly, fit results do not show large differences between the simple and first full fit model and most nuisance parameters added in the first full fit model are not pulled and are not constrained. This shows the sub-dominant effects of the additional processes and nuisance parameters as well as the resilience of the fit maximisation to small changes. The exceptions to this are the Wt theoretical uncertainties and the "Sherpa OL bb categories" nuisance parameter which corrects the relative fractions of the $t\bar{t}+ \geq 1b$ sub-categories to be approximately halfway between the SHERPA4F and nominal $t\bar{t}$ predications.

Examples of the effects of the first full model fit on the distributions used in each region are shown in Figure 6.19. The improvement in overall normalisation and reduction of shape mismodelling can be seen from the ratios of data to total predicted yield in each bin of the distributions. The reduction in the size of the uncertainty in each bin can also be seen.

Tests of SHERPA5F-based $t\bar{t}$ modelling uncertainties

Tests of replacing the NLO generator $t\bar{t}$ uncertainties with the uncertainties which are defined by comparing nominal predicted $t\bar{t}$ yields to the SHERPA5F dataset were also performed. These tests were performed with later versions of the simulated datasets. The main difference between the datasets used previously in this section and those used for the tests of including $t\bar{t}$ uncertainties defined using SHERPA5F was a change in the calibration from data of the c-jet mis-tagging efficiencies for the MV2c10 b-tagger. The

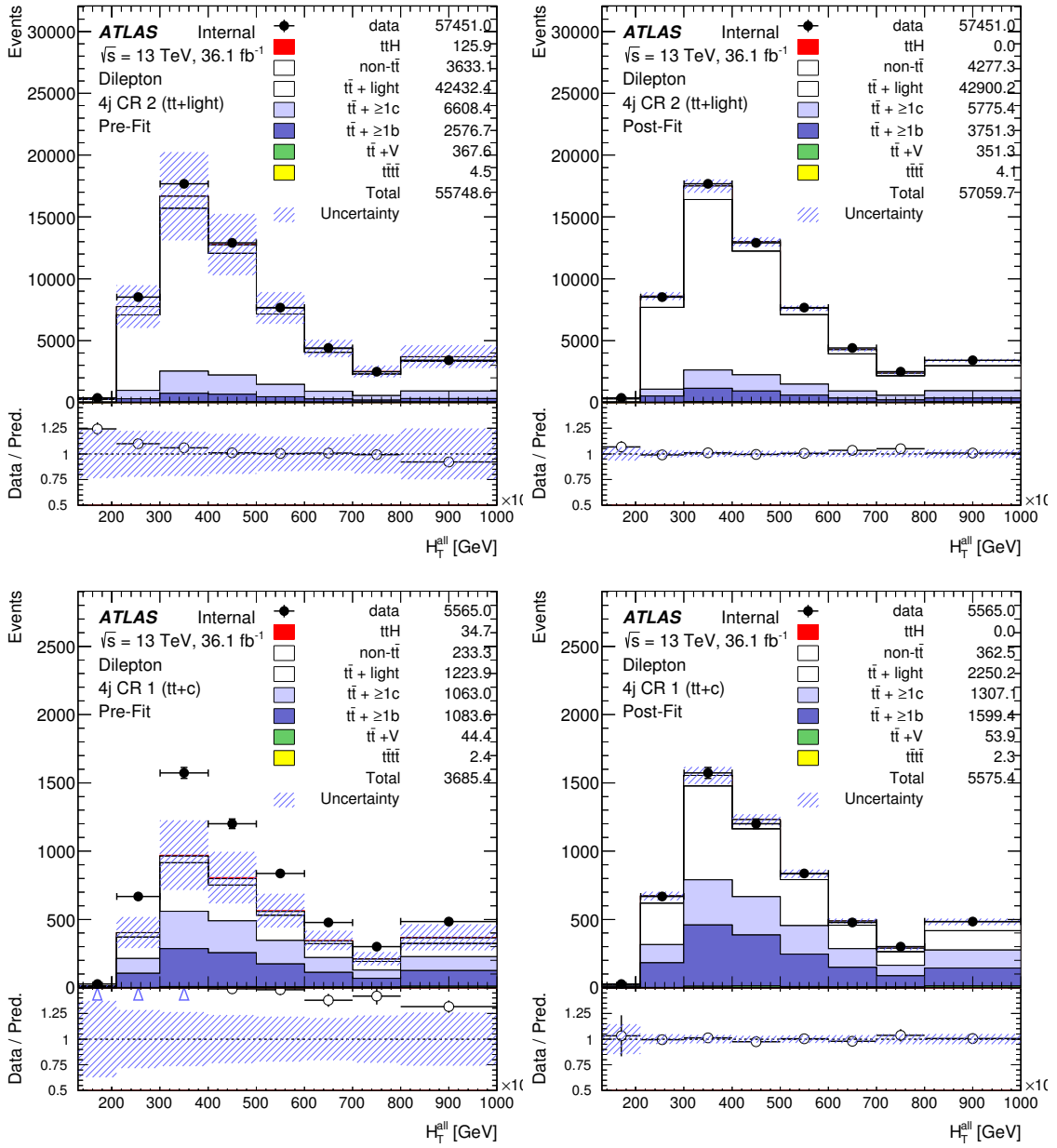


FIGURE 6.19: Fitted distributions for the first full fit model in which the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution is used in all analysis regions for a fit to data under the background-only hypothesis. In the left (right) column pre-fit (post-fit) distributions are shown. In the top (bottom) row results for the ≥ 4 jet $\mathrm{CR}_{t\bar{t}+\mathrm{light}}$ (≥ 4 jet $\mathrm{CR}_{t\bar{t}+\geq 1c}$) control region are shown. An error in plotting resulted in the drawing of both the non-$t\bar{t}$ and the $t\bar{t}$ +light categories with a white area. The non-$t\bar{t}$ category is plotted at the top of the stack of background histograms.

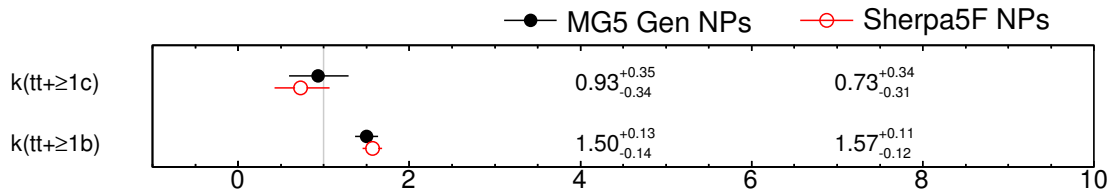


FIGURE 6.20: Fitted normalisation factors compared for MG5_aMC-based and SHERPA5F-based $t\bar{t}$ uncertainties for fits to data. The fits are performed under the background-only hypothesis and use the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in all regions.

dominant result of this difference is the increase in magnitude and change in sign of leading c-jet mis-tagging nuisance parameter pulls and the reduction of the fitted $k(t\bar{t}+\geq 1c)$ value to be slightly less than one. As well in this fit model the "Sherpa OL bb categories" nuisance parameter is almost unpulled and an upward pull is seen on jet energy resolution. A full comparison of results for this updated fit model with results from the first full fit model shown previously in this section can be seen in Appendix A.1.

Asimov fits were performed with the MG5_aMC+PYTHIA 8-based NLO generator uncertainties and with the SHERPA5F-based uncertainties. The uncertainty on μ was seen to increase from $\sigma_\mu = ^{+0.93}_{-0.89}$ to $\sigma_\mu = ^{+1.07}_{-1.04}$. The $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ SHERPA5F nuisance parameters were seen to be more constrained by the fit than the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ NLO generator uncertainties. The $t\bar{t}$ +light radiation nuisance parameter was also seen to be more constrained in the fit model with SHERPA5F $t\bar{t}$ uncertainties included. All other nuisance parameters showed little to no change in the size of their post-fit uncertainties. Data fits were then performed under the background-only hypothesis with $H_{\mathrm{T}}^{\mathrm{all}}$ used as the distribution in all regions. The comparison of background-only fits to data for both the MG5_aMC+PYTHIA 8-based NLO generator uncertainties and the SHERPA5F-based uncertainties are shown for the $k(t\bar{t}+\geq 1b)$ and $k(t\bar{t}+\geq 1c)$ normalisation factors in Figure 6.20 and for nuisance parameters which show significant changes between the two fit models in Figure 6.21. Smaller differences were also seen in other nuisance parameters such as other light-jet mis-tagging nuisance parameters.

It is seen that the value of $k(t\bar{t}+\geq 1c)$ decreases and most $t\bar{t}$ parton shower modelling and $t\bar{t}$ radiation modelling nuisance parameter pulls increase. The MG5-based nuisance parameters and SHERPA5F-based nuisance parameters show significantly different pulls. Significant changes are also seen for instrumental nuisance parameters. A number of

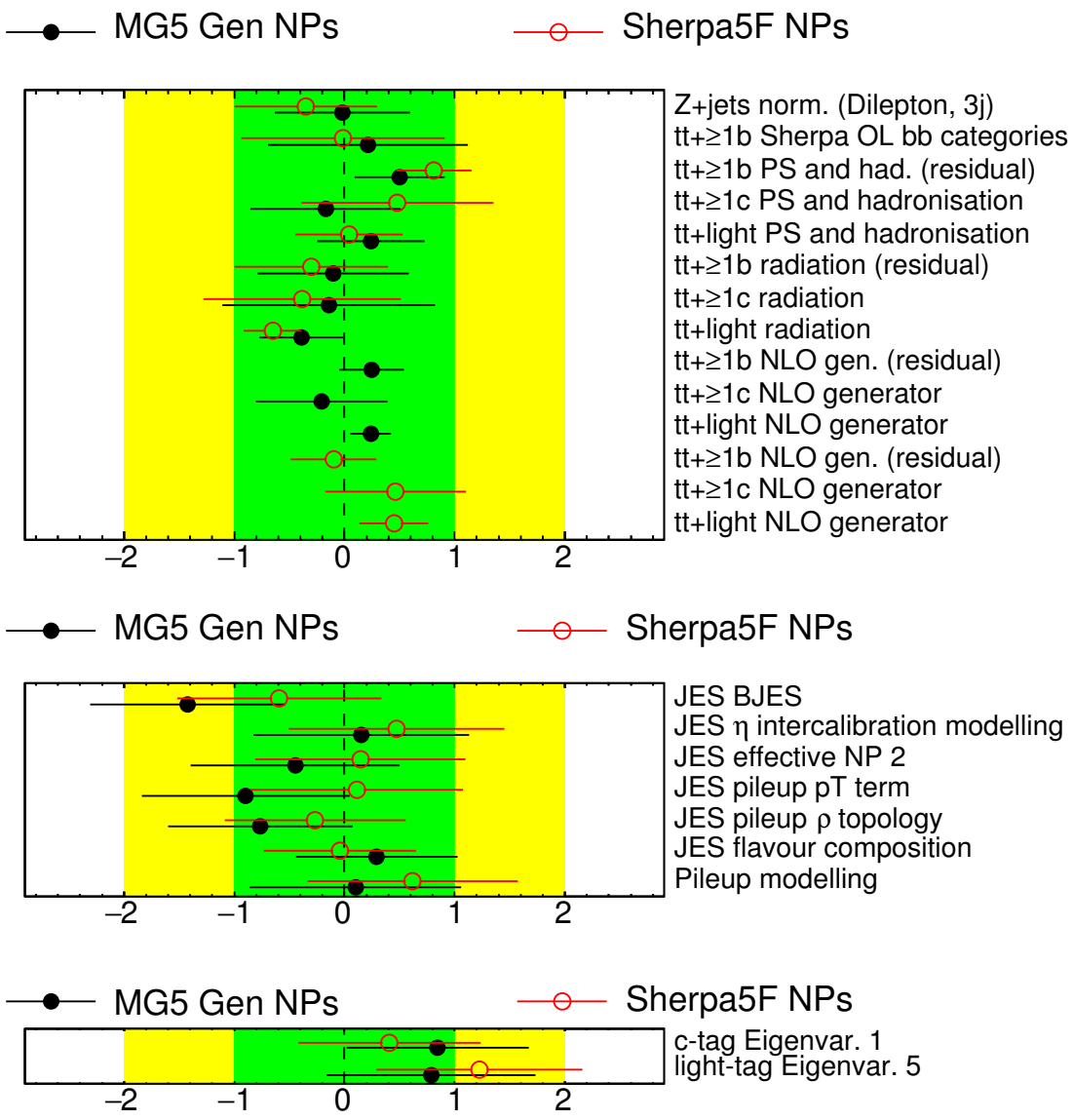


FIGURE 6.21: Fitted nuisance parameters compared for the use of MG5_aMC-based and SHERPA5F-based $t\bar{t}$ uncertainties for fits to data. The fits are performed under the background-only hypothesis and use the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in all regions. Points show the fitted values of the nuisance parameters and uncertainty bars show the posterior uncertainty of the nuisance parameters. The x-axis represents the pull of a nuisance parameter in units of the prior uncertainty. The green (yellow) region represents $\pm 1\sigma$ ($\pm 2\sigma$) uncertainties. In the top plot theoretical nuisance parameters are shown, in the middle plot non-b-tagging instrumental nuisance parameters are shown, and in the bottom plot b-tagging nuisance parameters are shown. Nuisance parameters labelled "NLO gen." or "NLO generator" that are plotted with red hollow points are actually the SHERPA5F-based nuisance parameters which model the effects of both matrix element generator and parton shower generator variations. Nuisance parameters with small changes or no changes between the two fit models are not shown.

jet energy scale nuisance parameter pulls are reduced while increases are seen for pile-up modelling, a light-jet mis-tagging nuisance parameter, and the Z+jets normalisation correction in the three jet analysis regions. These changes in pulls are an indication of the sensitivity of the likelihood function to the modelling of $t\bar{t}$ events and limitations in the fit model to correct for mismodelling of the $t\bar{t}$ process. Nuisance parameters that are not specific to $t\bar{t}$ may be pulled to correct residual mismodelling that is not fully corrected by pulls in $t\bar{t}$ modelling nuisance parameters. SHERPA5F $t\bar{t}$ uncertainties were chosen to be used in most of the remainder of the development of the $t\bar{t}H(b\bar{b})$ analysis. This choice was made due to the presence of event weights with negative signs in the MG5_aMC+PYTHIA 8 dataset. These negative weights had the effect of reducing the effective yields of this dataset which resulted in increased statistical uncertainties.

6.7.2 Choice of Binning in Control Regions

The choice of the positions and widths of distribution bins used in the binned likelihood function can result in significant differences to the resulting parameter pulls as well as affecting statistical uncertainties. The choice of binning in control regions should provide enough detail of the underlying distributions to correct mismodelling of background processes that are also present in signal regions. At the same time care must be taken to avoid being oversensitive to mismodelling in small regions of phase space. This could result in extreme nuisance parameter pulls due to the need to correct mismodelling using nuisance parameters that in most cases determine modelling over the full phase space of the $t\bar{t}H(b\bar{b})$ analysis. Changes in modelling outside of the region of phase space where mismodelling is seen will be compensated for by additional pulls in correlated nuisance parameters. Initial binning choices for the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in the control regions used in Section 6.7.1 were defined using bin widths of 100 GeV in most cases. The first (last) bin in three jet regions covered the inclusive range ≤ 200 GeV (≥ 500 GeV) and the first (last) bin in inclusive four jet regions covered the inclusive range ≤ 210 GeV (≥ 800 GeV). The only bins without a width of 100 GeV were the second bins in inclusive four jet regions, which had a low edge at 210 GeV and a high edge at 300 GeV.

Effects of merging the first two distribution bins

A first test of the effect of binning in control regions was performed by merging the two bins covering the smallest $H_{\mathrm{T}}^{\mathrm{all}}$ values in each control region. The first bins in each control region were seen to contain low yields and in some regions to show significant prior disagreement with data or agreement dissimilar to the magnitude seen for other bins in the same region. As such the first two bins in each were merged to test the effect of these low $H_{\mathrm{T}}^{\mathrm{all}}$ bins on nuisance parameter pulls. Comparisons of the initial control region binning and binning with the first two bins of each control region merged are shown in Figure 6.22. These tests were performed prior to the tests of SHERPA5F-based $t\bar{t}$ uncertainties described at the end of Section 6.7.1 and so the fit model used is the non-updated "first full fit model" (i.e. with MG5_aMC+PYTHIA 8 NLO generator $t\bar{t}$ uncertainties).

Fits to data were performed with the first two bins merged in each control region in turn and then with bins merged in all control regions. The comparison of background-only fits to data for each binning configuration is shown for the $k(t\bar{t}+\geq 1b)$ and $k(t\bar{t}+\geq 1c)$ normalisation factors in Figure 6.23 and for nuisance parameters which show significant changes between the different configurations in Figure 6.24. Smaller differences were also seen in other nuisance parameters such as non-leading b-tagging nuisance parameters.

It is seen that the large pulls seen for the "JES BJES" and JES pile-up nuisance parameters are significantly reduced when merging bins in $t\bar{t}$+light control regions and are effectively removed when merging bins in all control regions. Merging bins in $t\bar{t}$ +HF control regions results in significantly smaller changes to pulls; however, small reductions of "JES BJES" and JES pile-up nuisance parameters are seen as well as a reduction of the $t\bar{t}+\geq 1b$ radiation pull when merging bins in the 3 jet $\mathrm{CR}_{t\bar{t}+\geq 1b}$ region. However a number of other pulls are significantly increased as bins are merged in $t\bar{t}$+light control regions. In particular Z+jets normalisation in 3 jet regions is pulled up to twice the size of its prior uncertainty estimate when merging bins in all control regions. Increased magnitudes are also seen for the inclusive $t\bar{t}$ cross-section, Wt parton shower modelling and diagram subtraction, as well as leading mis-tag b-tagging nuisance parameters. $t\bar{t}$ modelling pulls are in general seen to reduce slightly however $t\bar{t}+\geq 1c$ parton shower modelling is seen to increase when bins are merged in all control regions.

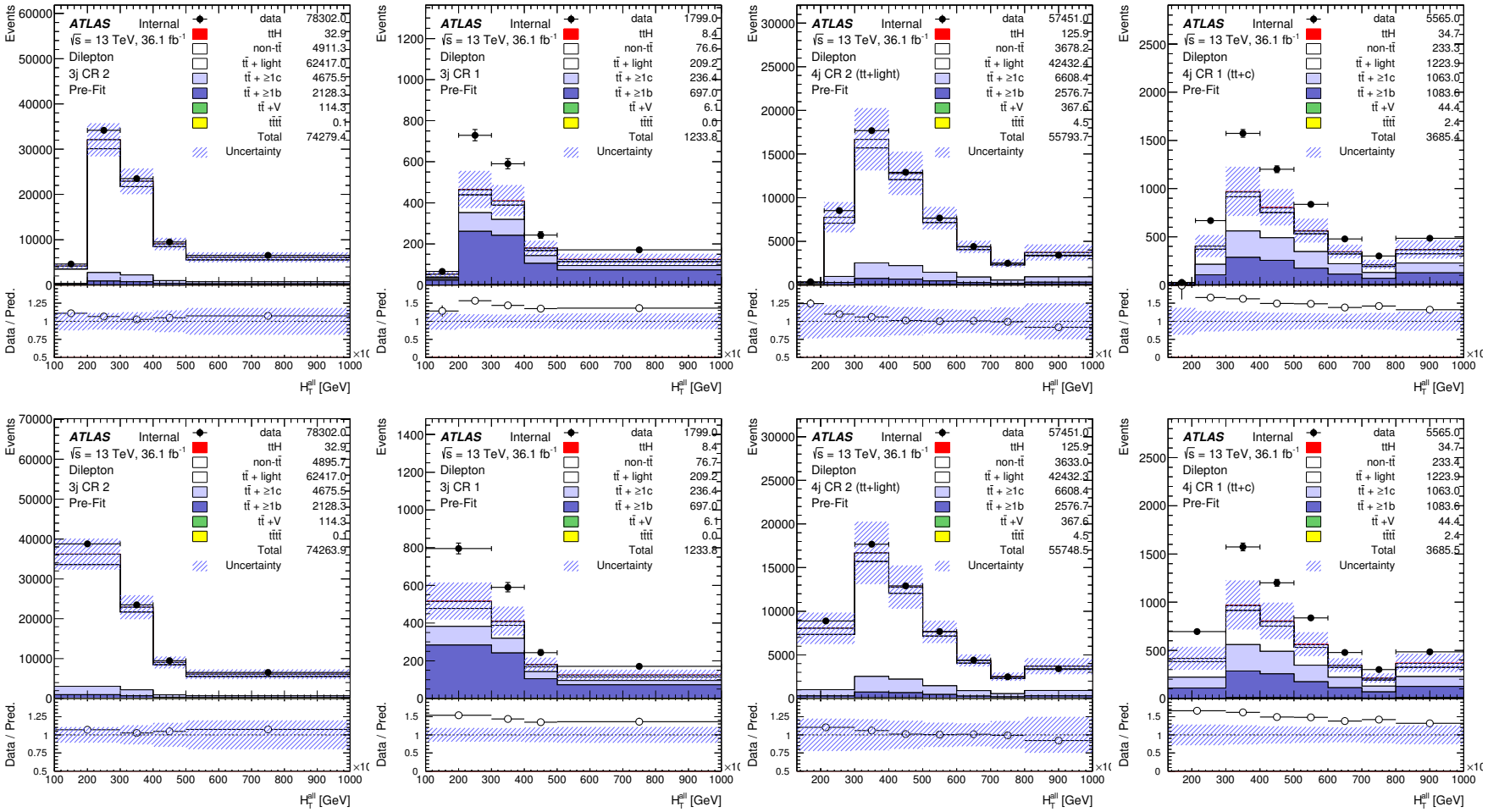


FIGURE 6.22: $H_{\mathrm{T}}^{\mathrm{all}}$ distributions in control regions. In the top (bottom) row distributions without (with) the first two bins merged are shown. Control regions are arranged left to right as: 3 jet $\mathrm{CR}_{t\bar{t}+\mathrm{light}}$, 3 jet $\mathrm{CR}_{t\bar{t}+\geq 1b}$, $\geq$ 4 jet $\mathrm{CR}_{t\bar{t}+\mathrm{light}}$, $\geq$ 4 jet $\mathrm{CR}_{t\bar{t}+\geq 1c}$. An error in plotting resulted in the drawing of both the non-$t\bar{t}$ and the $t\bar{t}$ +light categories with a white area. The non-$t\bar{t}$ category is plotted at the top of the stack of background histograms. A ratio scale of 0 to 2 is used for $t\bar{t}$+HF regions to allow all data points to be seen.

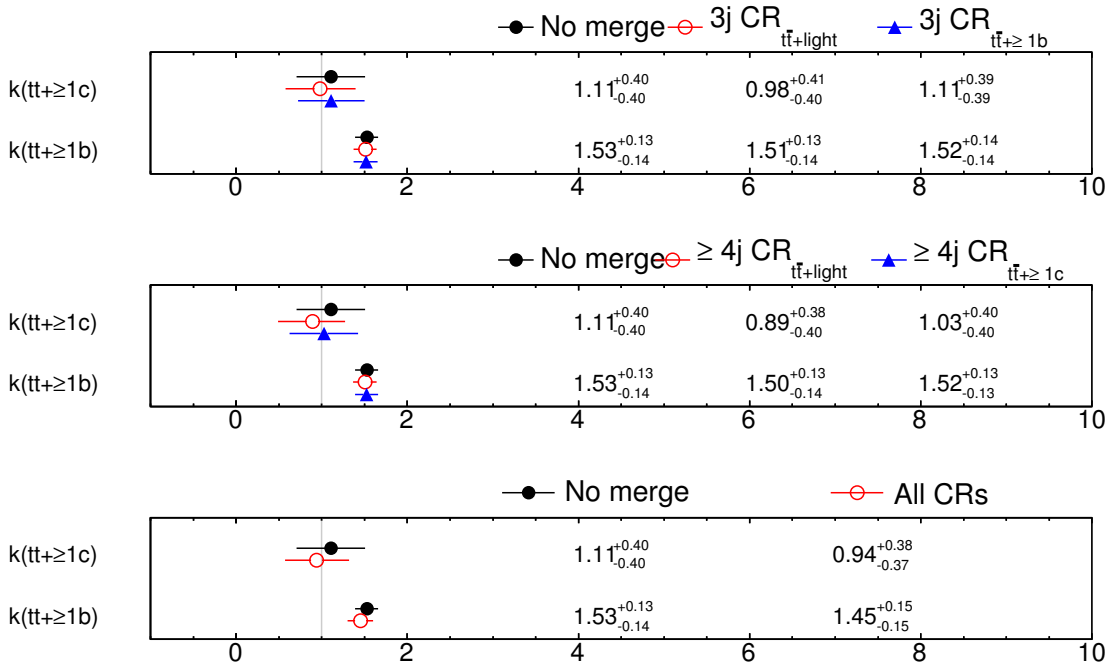


FIGURE 6.23: Fitted normalisation factors compared for the merging of the first two bins in different control regions. The fits are performed under the background-only hypothesis and use the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in all regions. Bins are merged in 3 jet regions in the top plot, in ≥ 4 jet regions in the middle plot, and in all regions in the bottom plot. In the top and middle plots the red circles and blue points represent fits where bins are merged in a single analysis region which is labelled at the top of the plot.

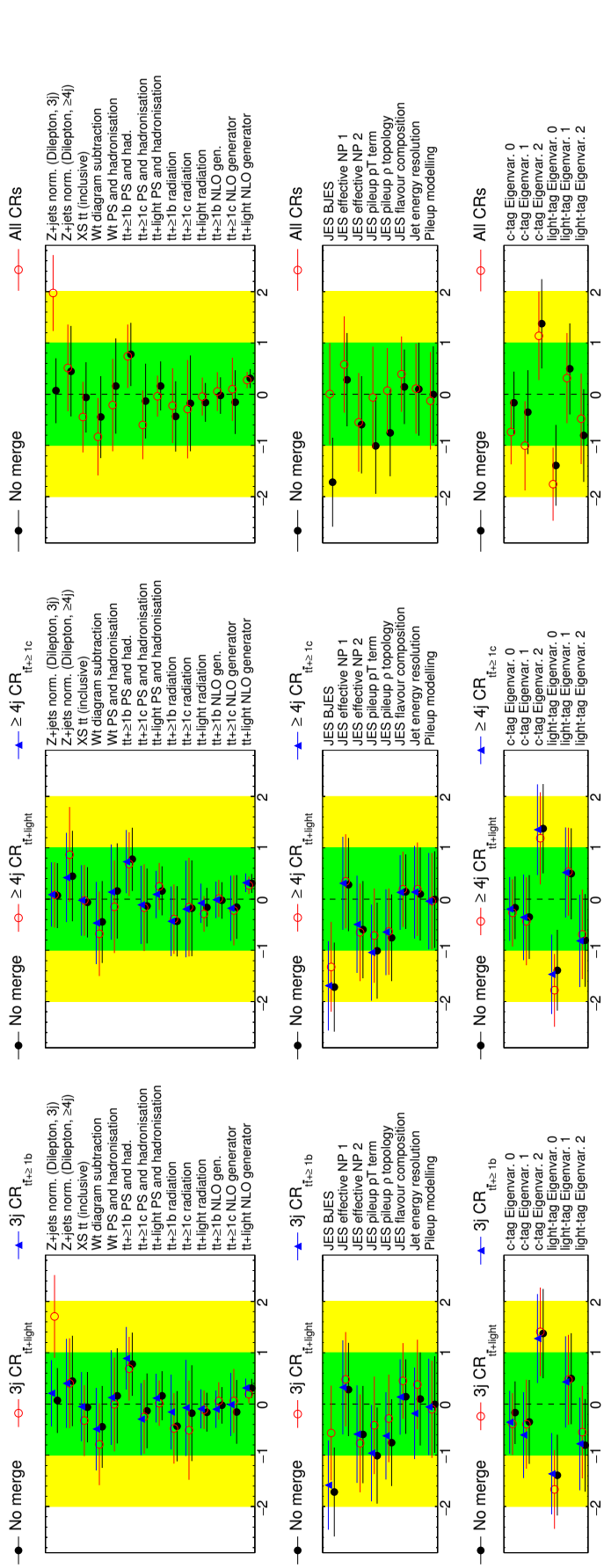


FIGURE 6.24: Fitted nuisance parameters compared for the merging of the first two bins in different control regions for fits to data. The fits are performed under the background-only hypothesis and use the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in all regions. Points show the fitted values of the nuisance parameters and uncertainty bars show the posterior uncertainty of the nuisance parameters. The x-axis represents the pull of a nuisance parameter in units of the prior uncertainty. The green (yellow) region represents $\pm 1\sigma$ ($\pm 2\sigma$) uncertainties. In the top row theoretical nuisance parameters are shown, in the middle row non-b-tagging instrumental nuisance parameters are shown, and in the bottom row b-tagging nuisance parameters are shown. Bins are merged in 3 jet regions in the left plot, in ≥ 4 jet regions in the middle plot, and in all regions in the right hand plot. In the left and middle plots the red circles and blue points represent fits where bins are merged in a single analysis region which is labelled at the top of the plot. Nuisance parameters with small changes or no changes between the two fit models are not shown.

The effective removal of jet energy scale BJES and pile-up nuisance parameter pulls shows that the source of these large corrections is the disagreement with data in the first bins of the control regions. In particular the source of these pulls is dominantly from $t\bar{t}$+light enriched control regions. The lack of pulls for these nuisance parameters after bins have been merged, as well as reductions in many $t\bar{t}$ modelling pulls, is an indication of corrections only being applied for these small sections of phase space and resulting tensions in the fit model. On the other hand the introduction of pulls or increases in pulls for other nuisance parameters after control region bins have been merged is an indication that corrections at low $H_{\mathrm{T}}^{\mathrm{all}}$ are still required but that the fit model is no longer flexible enough to perform these corrections individually. Instead corrections are performed by changing the modelling of processes in the newly merged bins. In particular the fact that the overall normalisation for the Z+jets process is dramatically pulled after merging bins when it was previously approximately unpulled shows how inflexibility in the fit model can result in somewhat spurious corrections. As such, care must be taken when defining updated control region bin definitions to allow the fit model to make appropriate corrections without being overly sensitive to fine details of mismodelling in small regions of phase space.

New definitions of distribution bins

To observe the underlying details of mismodelling in control regions prior to choosing new bin definitions, plots were produced of $H_{\mathrm{T}}^{\mathrm{all}}$ distributions in control regions with a large number of narrow bins. The corresponding plots are shown in Figure 6.25. Features are seen in the ratios of data to predicted yields such as relatively larger mismodelling at low $H_{\mathrm{T}}^{\mathrm{all}}$ in $t\bar{t}$+light regions and slopes in the value of the ratio as a function of $H_{\mathrm{T}}^{\mathrm{all}}$. Bin choices could then be made to separate different features in the ratios into different bins while keeping the number of bins to a minimum. Prior to choosing new definitions of control region bins, studies were also performed where two bins or even a single inclusive bin were used for distribution binning. These studies tested the effect of removing shape information from the likelihood function. Other studies were also performed with regions completely removed from the fit model; this tested the removal of both shape and normalisation information. By performing these tests the contribution of each region

to nuisance parameter pulls due to shape and normalisation information could be tested. From tests where the 3 jet $\mathrm{CR}_{t\bar{t}+\geq 1b}$ was modified or removed it was seen that results from fits to data were almost unchanged aside from an increase in the posterior uncertainty on the $t\bar{t}+\geq 1c$ SHERPA5F nuisance parameter (fits at this stage were performed after the inclusion of the SHERPA5F-based $t\bar{t}$ uncertainties). This suggested that the 3 jet $\mathrm{CR}_{t\bar{t}+\geq 1b}$ does not contribute additional information to the likelihood function that is not measured in other analysis regions aside from some additional sensitivity to $t\bar{t}+\geq 1c$ modelling. As a result very simplified bin choices could be used for this region without loss of fit performance.

Based on by-eye judgement of the finely binned control region distributions seen in Figure 6.25, and the test of reducing the number of bins in the 3 jet $\mathrm{CR}_{t\bar{t}+\geq 1b}$ region, new bin choices were defined as shown in Table 6.6. Comparisons of the initial control region bin choices and the new control region bin choices are shown in Figure 6.26. The initial control region choices used the merged first two bins configuration. New bin choices were also defined for $H_{\mathrm{T}}^{\mathrm{all}}$ distributions used in signal regions following similar principles as the new binning for the control regions. The new bin choices for signal regions were defined as shown in Table 6.7 The differences in initial bin choices and new bin choices for signal regions are shown in Figure 6.27. These control region and signal region plots and the resulting fits are defined using an updated fit model which now includes $t\bar{t}$ uncertainties defined from variations of the SHERPA4F dataset.

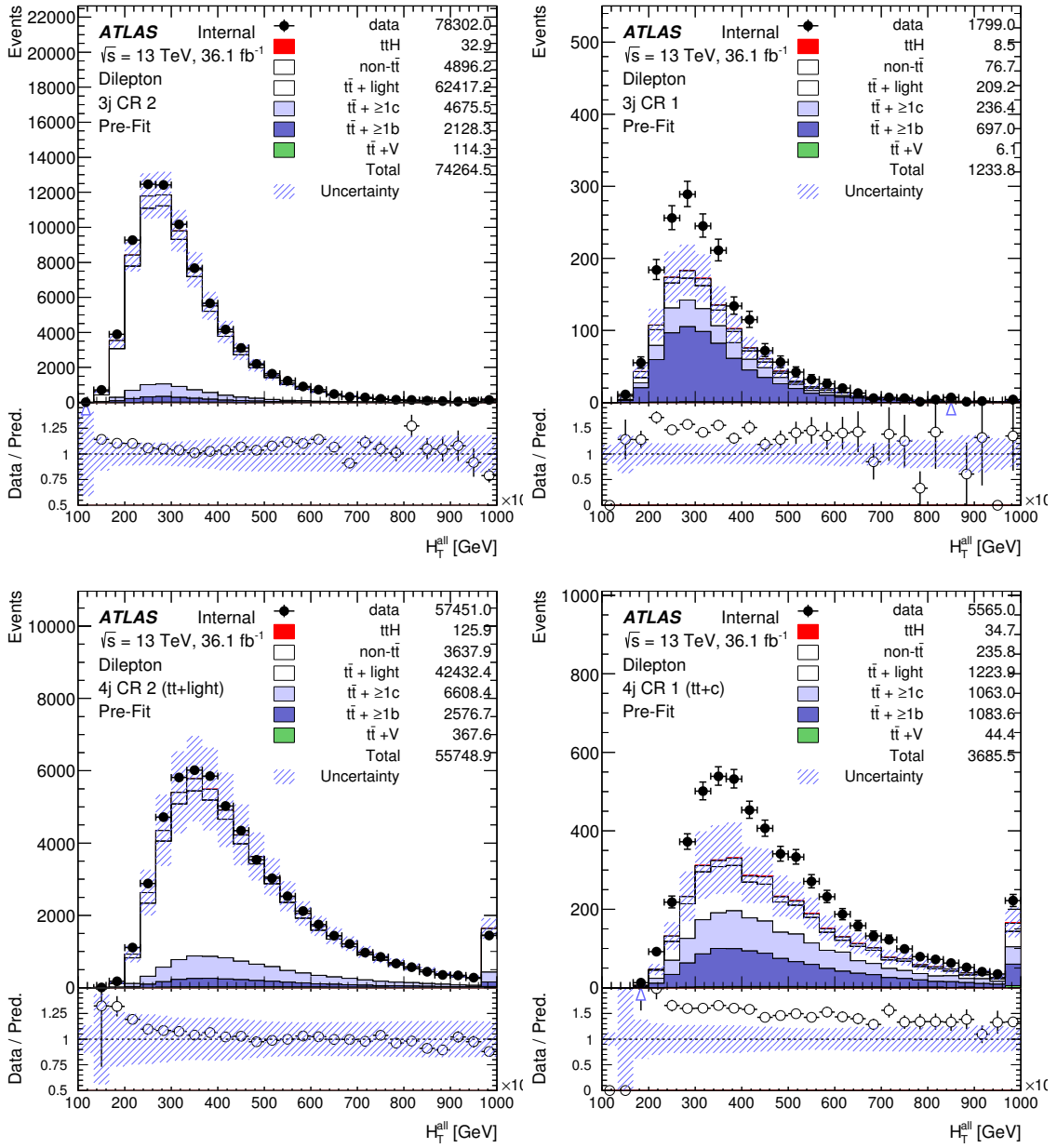


FIGURE 6.25: $H_{\mathrm{T}}^{\mathrm{all}}$ distributions with very fine binning in control regions. In the top row the 3 jet $\mathrm{CR}_{t\bar{t}+\mathrm{light}}$ and 3 jet $\mathrm{CR}_{t\bar{t}+\geq 1b}$ regions are shown. In the bottom row $\geq$ 4 jet $\mathrm{CR}_{t\bar{t}+\mathrm{light}}$ and $\geq$ 4 jet $\mathrm{CR}_{t\bar{t}+\geq 1c}$ regions are shown. An error in plotting resulted in the drawing of both the non-$t\bar{t}$ and the $t\bar{t}$ +light categories with a white area. The non-$t\bar{t}$ category is plotted at the top of the stack of background histograms. A ratio scale of 0 to 2 is used for $t\bar{t}$+HF regions to allow all data points to be seen.

TABLE 6.6: New binning definitions in control regions. Bin edges with a less-than-or-equal (greater-than-or-equal) sign in the value represent bin edges that define inclusive bins below (above) the bin edge value.

Region	Bin edges [GeV]
3 jet $CR_{t\bar{t}+\text{light}}$	$\leq 250, 350, 450, \geq 550$
3 jet $CR_{t\bar{t}+\geq 1b}$	$\leq 450, \geq 450$
≥ 4 jet $CR_{t\bar{t}+\text{light}}$	$\leq 250, 400, 600, \geq 800$
≥ 4 jet $CR_{t\bar{t}+\geq 1c}$	$\leq 300, 450, 600, \geq 850$

TABLE 6.7: New binning definitions in signal regions. Bin edges with a less-than-or-equal (greater-than-or-equal) sign in the value represent bin edges that define inclusive bins below (above) the bin edge value.

Region	Bin edges [GeV]
≥ 4 jet SR_3	$\leq 350, 450, 550, \geq 700$
≥ 4 jet SR_2	$\leq 350, 500, 650, \geq 800$
≥ 4 jet SR_1	$\leq 400, \geq 600$

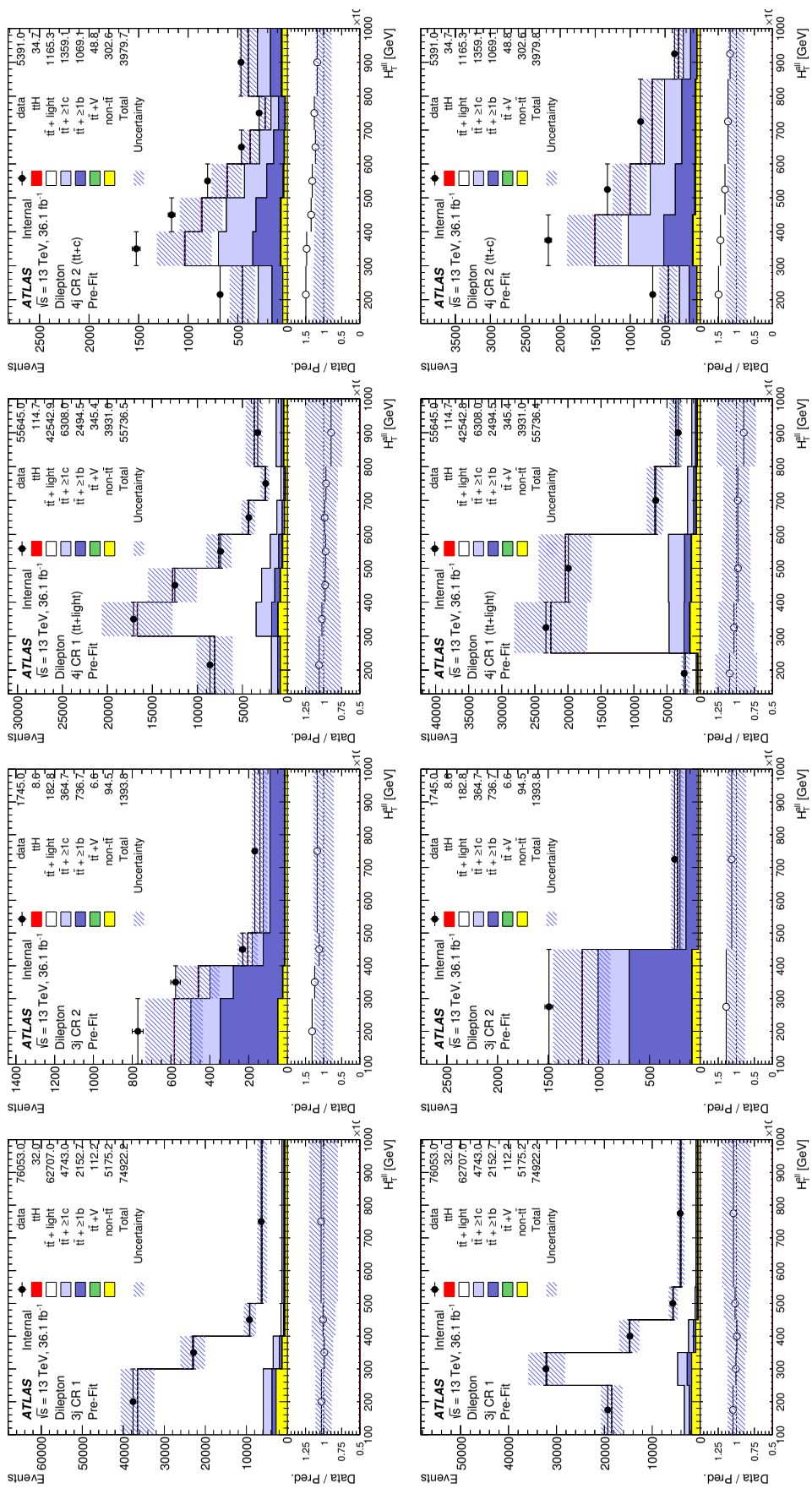


FIGURE 6.26: $H_{\mathrm{T}}^{\mathrm{all}}$ distributions in control regions. In the top (bottom) row distributions with initial (new) bin choices are shown. Control regions are arranged left to right as: 3 jet $\mathrm{CR}_{t\bar{t}+\mathrm{light}}$, 3 jet $\mathrm{CR}_{t\bar{t}+\geq 1b}$, ≥ 4 jet $\mathrm{CR}_{t\bar{t}+\mathrm{light}}$, ≥ 4 jet $\mathrm{CR}_{t\bar{t}+\geq 1c}$. A ratio scale of 0 to 2 is used for $t\bar{t}$+HF regions to allow all data points to be seen.

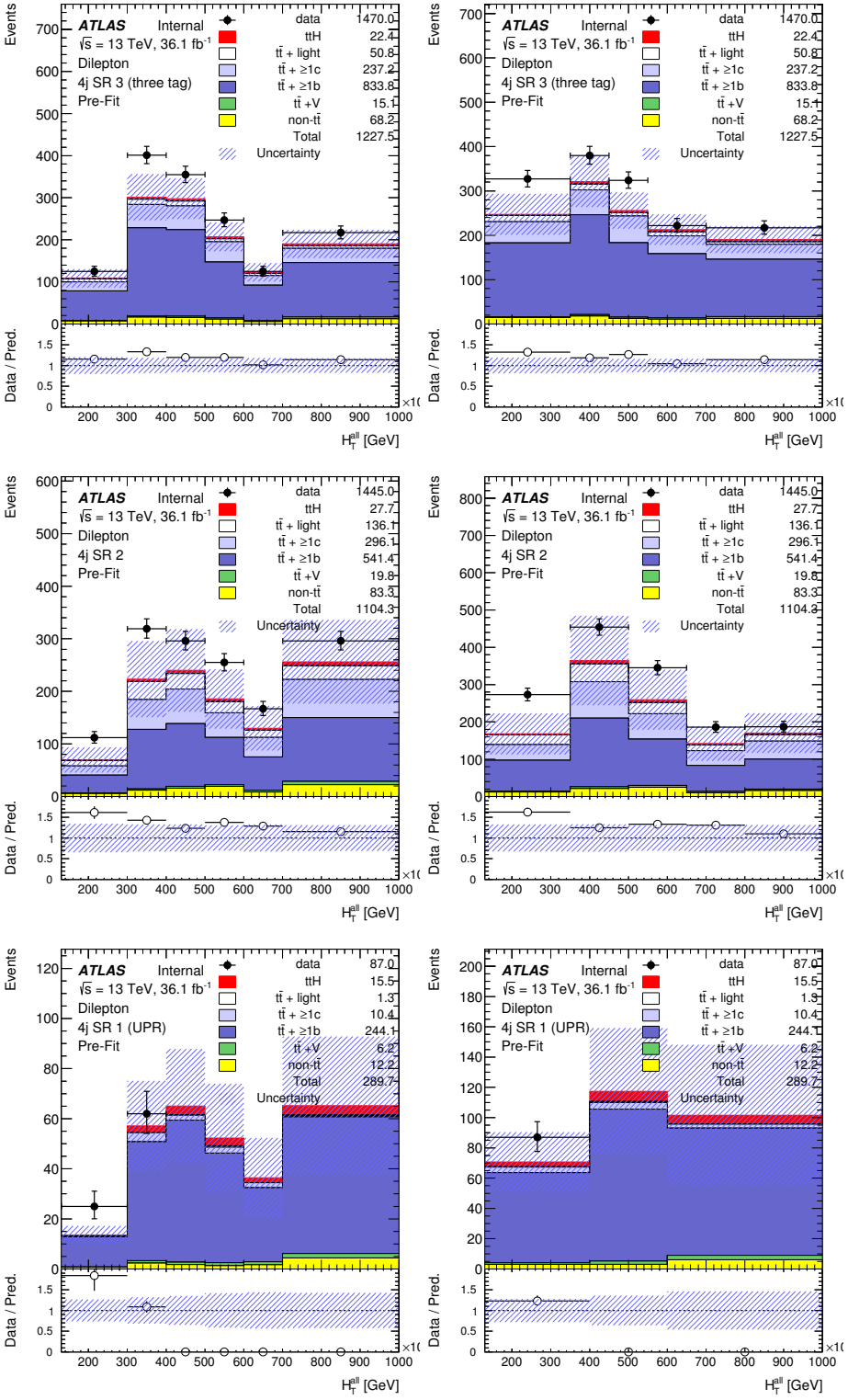


FIGURE 6.27: $H_{\mathrm{T}}^{\mathrm{all}}$ distributions in signal regions. In the left (right) column distributions with initial (new) bin choices are shown. In the top row results for the $\geq$ 4 jet SR_3 signal region are shown. In the middle row results for the $\geq$ 4 jet SR_2 signal region are shown. In the bottom row results for the $\geq$ 4 jet SR_1 signal region are shown. Bins without data points are blinded due to a signal purity of $> 5\%$ (points are still plotted in the ratio of data and prediction for these bins, however these points are not meaningful).

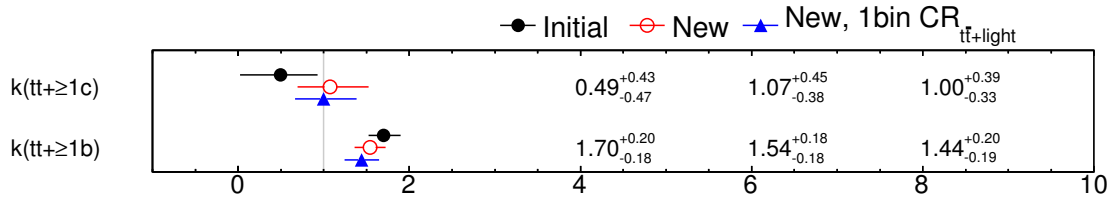


FIGURE 6.28: Fitted normalisation factors compared for initial and new bin choice configurations for fits to data. The fits are performed under the background-only hypothesis and use the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in all regions.

Tests of new bin choice configurations

Fits to data under the background-only hypothesis were performed using the initial bin choice configuration and using the new bin choice configuration. Fits were also performed with the new bin choices but with a single inclusive bin in the $t\bar{t}$ +light control regions (i.e. fitting only normalisation in these regions). This was tested due to the large pulls associated with shape modelling in these regions as well as tensions with fitted nuisance parameter values in other regions when fits which defined nuisance parameters individually in each region were trialled in the single-lepton channel. The comparison of each bin choice configuration is shown for the $k(t\bar{t}+\geq 1b)$ and $k(t\bar{t}+\geq 1c)$ normalisation factors in Figure 6.28 and for nuisance parameters which show significant changes between the different configurations in Figure 6.29.

It is seen that the new bin choice configuration reduces the very large pulls seen for the normalisation of Z+jets in 3 jet regions, the first light-jet mis-tagging nuisance parameter, and for Wt diagram subtraction. In other instrumental nuisance parameters, pulls remain similar or in some cases change sign with similar magnitude. The nuisance parameter for luminosity which controls the normalisation of all processes simultaneously is pulled downwards which suggests inflexibility in the fit model that is compensated for by correcting total normalisation instead. A significant increase in pull for the $t\bar{t}$ +light SHERPA5F-based uncertainty is also seen as well as an increase in the large pull on the $t\bar{t}+\geq 1b$ parton shower modelling nuisance parameter.

It is seen that when single inclusive bins are used in $t\bar{t}$ +light control regions that many instrumental nuisance parameter pulls are reduced in magnitude. The large pulls for the $t\bar{t}$ +light SHERPA5F-based uncertainty, Z+jets normalisation in 3 jet regions, and inclusive $t\bar{t}$ cross-section nuisance parameters are also greatly reduced. These changes in

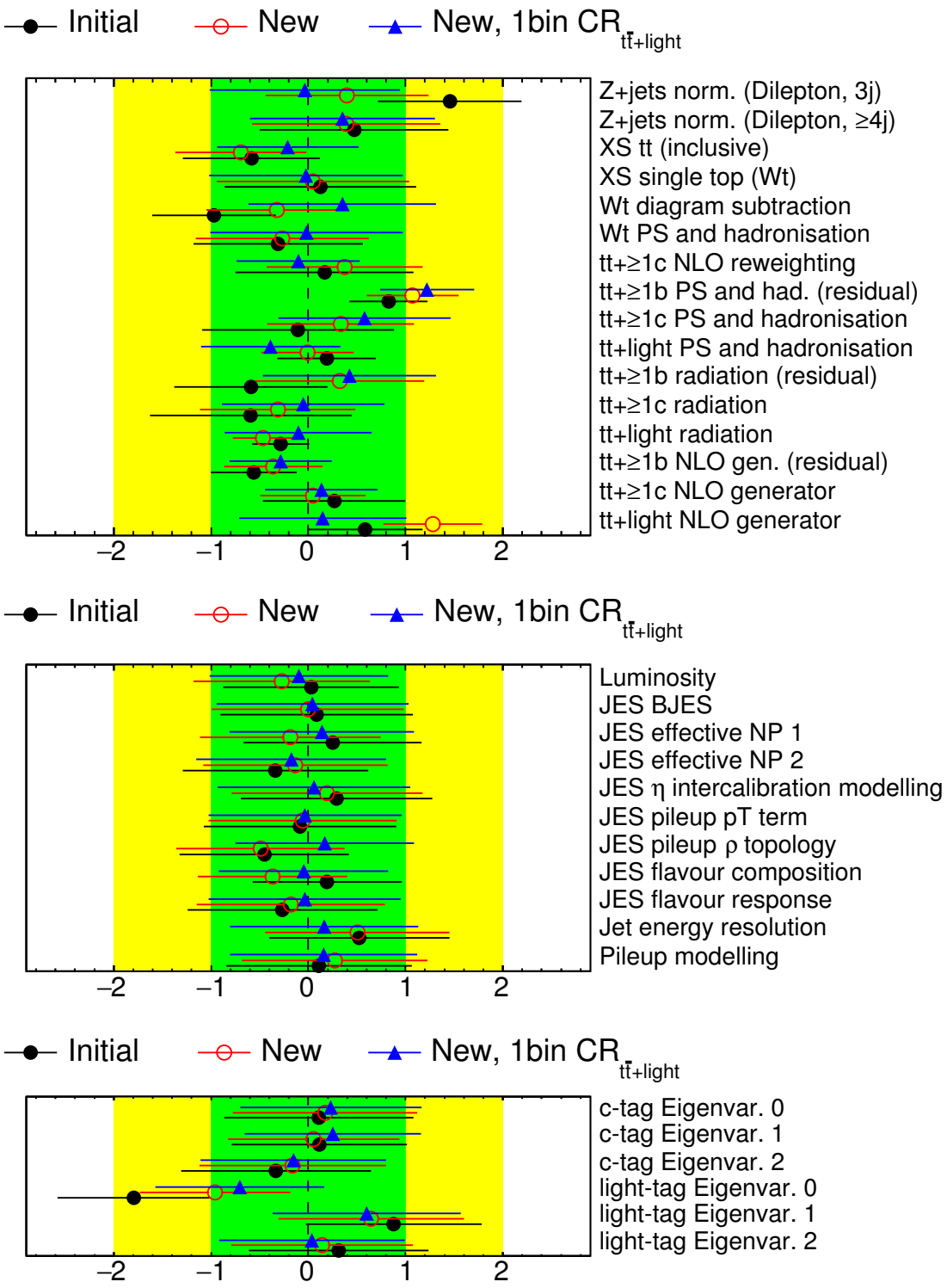


FIGURE 6.29: Fitted nuisance parameters compared for initial and new bin choice configurations for fits to data. The fits are performed under the background-only hypothesis and use the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in all regions. Points show the fitted values of the nuisance parameters and uncertainty bars show the posterior uncertainty of the nuisance parameters. The x-axis represents the pull of a nuisance parameter in units of the prior uncertainty. The green (yellow) region represents $\pm1\sigma$ ($\pm2\sigma$) uncertainties. In the top row theoretical nuisance parameters are shown, in the middle row non-b-tagging instrumental nuisance parameters are shown, and in the bottom row b-tagging nuisance parameters are shown. Nuisance parameters labelled "NLO gen." or "NLO generator" are actually the SHERPA5F-based nuisance parameters which model the effects of both matrix element generator and parton shower generator variations. Nuisance parameters with small changes or no changes between the two fit models are not shown.

nuisance parameter pulls show that the corrections associated with the initial pulls are dominantly coming from correcting distribution shapes in $t\bar{t}$ +light control regions and are not needed to correct mismodelling in other regions. Agreement between data and posterior predicted yields for the fitted distributions in each region was seen to be good for the use of new bin choices with and without inclusive single bins in $t\bar{t}$ +light control regions and little difference was seen between the two configurations. This confirmed that posterior agreement with data, particularly in signal regions, was not degraded by the exclusion of distribution shapes in $t\bar{t}$ +light control regions. As a result, the new binning configuration with inclusive single bins in $t\bar{t}$ +light control regions was adopted as the final binning used in control regions. The bin choices in the ≥ 4 jet SR_1 region were also changed in subsequent fits to data with H_T^{all} distributions in signal regions so that four bins were used in this region. The final bin edges for the ≥ 4 jet SR_1 region were: ≤ 400 GeV, 500 GeV, ≥ 700 GeV. This was done to match the expected number of bins for the use of BDT outputs in signal regions.

6.7.3 First Fits to Data Using Discriminant Distributions In Signal Regions

Once new classification BDTs had been trained and optimised for the pseudo-continuous b-tagging analysis region definitions, they could be tested in fits to data (classification BDTs from the 13.2 fb^{-1} analysis had been used instead when previously performing Asimov fits, as discussed at the start of Section 6.7.1). The agreement with data and resulting nuisance parameter pulls could then be compared to results from fits with H_T^{all} distributions in signal regions. The new BDTs were similar to the old BDTs in structure and in the list of input variables used, however the training of the new BDTs in the pseudo-continuous b-tagging analysis regions was expected to result in improved separation of signal and background. Asimov fits were performed comparing the performance of using the older unoptimised BDT outputs to using the new BDT outputs. It was seen that nuisance parameter constraints did not significantly change aside from a reduction in the posterior uncertainty on the $t\bar{t}+ \geq 1b$ SHERPA5F-based uncertainty. The uncertainty on μ was seen to reduce from $\sigma_\mu = ^{+1.08}_{-1.02}$ to $\sigma_\mu = ^{+0.83}_{-0.77}$. This decrease in uncertainty comes from the improved separation of signal events and background events when using the new optimised BDT.

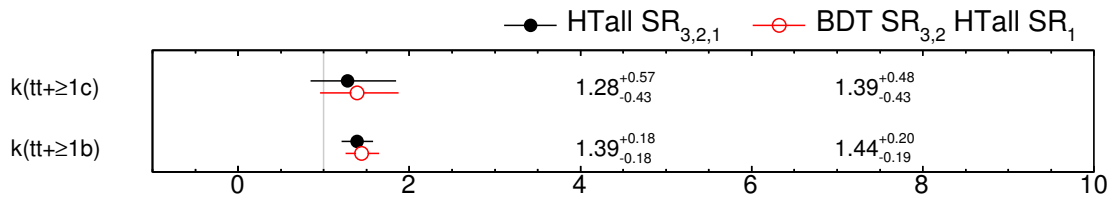


FIGURE 6.30: Fitted normalisation factors compared for fits to data with and without BDT outputs used in $\geq$ 4 jet SR_2 and $\geq$ 4 jet SR_3 regions. The fits are performed under the background-only hypothesis and use the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in all regions other than those using BDT outputs.

At this stage of the development of the $t\bar{t}H(b\bar{b})$ analysis results were still blinded to the presence of $t\bar{t}H$ signal. In prior fit results this involved performing fits under the background-only hypothesis and simply removing data yields from any bins with a signal purity in excess of 5% when plotting distributions. However due to the separating power of the BDT outputs additional care had to be taken when performing fits to data with BDT outputs used as distributions in signal regions. Due to the relatively high purity of signal in bins at high BDT output values, the presence of $t\bar{t}H$ signal could result in noticeably different nuisance parameter pulls in background-only fits. This would be particularly true if there was a $t\bar{t}H$ signal in data with a production cross-section larger than that of the Standard Model prediction. To avoid this sensitivity to signal while still observing corrections required for BDT output modelling the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution was used in the $\geq$ 4 jet SR_1 region while BDT outputs were used in the $\geq$ 4 jet SR_2 and $\geq$ 4 jet SR_3 regions. Background-only fits to data could then be performed and compared to background-only fits which used $H_{\mathrm{T}}^{\mathrm{all}}$ distributions in all signal regions. The comparison of background-only fits to data for the use of $H_{\mathrm{T}}^{\mathrm{all}}$ in all regions and for the use of classification BDTs in SR_3 and SR_2 are shown for the $k(t\bar{t}+\geq 1b)$ and $k(t\bar{t}+\geq 1c)$ normalisation factors in Figure 6.30 and for nuisance parameters which show significant changes between the different configurations in Figure 6.31. Smaller differences were also seen in other nuisance parameters such as reductions in pulls of non-leading light-jet mis-tagging nuisance parameters. Pre-fit and post-fit distributions for the fit using BDT outputs in signal regions 2 and 3 are shown in Figure 6.32.

In most cases posterior values and uncertainties for instrumental nuisance parameters were not seen to vary significantly between the use of $H_{\mathrm{T}}^{\mathrm{all}}$ distributions and BDT outputs in signal regions 2 and 3. However the significant pulls to leading light-jet mis-tagging nuisance parameters were seen to reduce. The fit model with BDTs included is

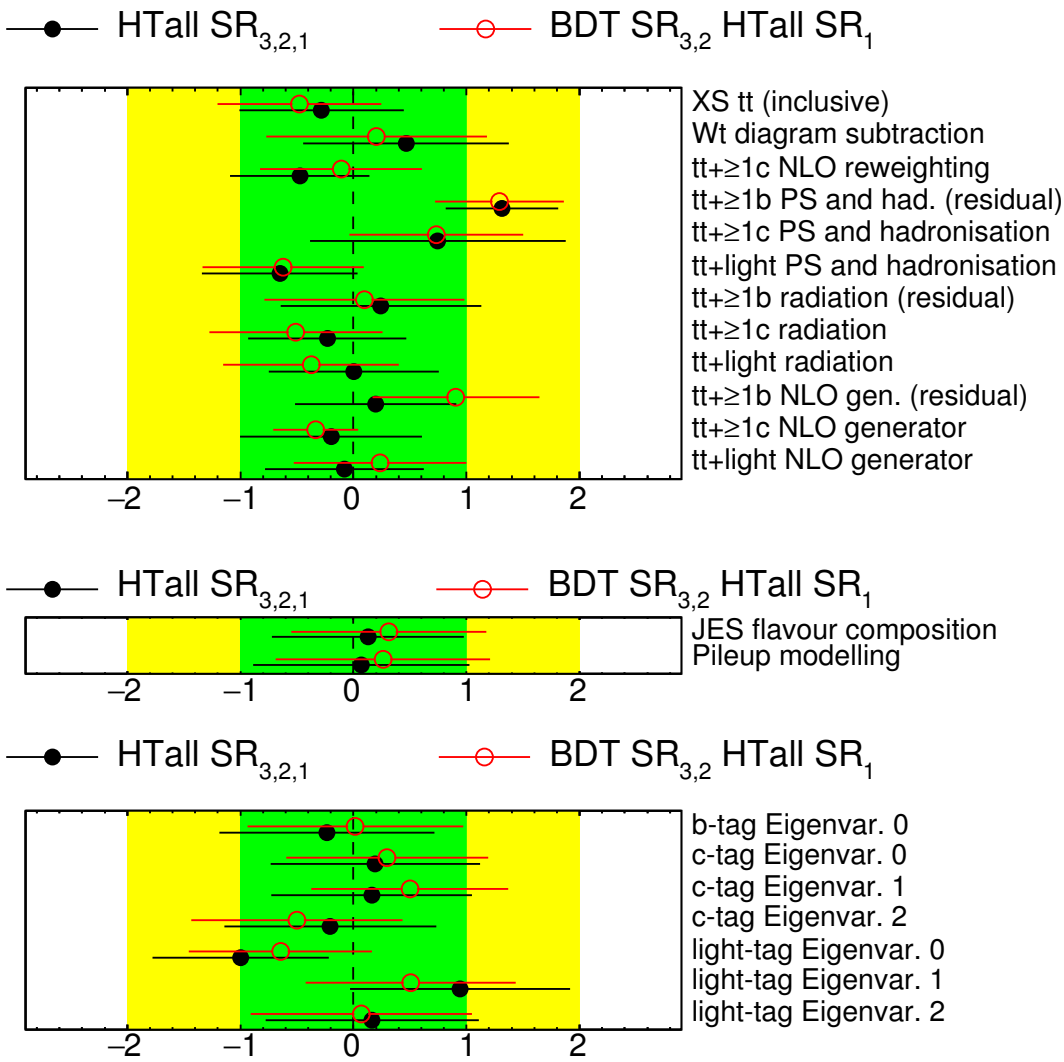


FIGURE 6.31: Fitted nuisance parameters compared for fits to data with and without BDT outputs used in $\geq$ 4 jet SR_2 and $\geq$ 4 jet SR_3 regions. The fits are performed under the background-only hypothesis and use the H_T^{all} distribution in all regions other than those using BDT outputs. Points show the fitted values of the nuisance parameters and uncertainty bars show the posterior uncertainty of the nuisance parameters. The x-axis represents the pull of a nuisance parameter in units of the prior uncertainty. The green (yellow) region represents $\pm1\sigma$ ($\pm2\sigma$) uncertainties. In the top row theoretical nuisance parameters are shown, in the middle row non-b-tagging instrumental nuisance parameters are shown, and in the bottom row b-tagging nuisance parameters are shown. Nuisance parameters labelled "NLO gen." or "NLO generator" are actually the SHERPA5F-based nuisance parameters which model the effects of both matrix element generator and parton shower generator variations. Nuisance parameters with small changes or no changes between the two fit models are not shown.

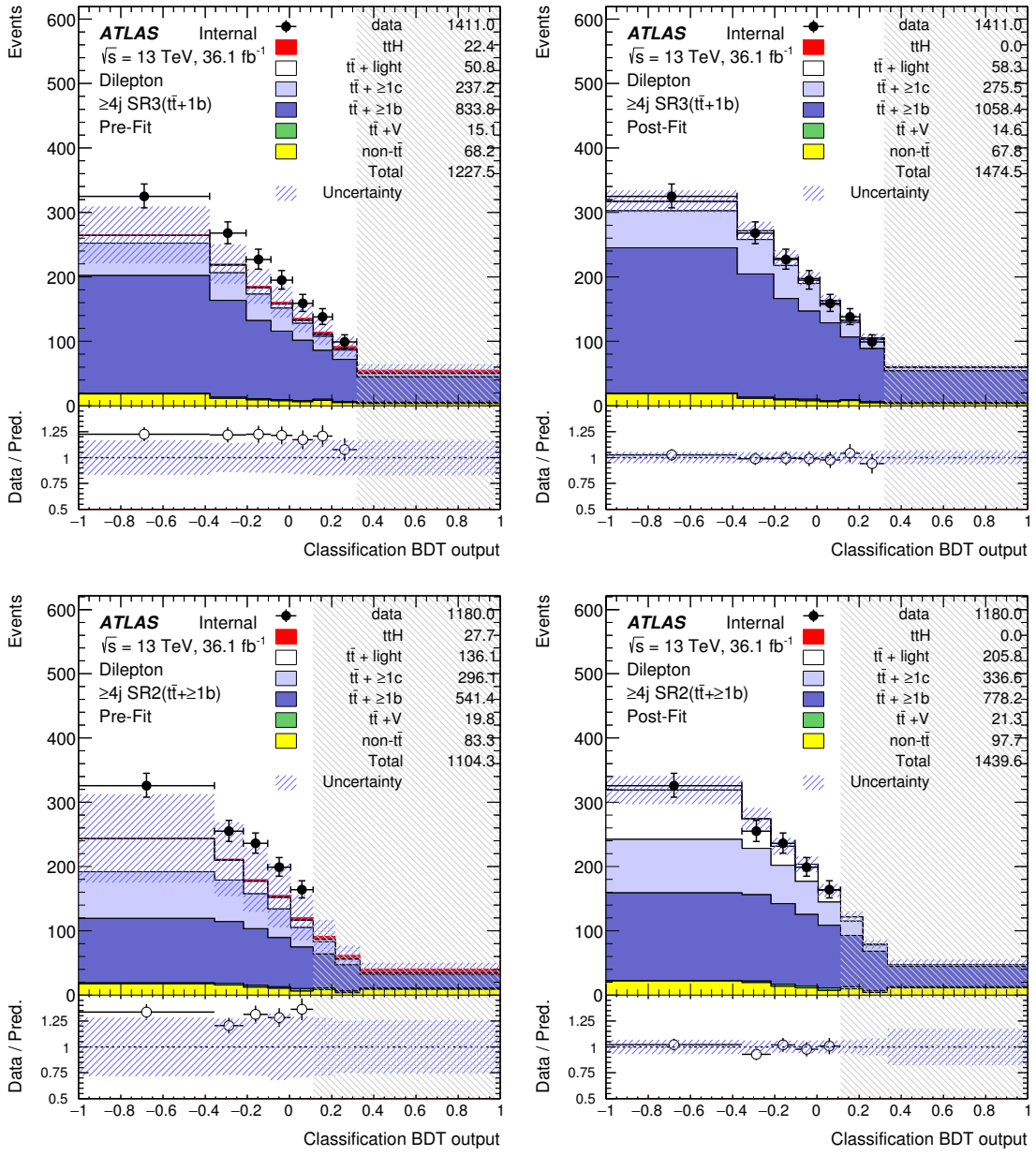


FIGURE 6.32: BDT output distributions in signal regions for fits to data under the background-only hypothesis in which $H_{\mathrm{T}}^{\mathrm{all}}$ was used in the ≥ 4 jet SR_1 signal region. In the left (right) column pre-fit (post-fit) distributions are shown. In the top (bottom) row results for the ≥ 4 jet SR_3 (≥ 4 jet SR_2) signal region are shown. Shaded grey regions are bins that are blinded due to a signal purity of $> 5\%$

seen to be significantly more sensitive to $t\bar{t}+ \geq 1c$ parton shower and SHERPA5F uncertainties and $t\bar{t}+ \geq 1c$ and c-jet mis-tagging nuisance parameter pulls are generally seen to increase in magnitude. The pull for the $t\bar{t}+ \geq 1b$ SHERPA5F nuisance parameter is seen to significantly increase and is seen to use a posterior value similar to the magnitude of the upward prior uncertainty.

It is therefore seen that BDTs in signal regions have similar sensitivity to instrumental modelling and similar levels of agreement with data, aside from changes to flavour tagging parameters. It is likely that the increase in c-jet mis-tagging nuisance parameter pulls originates from the fact that the BDT output will have better separation between $t\bar{t}$light, $t\bar{t}+ \geq 1c$, and $t\bar{t}+ \geq 1b$ compared to the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution, resulting in increased sensitivity to the modelling of $t\bar{t}+ \geq 1c$ events. Related to this fact, it is also seen that the use of BDTs in signal regions results in increased sensitivity to $t\bar{t}$+HF modelling parameters.

6.7.4 Decorrelation Of Jet Energy Resolution Uncertainties

An important test of the sensitivity of the statistical fit to the presence of $t\bar{t}H$ signal is the breakdown of the posterior uncertainty on μ into components from individual nuisance parameters. This breakdown, described as 'ranking', is performed by fixing each nuisance parameter to its upward and downward prior and posterior values and performing a new fit with the remaining parameters for each of these four value choices. The measured value of μ from each fit can then be compared to the measured value from the original fit to determine the effect of each nuisance parameter on μ.

Ranking plots listing the nuisance parameters with largest contributions to the uncertainty on μ were produced using a fit model similar to the one used in Section 6.7.3 for fits to the Asimov dataset. The fit model differed by the use of 12 bins in the ≥ 4 jet SR_2 region instead of eight. The ranking plot from the Asimov fit is shown in Figure 6.33. It can be seen that the nuisance parameter with the largest effect on the fitted value of μ is the jet energy resolution. This sensitivity to jet energy resolution was not expected due to the previously low ranking of this nuisance parameter in the 13.2 fb^{-1} analysis results. Studies were therefore performed to understand the source of the large effect on μ.

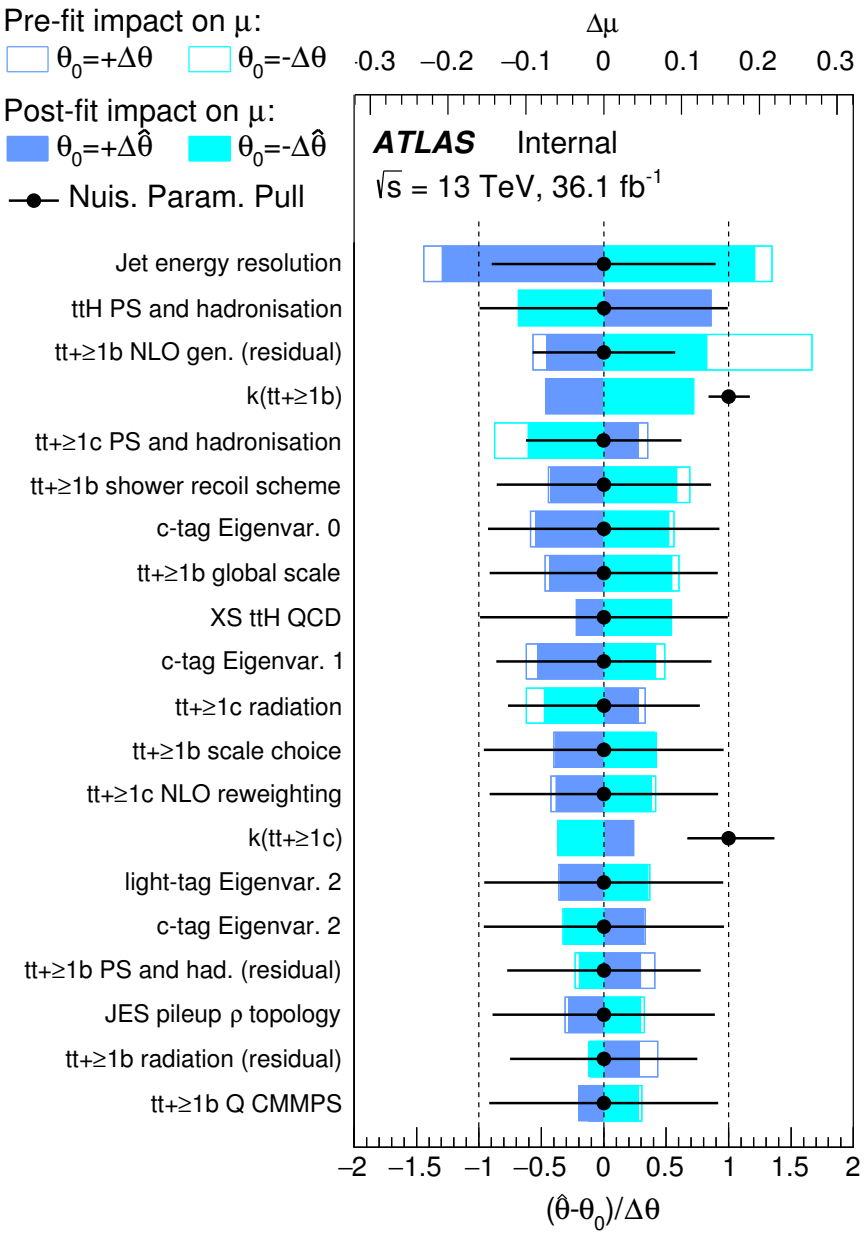


FIGURE 6.33: Ranking plot for a fit to the Asimov dataset. The top axis shows the change in the measured value of μ when the nuisance parameters are varied by the size of their initial (final) uncertainties, with the corresponding results shown as hollow (solid) blue bars. The bottom axis shows the change in the posterior values of the nuisance parameters relative to the size of their prior uncertainty. The black bars on each nuisance parameter show the size of their posterior uncertainty.

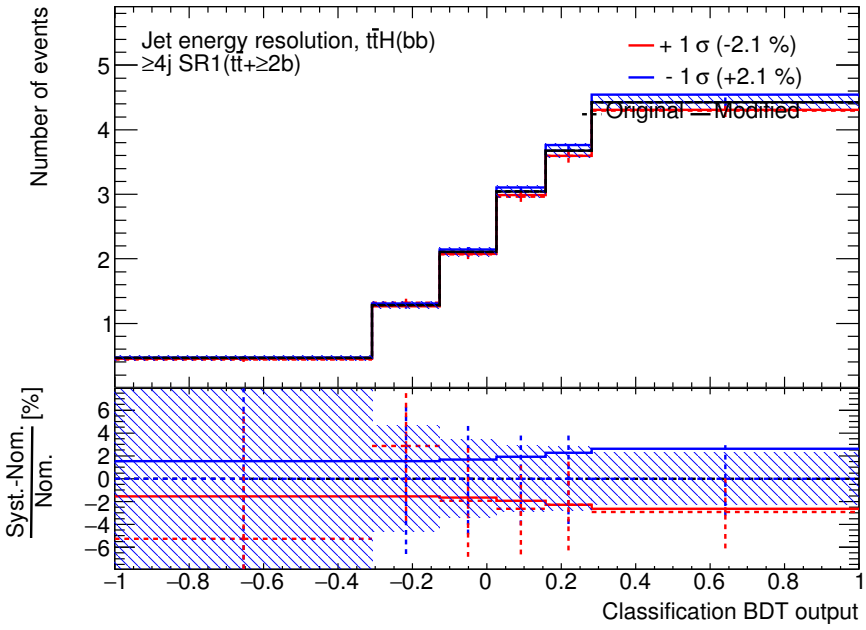
TABLE 6.8: Covariance between $t\bar{t}H$ signal strength and jet energy resolution when excluding the effect of jet energy resolution in signal regions. θ_{JER} is the nuisance parameter associated with jet energy resolution. "Change in covariance" is the actual change relative to the "None" scenario.

Region excluded from θ_{JER}	Covariance(μ, θ_{JER})	Change in covariance
None	-28.0%	-
≥ 4 jet SR_3	-24.4%	3.6%
≥ 4 jet SR_2	-25.2%	2.8%
≥ 4 jet SR_1	-10.3%	17.7%

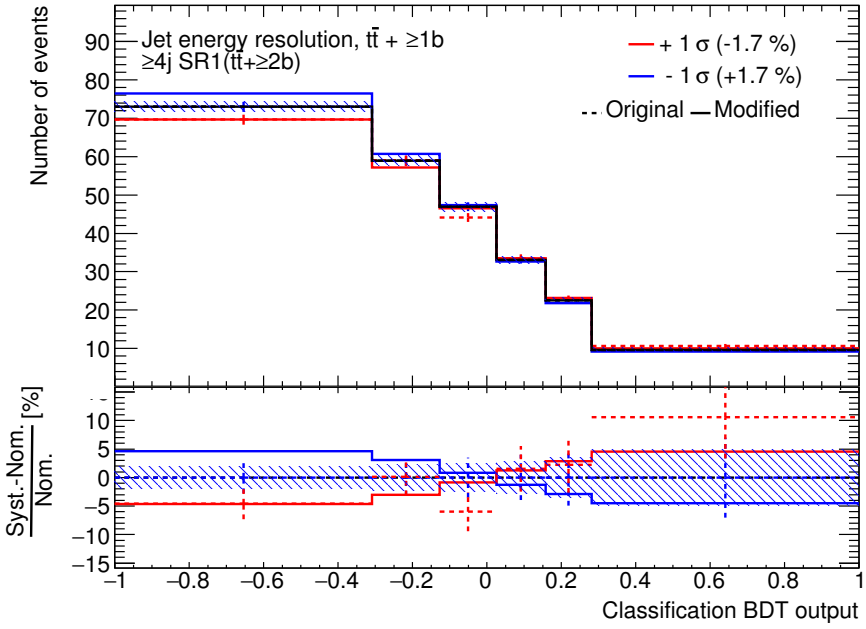
To understand this effect the covariance between μ and the jet energy resolution resolution was studied. The covariance between μ and jet energy resolution was seen to be -28%. This was significantly greater than the next largest covariance of 18.1% associated with the $t\bar{t}H$ parton shower modelling nuisance parameter, which was ranked second. To understand the source of this strong anti-correlation in more detail the effect of the jet energy resolution nuisance parameter was removed from each signal region in turn and the fit was rerun. By observing the decrease in covariance the contribution from each region could be determined. The resulting values for the covariance of μ with the jet energy resolution nuisance parameter are shown in Table 6.8.

It is seen that the dominant contribution to the covariance comes from the ≥ 4 jet SR_1 region while regions SR_2 and SR_3 provide small and roughly similar contributions. A ranking plot produced from the fit with the effect of jet energy resolution excluded in the SR_1 region also showed that the jet energy resolution parameter was ranked 13^{th} under this configuration while it remained the leading effect in the other two configurations. To understand the strong correlation in the SR_1 region the effect of the jet energy resolution uncertainty was plotted for both the $t\bar{t}H(b\bar{b})$ signal and $t\bar{t}+ \geq 1b$ process which is dominant in the SR_1 region. The resulting systematic uncertainty plots are shown in Figure 6.34.

It can be seen that the effect of the jet energy resolution systematic uncertainty on binned yields is in the opposite direction for $t\bar{t}H(b\bar{b})$ and $t\bar{t}+ \geq 1b$ with the relatively largest effect in the bin with highest signal purity. This results in a change to signal purity of approximately 3% in the bin covering the highest BDT output values. The fact that the effect of jet energy resolution seen in the ranking plot (approximately 20%) is much larger



a)



b)

FIGURE 6.34: Effect of the jet energy resolution systematic uncertainty in the $\geq$ 4 jet SR_1 signal region on the BDT output distribution. The effects are shown for *a)* the $t\bar{t}H(b\bar{b})$ process and *b)* the $t\bar{t}+\geq 1b$ process. The black line indicates the nominal distribution, the red line indicates the upward uncertainty variation, and the blue line indicates the downward uncertainty variation. The dashed red lines indicate the original systematic uncertainty estimates before smoothing is applied to allow the effect of smoothing to be seen. No dashed blue lines are shown because the original uncertainty estimate is only defined in one direction; the presence of the solid blue line is the result of symmetrisation after smoothing has been applied.

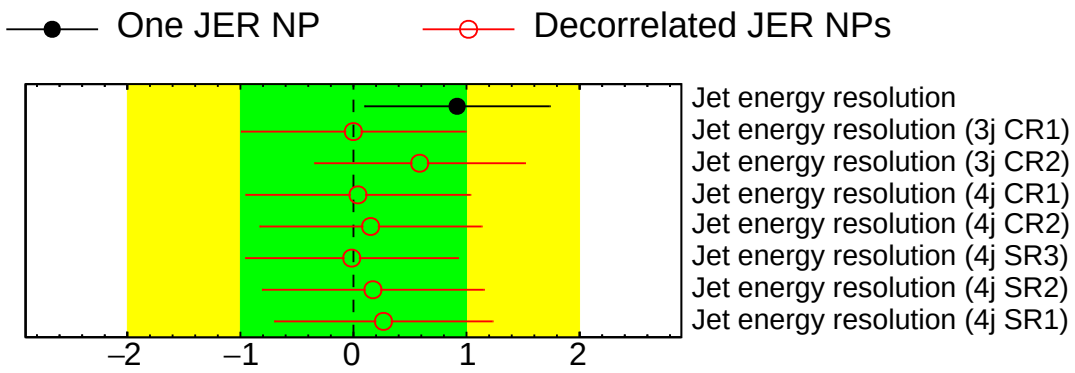


FIGURE 6.35: Fitted jet energy resolution nuisance parameters for fits to data with a single nuisance parameter and with decorrelated nuisance parameters. The fits are performed under the background-only hypothesis and use BDT outputs in signal regions with bins that have a signal purity of $> 5\%$ removed from the fit model. Points show the fitted values of the nuisance parameters and uncertainty bars show the posterior uncertainty of the nuisance parameters. The x-axis represents the pull of a nuisance parameter in units of the prior uncertainty. The green (yellow) region represents $\pm 1\sigma$ ($\pm 2\sigma$) uncertainties. "CR1" refers to $t\bar{t}$ +light control regions and "CR2" refers to $t\bar{t}$ +HF control regions.

than this change suggests that the fit model is not flexible enough to correct the changes to $t\bar{t}$ in all bins in the SR_1 region appropriately and instead results in significantly more or less $t\bar{t}$ in the most signal-sensitive bins which implies a significant increase in $t\bar{t}H$ to match the modified dataset.

Studies of jet energy resolution nuisance parameters in individual regions

Given the potential for pulls of the jet energy resolution nuisance parameter to affect the final measured value of $t\bar{t}H$ signal strength, studies were later performed in fits to data to understand the source and magnitude of any pull seen. These tests were performed with a later fit model which was similar to the one used for the studies of nuisance parameter ranking aside from minor updates. The studies were performed by defining the effects of the jet energy resolution uncertainties as a number of decorrelated nuisance parameters which control the variation only in a single analysis region. From the individual-region nuisance parameters it could be observed which regions contributed pulls to the combined jet energy resolution nuisance parameter. Fits to data were performed under the background-only hypothesis with BDT outputs in the signal regions. To avoid sensitivity to the presence of signal in nuisance parameter pulls any bins with a signal purity of $> 5\%$ are removed from the fit model prior to performing the fit. The resulting measured jet energy resolution nuisance parameters are shown in Figure 6.35.

It is seen that the single jet energy resolution nuisance parameter has a large upward pull. From the decorrelated nuisance parameters it is seen that this pull is dominated by the 3 jet $CR_{t\bar{t}+\geq 1b}$ region and has smaller contributions of approximately similar magnitude from the $\geq$ 4 jet $CR_{t\bar{t}+\geq 1c}$, SR_2, and SR_1 regions. Given the discrepancy between the nuisance parameter pulls and the sensitivity of the measured $t\bar{t}H$ signal strength to jet energy resolution it was decided to use a separate nuisance parameter for jet energy resolution in the 3 jet $CR_{t\bar{t}+\geq 1b}$ region and a single combined nuisance parameter across all other regions.

6.7.5 Understanding The Effects of Parton Shower Nuisance Parameter Pulls Using Modified Asimov Fits

To further understand the effect of nuisance parameter pulls on the measured $t\bar{t}H$ signal strength the effect of the large pull for the $t\bar{t}+\geq 1b$ parton shower and hadronisation modelling nuisance parameter was studied. To check the effect of the posterior nuisance parameter value seen in fits to data, tests were performed using Asimov datasets that were modified to use the value from data fits. Two tests were performed: the first, labelled as an 'injected nuisance parameter pull' test, created a modified Asimov dataset with the $t\bar{t}+\geq 1b$ parton shower nuisance parameter set to the value seen from data fits. This modified dataset could then be treated as data and the unmodified Asimov dataset could be fit to it. In this way the ability of the fit model to correct for a true value of the nuisance parameter that is different from the prior value could be tested and any resulting change in the measured value of μ could be observed. The second test, labelled as a 'fixed nuisance parameter pull' test, uses the same method as the calculation of nuisance parameter ranking except that the value from fits to data is used instead of the values associated with upward and downward nuisance parameter uncertainties. The prior dataset was created with the $t\bar{t}+\geq 1b$ parton shower nuisance parameter fixed to the value extracted from data fits. This modified prior dataset was then fit to the standard Asimov dataset and any resulting change in the measured value of μ could be observed. This measured the correlation between the parton shower nuisance parameter and μ. The value of the $t\bar{t}+\geq 1b$ parton shower and hadronisation modelling nuisance parameter taken from data was 1.45.

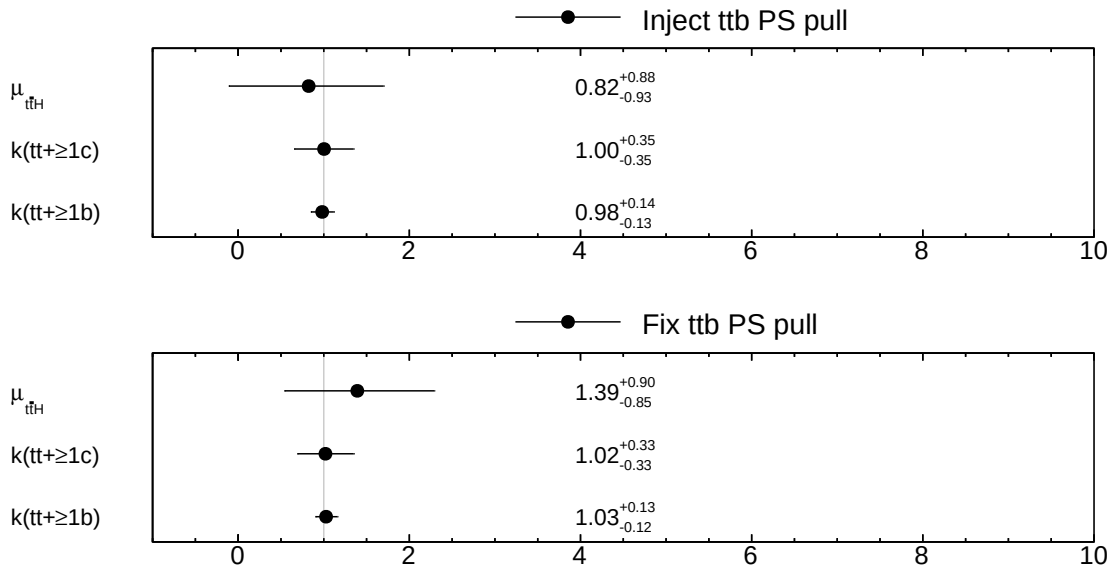


FIGURE 6.36: Fitted $t\bar{t}H$ signal strength and normalisation factors for modified Asimov fits with $t\bar{t}+\geq 1b$ parton shower nuisance parameter pulls. In the top plot the fitted parameters are shown for a fit to an Asimov dataset modified with the parton shower nuisance parameter pull injected. In the bottom plot the fitted parameters are shown for a fit to the Asimov dataset using a fit model where the prior parton shower nuisance parameter has been fixed to the pulled value.

The results from each test are shown for the fitted $t\bar{t}H$ signal strength and $k(t\bar{t}+\geq 1b)$ and $k(t\bar{t}+\geq 1c)$ normalisation factors in Figure 6.36 and for nuisance parameters which show significant pulls in Figure 6.37. For the injected nuisance parameter pull test pulls for nuisance parameters which aren't plotted are very small. However for the fixed nuisance parameter pull test a number of instrumental nuisance parameters such as pile-up modelling and jet energy scale nuisance parameters show noticeable though still small (approximately less than ± 0.2) pulls.

It is seen for the injected nuisance parameter pull test that the $t\bar{t}+\geq 1b$ parton shower nuisance parameter is not measured to be the same as the injected value. Instead a value of 0.74 is measured and a number of smaller pulls are seen in other $t\bar{t}$ modelling nuisance parameters. The inability of the fit model to correct for a single very large pull is not surprising due to the penalty associated with large changes to individual nuisance parameters due to the constraint term in Equation 6.6. Instead the value of $\ln L$ is maximised for smaller changes to a number of nuisance parameters. It is seen that the $t\bar{t}H$ signal strength is pulled away from the true value of one by approximately -0.2. The effect of an upward pull for the $t\bar{t}+\geq 1b$ parton shower nuisance parameter is to reduce the $t\bar{t}+\geq 1b$ yield in signal-rich bins so it is likely that the reduction of μ is performing

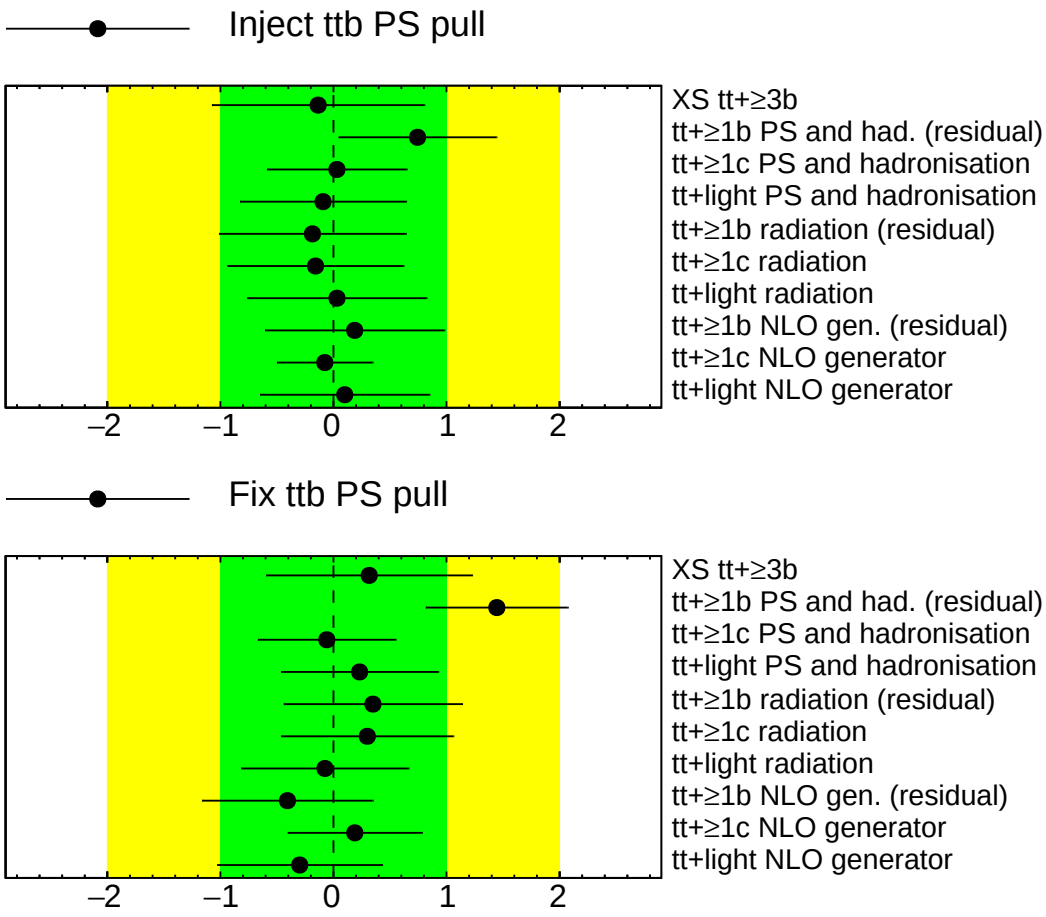


FIGURE 6.37: Fitted nuisance parameters for modified Asimov fits with $t\bar{t}+ \geq 1b$ parton shower nuisance parameter pulls. Points show the fitted values of the nuisance parameters and uncertainty bars show the posterior uncertainty of the nuisance parameters. The x-axis represents the pull of a nuisance parameter in units of the prior uncertainty. The green (yellow) region represents $\pm 1\sigma$ ($\pm 2\sigma$) uncertainties. In the top plot the fitted parameters are shown for a fit to an Asimov dataset modified with the parton shower nuisance parameter pull injected. In the bottom plot the fitted parameters are shown for a fit to the Asimov dataset using a fit model where the prior parton shower nuisance parameter has been fixed to the pulled value. Nuisance parameters labelled "NLO gen." or "NLO generator" that are plotted with red hollow points are actually the SHERPA5F-based nuisance parameters which model the effects of both matrix element generator and parton shower generator variations. Nuisance parameters with small changes or no changes away from their initial estimates are not shown.

residual corrections of yields in these bins.

For the fixed nuisance parameter pull test it is seen that pulls are roughly mirrored from the injected nuisance parameter pull test but have larger magnitudes. The definitions of the two tests are effectively inversions of each other aside from the fixing of the parton shower nuisance parameter in the second test and so this shows they are approximately consistent with one another. Due to the need to correct for the prior effect of the fixed parton shower nuisance parameter when fitting to the standard Asimov dataset the pulls are inverted and the increase in magnitude follows from the larger parton shower pull applied. It is seen that the $t\bar{t}H$ signal strength is pulled away from the true value of one by approximately 0.4. As such it can be said that the value for the $t\bar{t}+ \geq 1b$ parton shower nuisance parameter which is measured in data results in an increase to the measured $t\bar{t}H$ signal strength of approximately 0.4.

6.8 Results

To determine a final measured value for $t\bar{t}H$ signal strength the single-lepton and dilepton analyses are combined into a single binned profile likelihood fit with most systematic uncertanties included as single nuisance parameters applied to both analyses [90]. Some systematic uncertanities were only defined in one analysis or were defined differently in each analysis, such as different uncertainties for the different fake lepton estimation methods, and as such were included as nuisance parameters which were individual to each analysis. In this way information from one channel can be used to constrain measurements in the second channel. The likelihood fit is performed with a single $t\bar{t}H$ signal strength parameter of interest but is also checked with a fit where each channel determines a separate $t\bar{t}H$ signal strength. To test the agreement of the observed data with the background-only hypothesis and to determine upper limits on the $t\bar{t}H$ signal strength a test statistic, t_μ, is defined as the profile likelihood ratio:

$$t_\mu = -2\ln\lambda(\mu) \equiv -2\ln\left[\frac{L(\mu,\hat{\hat{\theta}}(\mu))}{L(\hat{\mu},\hat{\hat{\theta}})}\right] \tag{6.10}$$

where the numerator is the same type of profile likelihood as used for MINOS uncertainties in Equation 6.5 and the denominator is the maximised likelihood resulting from

the binned profile likelihood fit. MINOS uncertainties on μ can be defined as the changes in μ away from $\hat{\mu}$ that increase t_μ by one.

A number of changes were made to the dilepton analysis between the fit studies presented previously in this section and the final published results. The most significant of these changes was reduction in the number of bins in the ≥ 4 jet SR_1 region, the use of single inclusive bins in all dilepton analysis control regions, and the correction of a bug in the evaluation of the classification BDT outputs. For the first of these changes the number of bins in the ≥ 4 jet SR_1 region was reduced from six to four. This resulted in decreased sensitivity to the jet energy resolution nuisance parameter for the measured $t\bar{t}H$ signal strength. The second change of using single inclusive bins in all dilepton analysis control regions was done after further tests were performed decorrelating nuisance parameters between regions as well as decorrelating the shape and normalisation effects of nuisance parameters and it was seen that mismodelling corrections from control regions were in tension with nuisance parameter pulls from signal regions. This suggested that the kinematic modelling in control regions and signal regions was too different to extrapolate between these regions and so single inclusive bins were used in control regions instead to only constrain the normalisation of processes. The third change of the correction of a bug in the evaluation of the classification BDT outputs is the same issue described for the 13.2 fb^{-1} $t\bar{t}H(b\bar{b})$ analysis in Section 6.1: due to a file-naming error the BDT output was calculated using a BDT that was trained with the same events that were used for the evaluation of the BDT output, which resulted in strongly biased BDT output values. The use of the correct BDT resulted in an increase on the uncertainty on μ in Asimov fits from $\sigma_\mu = ^{+0.89}_{-0.82}$ to $\sigma_\mu = ^{+1.19}_{-1.19}$. The discovery of this bug resulted in the subsequent discovery of the same bug in the 13.2 fb^{-1} analysis and the retraction of that result.

The single-lepton analysis also incorporates additional signal-sensitive variables and a sub-channel which allows easier identification of jets from the top quarks and from the Higgs boson. The additional signal-sensitive variables are a likelihood discriminant built from the probability density functions for the signal and background predictions and a matrix element method (MEM) discriminant which calculates signal and background likelihoods for each event using matrix element level calculations from simulation. These variables are used as inputs into classification BDTs in the single-lepton analysis to improve the separation of signal and background in the output. The sub-channel used in

the single-lepton analysis is labelled as the 'boosted' channel. This sub-channel selects high p_{T} events where the hadronic jets originating from the top quarks or Higgs boson are collimated enough that they can be reconstructed as a single grouped 'large-R jet'. By selecting events containing large-R jets, a classification BDT can be built using the kinematic properties of the large-R jets without the need for a reconstruction BDT to identify top quark and Higgs boson candidates. The boosted channel is treated as an additional analysis region in the single-lepton channel.

Pre-fit and post-fit distributions in the dilepton channel for the final fit to data using both the dilepton and single-lepton channels and performed under the signal-plus-background hypothesis are shown in Figure 6.38. The $k(t\bar{t}+\geq 1b)$ and $k(t\bar{t}+\geq 1c)$ normalisation factors are measured to be 1.24 ± 0.10 and 1.63 ± 0.23 respectively. The measured $t\bar{t}H$ signal strengths from the combined fit as well as from the fit using individual signal strengths for the dilepton and single-lepton channels are shown in Figure 6.39. It is seen that the combined $t\bar{t}H$ signal strength is measured to be $\mu = 0.84^{+0.64}_{-0.61}$ and the dilepton channel signal strength from the individual channel signal strengths fit is seen to be $\mu_{\text{dilep.}} = -0.24^{+1.02}_{-1.05}$. It was also observed that the dilepton signal strength from a fit exclusively to the dilepton channel was observed to be $\mu = 0.11^{+1.36}_{-1.41}$.

To determine the significance of the measured $t\bar{t}H$ signal strength a second test statistic, q_0, based on the test statistic defined in Equation 6.10 is used [152]. The test statistic calculates the profile likelihood ratio for the background-only hypothesis under the definition that $\hat{\mu} \geq 0$. The test statistic is defined as:

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{6.11}$$

With sufficiently large data yields the measurement is said to be in the 'asymptotic regime' and the probability distribution function for q_0 under the background-only hypothesis is a chi-square distribution. By appropriately integrating this distribution a p-value can be calculated for the measured $t\bar{t}H$ signal strength. The p-value can then be turned into a measure of discovery significance by defining the p-value as the integral of the upper-tail of a normal gaussian distribution and defining the discovery significance as the number of gaussian σ that the lower bound of the p-value integral is away from

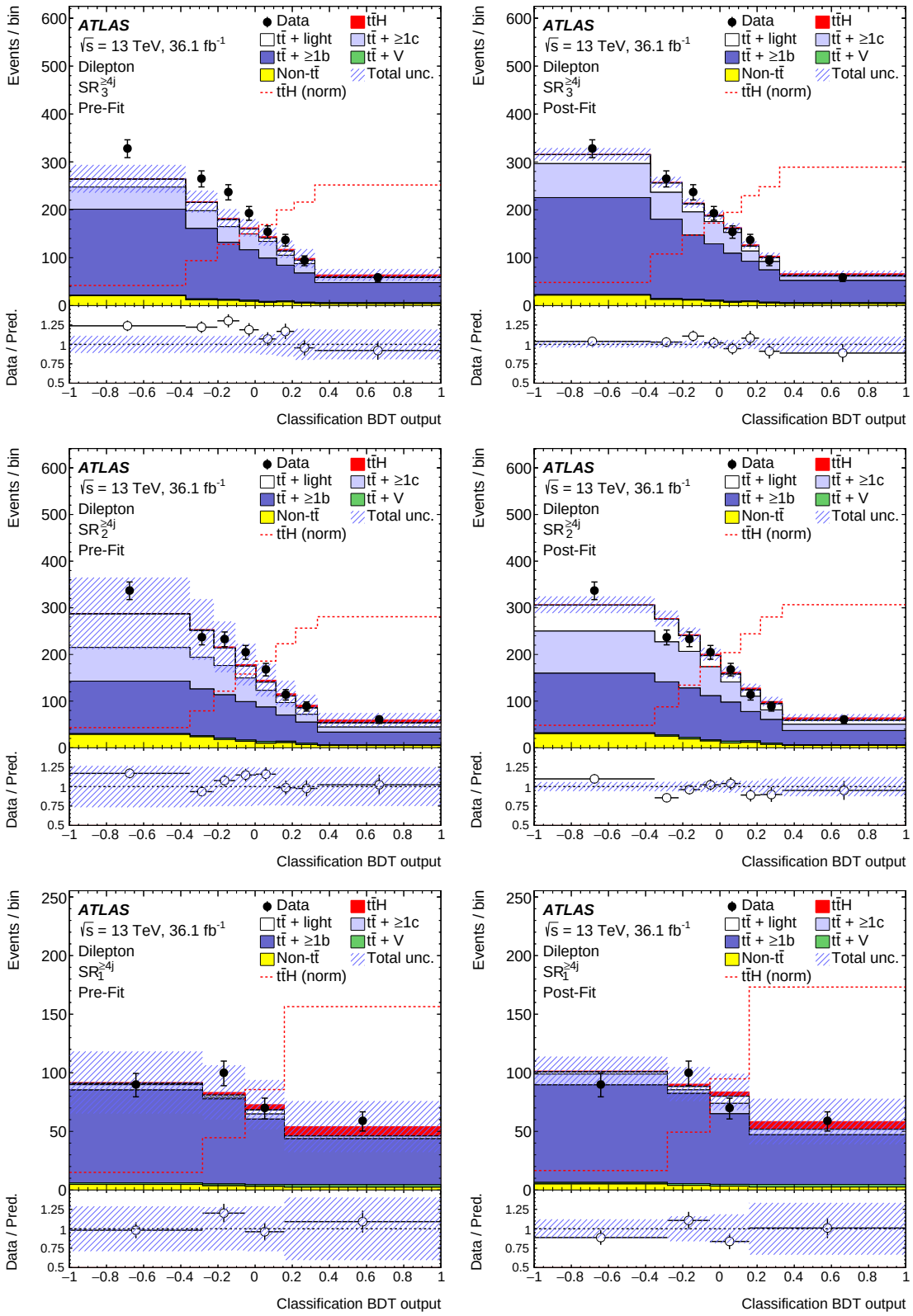


FIGURE 6.38: BDT output distributions in signal regions for the final fit to data under the signal-plus-background hypothesis. In the left (right) column pre-fit (post-fit) distributions are shown. In the top row results for the ≥ 4 jet SR_3 signal region are shown. In the middle row results for the ≥ 4 jet SR_2 signal region are shown. In the bottom row results for the ≥ 4 jet SR_1 signal region are shown. The dashed red line shows the $t\bar{t}H$ distribution normalised to the total background prediction, to allow for the shape of the $t\bar{t}H$ distribution to be clearly seen. [90]

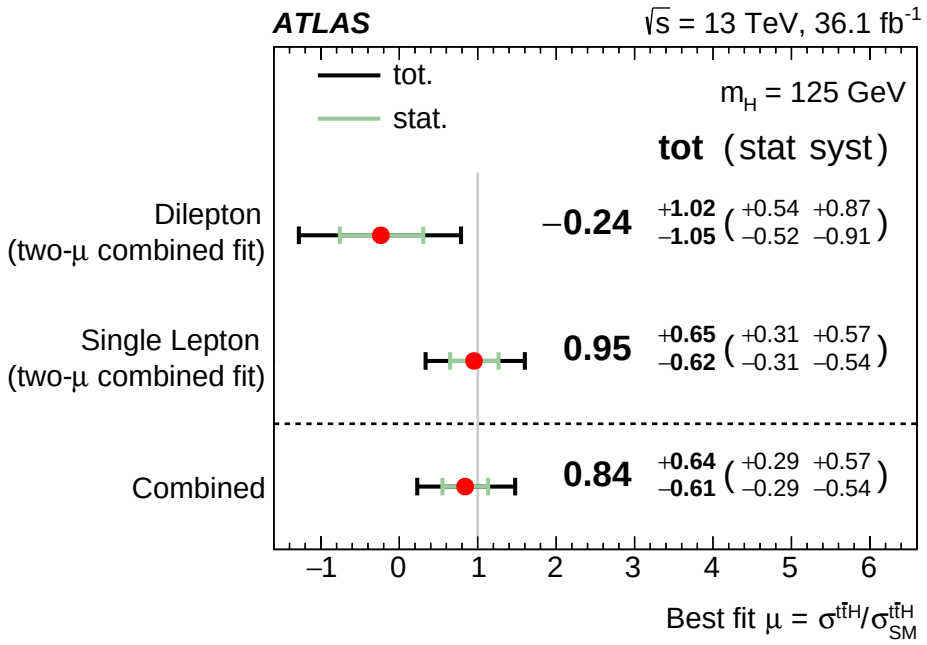


FIGURE 6.39: Measured $t\bar{t}H$ signal strengths from the final fit to data under the signal-plus-background hypothesis. The combined signal strength is measured in the combined fit which uses a single $t\bar{t}H$ signal strength parameter while the individual channel signal strengths are from the same fit except with individual $t\bar{t}H$ signal strength parameters used in each channel. [90]

the gaussian mean. This can be calculated for the fit to the Asimov dataset to provide an expected discovery significance assuming the presence of $t\bar{t}H$ at the Standard Model expectation and for the fit to data to provide the measured discovery significance. By performing this calculation using the combined measurement of the $t\bar{t}H$ signal strength an observed (expected) significance of 1.4σ (1.6σ) is measured.

Upper limits on the $t\bar{t}H$ signal strength are calculated using the CL_S method [152–154]. Exclusion limits calculated using this method for the background-only hypothesis, the presence of $t\bar{t}H$ at the Standard Model expectation, and for the measured $t\bar{t}H$ signal strength are shown in Figure 6.40. A signal strength larger than 2.0 is excluded at the 95% confidence limit from the combined fit. All discovery significance and limit calculations are performed using the posterior background predictions.

The results from the ATLAS $t\bar{t}H(b\bar{b})$ analysis were then combined with results from ATLAS $t\bar{t}H$ searches looking for other decays of the Higgs boson. This combination is performed by forming a combined binned profile likelihood of analysis regions from all analyses in the same way that the dilepton and single-lepton $t\bar{t}H(b\bar{b})$ analyses were combined. The $t\bar{t}H(b\bar{b})$ analysis is combined with analyses searching for $H \to \gamma\gamma$ [155],

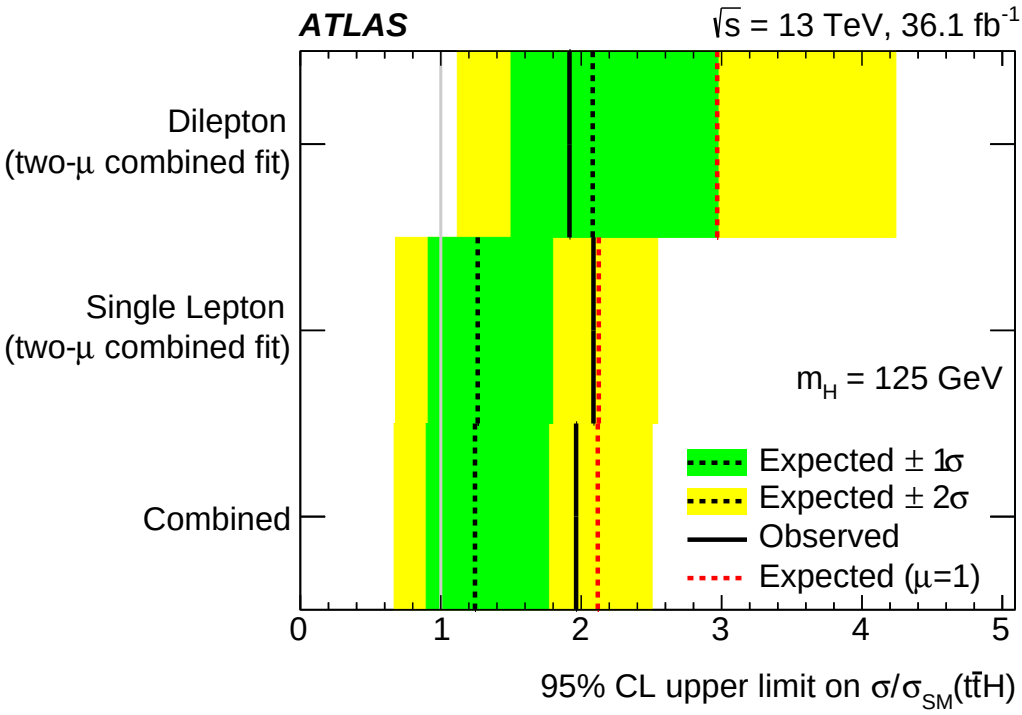


FIGURE 6.40: 95% confidence limit upper limits on the $t\bar{t}H$ signal strength under a number of hypotheses. The black dashed lines show the upper limits for the background-only hypothesis. The dotted red lines show the upper limits for the presence of $t\bar{t}H$ signal at the Standard Model prediction. The solid black lines show the upper limits for the measured value of $t\bar{t}H$ signal strength. The combined signal strength is measured in the combined fit which uses a single $t\bar{t}H$ signal strength parameter while the individual channel signal strengths are from the same fit except with individual $t\bar{t}H$ signal strength parameters used in each channel. [90]

$H \to ZZ^* \to 4l$ [156], and $H \to$ multilepton [157], referring to a range of multilepton final states arising from Higgs boson decays to WW^*, $\tau\tau$, and ZZ^*. A combination with these analyses performed with the same dataset as the $t\bar{t}H(b\bar{b})$ analysis resulted in a measured $t\bar{t}H$ signal strength of $\mu = 1.17^{+0.33}_{-0.30}$ [157]. This corresponds to a measured (expected) discovery significance of 4.2σ (3.8σ) and as such constitutes evidence of the $t\bar{t}H$ production mechanism. A later combination of the $t\bar{t}H$ searches was performed which used the 36.1 fb^{-1} $t\bar{t}H(b\bar{b})$ analysis presented in this thesis as well as the $H \to$ multilepton analysis with the same dataset with $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ analyses which used a larger 79.8 fb^{-1} dataset [158]. From this combination a measured $t\bar{t}H$ signal strength of $\mu = 1.32^{+0.28}_{-0.26}$ was observed. This corresponds to a measured (expected) discovery significance of 5.8σ (4.9σ) and constitutes observation of the $t\bar{t}H$ production mechanism. The $t\bar{t}H$ production cross-section was measured to be 670^{+142}_{-135} fb at a centre-of-mass energy of 13 TeV. This is in agreement with the Standard Model prediction of 507^{+35}_{-50} fb. The measured $t\bar{t}H$ signal strengths in each combination are shown in Figure 6.41.

Higgs coupling measurements have also been produced for the inclusion of the ATLAS $t\bar{t}H$ measurements with other ATLAS Higgs results [159]. From this combined result the Yukawa top coupling is measured to be $1.03^{+0.12}_{-0.11}$ relative to the Standard Model prediction, assuming no Beyond the Standard Model (BSM) particles contribute to loops or particle decays. If no assumptions on loop content are made and it is assumed that no BSM particles contribute to decays, the Yukawa top coupling is measured as $1.09^{+0.15}_{-0.14}$ relative to the Standard Model prediction. If no assumptions on loop content are made and it is assumed that BSM particles can contribute to decays, the Yukawa top coupling is measured as $1.05^{+0.14}_{-0.13}$ relative to the Standard Model prediction.

The CMS Collaboration also performed searches for $t\bar{t}H$ production and the $t\bar{t}H(b\bar{b})$ channel using $\sqrt{s} = 13$ TeV LHC p-p collision data. The search for $t\bar{t}H$ production in the $t\bar{t}H(b\bar{b})$ channel was performed using 35.9 fb^{-1} of data and measured a $t\bar{t}H$ signal strength of $\mu = 0.72 \pm 0.45$, which corresponded to an observed (expected) significance of 1.6σ (2.2σ) [160]. A 95% confidence limit was set: a signal strength larger than 1.5 was excluded. The search for $t\bar{t}H$ production using the combination of a number of $t\bar{t}H$ analyses also used 35.9 fb^{-1} of $\sqrt{s} = 13$ TeV data and measured a $t\bar{t}H$ signal strength of $\mu = 1.14^{+0.31}_{-0.27}$ [161]. When the 13 TeV data was combined with 5.1 fb^{-1} and

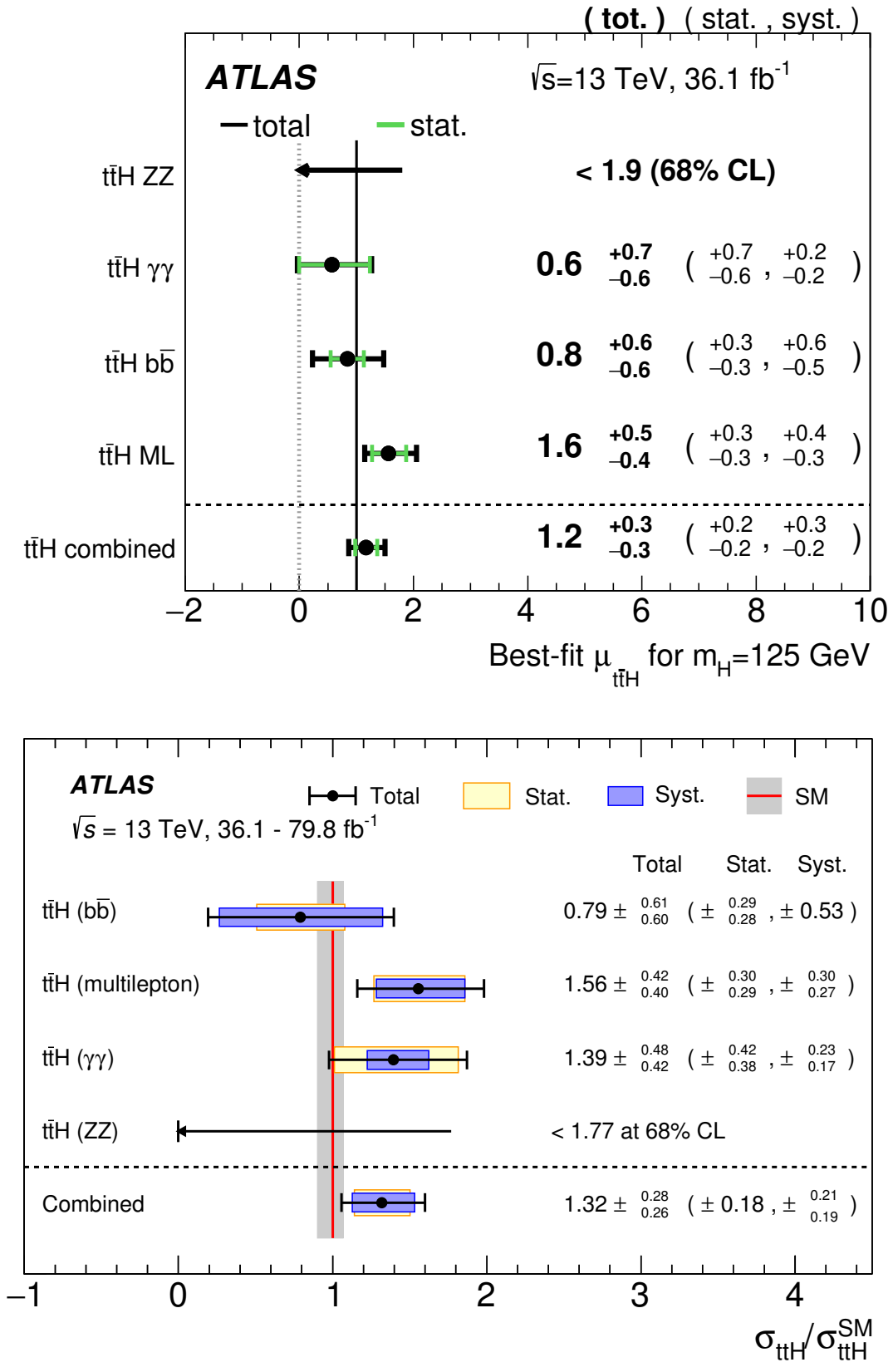


FIGURE 6.41: Measured $t\bar{t}H$ signal strengths for the combination of a number of ATLAS $t\bar{t}H$ searches. In the top plot the results for the combination of analyses using 36.1 fb^{-1} of data are shown. In the bottom plot the results for the combination of $t\bar{t}H(b\bar{b})$ and $t\bar{t}H$(multilepton) searches using 36.1 fb^{-1} with $t\bar{t}H(\gamma\gamma)$ and $t\bar{t}H(ZZ^*)$ searches using 79.8 fb^{-1} of data are shown. The combined signal strength is measured in the combined fit which uses a single $t\bar{t}H$ signal strength parameter while the individual search signal strengths are from the results of the individual searches. [157, 158]

19.7 fb^{-1} of 7 TeV and 8 TeV data respectively, the $t\bar{t}H$ signal strength was measured to be $\mu = 1.26^{+0.31}_{-0.26}$. This corresponds to a significance of 5.2σ and therefore constitutes an observation of the $t\bar{t}H$ production mechanism that is independent of the results from ATLAS. The ATLAS and CMS measurements agree within uncertainties, and both are in agreement with the Standard Model prediction within approximately 1σ uncertainties.

The CMS collaboration also produced Higgs coupling measurements for the inclusion of the CMS $t\bar{t}H$ results with other CMS Higgs results [162]. For the assumption of no BSM particle contributions to loops or decays the Yukawa top coupling is measured to be $1.11^{+0.12}_{-0.10}$ relative to the Standard Model prediction. If no assumptions of loop content are made and it is assumed that no BSM particles contribute to decays, the Yukawa top coupling is measured as $0.98^{+0.14}_{-0.14}$ relative to the Standard Model prediction. If no assumptions of loop content are made and it is assumed that BSM particles can contribute to decays, the Yukawa top coupling is measured as $1.02^{+0.19}_{-0.15}$ relative to the Standard Model prediction. The coupling results from ATLAS and CMS are in agreement within uncertainties, and are also in agreement with the Standard Model prediction within uncertainties.

Chapter 7

Conclusions

The search for the $t\bar{t}H$ production mechanism in the channel in which the Higgs boson decays to bottom quarks has been presented, with focus on the dileptonic channel. The author's work in a number of aspects of this analysis were shown:

- a study of the effect of including dilepton triggers in the analysis and subsequently lowering lepton p_{T} thresholds was performed, and no significant gain in signal purity or signal sensitivity was seen from doing this;
- the estimation of yields and distribution shapes for events containing fake or non-prompt leptons was performed using a simulation-based method which was validated in an analysis region with the opposite-sign lepton charge requirement inverted;
- shape and acceptance uncertainties arising from PDF choice were estimated and it was seen that these uncertainties are approximately 1% or less for both the signal $t\bar{t}H$ process and the $t\bar{t}$ background process;
- detailed studies and development of the statistical fit model have been presented, showing the process of refinement that results in stable and understood fit results.

Results were presented for 36.1 fb^{-1} of data collected by the ATLAS detector in 2015 and 2016. The $t\bar{t}H$ signal strength from the combined single-lepton channel and dilepton channel fit has been measured as $\mu = 0.84^{+0.64}_{-0.61}$. An upper limit of 2.0 has been set for the $t\bar{t}H$ signal strength at the 95% confidence limit. The combination of the $t\bar{t}H(b\bar{b})$ analysis with other searches for $t\bar{t}H$ which target Higgs boson decays to other particles measured a $t\bar{t}H$ signal strength of $\mu = 1.32^{+0.28}_{-0.28}$. This constitutes observation of the $t\bar{t}H$ production

mechanism. Results have been compared with results from the CMS Collaboration and good agreement was seen.

Developments by the author to inner detector trigger algorithms used in tau triggers have also been presented. Improvements were made to the selection of the track used to seed the position of second-stage RoIs, resulting in an improvement in the efficiency of correct RoI placement of approximately 2%. Reductions in the track finding efficiency for low track p_{T} in first-stage RoIs were studied and were understood to originate from tracks which were bent out of the RoIs by the solenoidal magnetic field. Loss of these tracks did not significantly affect the efficiency of correct placement of second-stage RoIs and therefore no changes were made to the inner detector algorithms. Parametrised track-finding methods were developed around measurements made of z resolutions for track seed-doublets. These methods reduced computation time for track finding by up to 50% for reductions in track finding efficiency for tracks from taus of approximately 1% to 2%.

7.1 Outlook

The observation of the $t\bar{t}H$ production mechanism at a rate in agreement with the Standard Model prediction further cements the consistency of the observed scalar boson with the Standard Model Higgs boson. Now that the mechanism has been observed, focus can be placed on the precise measurement of the $t\bar{t}H$ cross-section, and therefore on the Yukawa top coupling. As described in Section 2.3, the precise value of the Yukawa top coupling can have profound implications for the fate of the Higgs vacuum state. Furthermore, although the measured cross-section is in agreement with the Standard Model expectation at the level of precision needed for observation, the size of the uncertainties does not exclude small deviations from the Standard Model prediction. If the $t\bar{t}H$ cross-section, and therefore the Yukawa coupling, is different from the Standard Model prediction this could be an indication of new physics and may indicate the expected scale of new physics.

To estimate the expected precision of future $t\bar{t}H$ results, particularly in regards to the upgrade of the LHC to the High Luminosity LHC (HL-LHC) which will significantly increase the collision luminosity and will increase the collision energy to $\sqrt{s} = 14$ TeV [163],

an extrapolation was performed based on $\sqrt{s} = 13$ TeV ATLAS Higgs boson analyses [164]. Results are produced for two scenarios, where each extrapolates current analyses to an integrated luminosity of 3000 fb^{-1} that is expected from the HL-LHC: a scenario where systematic uncertainties are reduced compared to present uncertainties, including the halving of all theoretical uncertainties, labelled the S2 scenario; and a scenario where systematic uncertainties are kept at their current values, labelled the S1 scenario. This is done to highlight the importance of reducing systematic uncertainties in future results. Uncertainties associated with the statistical size of simulated datasets are neglected in both scenarios. For scenario S1 the expected uncertainty on the $t\bar{t}H$ signal strength for the $t\bar{t}H(b\bar{b})$ single-lepton channel is $^{+0.25}_{-0.20}$, while for the dilepton channel the expected uncertainty is $^{+0.32}_{-0.26}$. For scenario S2 the expected single-lepton channel uncertainty is $^{+0.18}_{-0.15}$, and the expected dilepton channel uncertainty is $^{+0.23}_{-0.20}$.

When using these extrapolated $t\bar{t}H(b\bar{b})$ results in combination with other similarly improved Higgs boson analyses, the expected uncertainties for a measurement of the $t\bar{t}H$ production cross-section relative to the Standard Model cross-section are $^{+0.054}_{-0.052}$ ($^{+0.069}_{-0.066}$) for the S2 (S1) scenario. Uncertainties are also estimated from this combination for the measurement of the $t\bar{t}H(b\bar{b})$ cross-section times branching ratio relative to the Standard Model: an uncertainty of $^{+0.151}_{-0.133}$ ($^{+0.218}_{-0.181}$) is estimated for the S2 (S1) scenario, placing an observation of specifically the $t\bar{t}H(b\bar{b})$ process within grasp at the HL-LHC. Uncertainties for the measurement of the Yukawa top coupling from this combination are also estimated. Assuming no BSM contributions to loops or decays an uncertainty of $^{+0.030}_{-0.029}$ ($^{+0.043}_{-0.041}$) is estimated in scenario S2 (S1) for the measurement of the Yukawa top coupling relative to the Standard Model expectation. If no assumptions are made about loop content and no BSM contributions to decays are assumed, an uncertainty of $^{+0.043}_{-0.041}$ ($^{+0.068}_{-0.058}$) is estimated for the S2 (S1) scenario. If no assumptions are made about loop content and BSM contributions to decays are allowed, an uncertainty of $^{+0.039}_{-0.040}$ ($^{+0.063}_{-0.058}$) is estimated for the S2 (S1) scenario.

The outlook for future $t\bar{t}H$ measurements is therefore very bright. Even without significant reductions in systematic uncertainties, results from the HL-LHC will be sensitive to smaller deviations from the Standard Model. In the near future, the ATLAS $t\bar{t}H(b\bar{b})$ analysis will be updated to produce results with the full Run 2 dataset, providing further precision for the impressive suite of Run 2 $t\bar{t}H$ measurements.

Appendix A

Additional Results From the Statistical Analysis Using a Binned Profile Likelihood Fit

A.1 Comparison Of First Full Fit Models

A comparison of the "first full fit model" initially used in Section 6.7.1 (referred to here as the 'initial' model) with the updated version of the same fit model (referred to here as the 'updated' model) are shown for the $k(t\bar{t}+\geq 1b)$ and $k(t\bar{t}+\geq 1c)$ normalisation factors in Figure A.1 and for the nuisance parameters in Figure A.2. Results are shown for fits to data in a background-only fit model which uses $H_{\mathrm{T}}^{\mathrm{all}}$ binned distributions in all regions. The main difference between the two fit models is the use of updated calibrations of the c-jet mis-tag efficiencies for the MV2c10 b-tagging algorithm in the latter fit model.

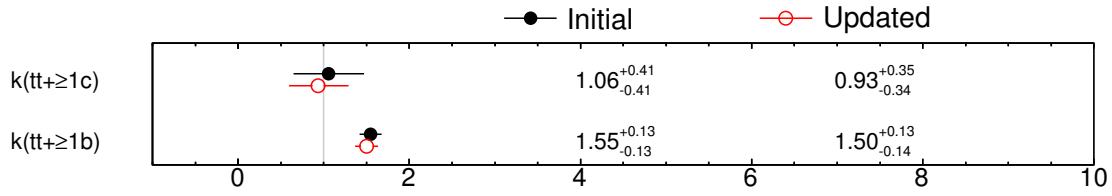


FIGURE A.1: Fitted normalisation factors compared for the initial and updated versions of the first full fit model for fits to data. The fits are performed under the background-only hypothesis and uses the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution used in all regions.

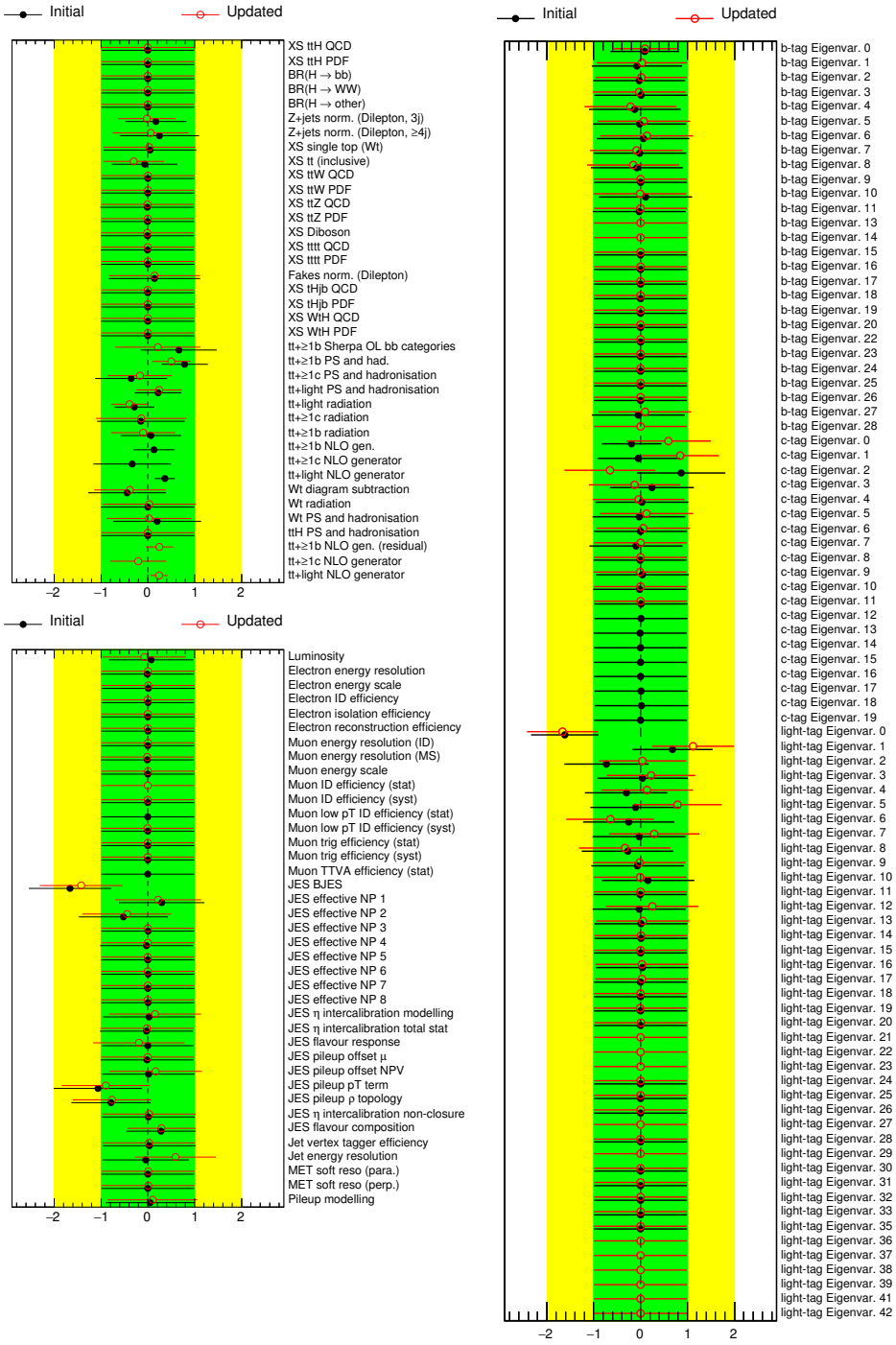


FIGURE A.2: Fitted nuisance parameters compared for the the initial and updated versions of the first full fit model for fits to data. The fits are performed under the background-only hypothesis and use the $H_{\mathrm{T}}^{\mathrm{all}}$ distribution in all regions. In the top left plot theoretical nuisance parameters are shown, in the bottom left plot non-*b*-tagging intrumental nuisance parameters are shown, and in the right plot *b*-tagging nuisance parameters are shown. Differences in the fit models result in some nuisance parameters being pruned or included in only one of the fit models. "NLO generator" uncertainties are the same between both fit models, but are plotted as separate parameters due to changes in internal parameter names within the fitting software.

Bibliography

[1] L. Evans and P. Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

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

[3] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett. B* 716 (2012), p. 30. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[4] S. L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579 –588. ISSN: 0029-5582. DOI: 10.1016/0029-5582(61)90469-2. URL: http://www.sciencedirect.com/science/article/pii/0029558261904692.

[5] S. Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[6] A. Salam. "Weak and Electromagnetic Interactions". In: *Conf. Proc.* C680519 (1968), pp. 367–377.

[7] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Phys. Rev. D* 2 (7 1970), pp. 1285–1292. DOI: 10.1103/PhysRevD.2.1285. URL: https://link.aps.org/doi/10.1103/PhysRevD.2.1285.

[8] ATLAS Collaboration 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". In: *Phys. Rev. Lett.* 114 (2015), p. 191803. DOI: 10.1103/PhysRevLett.114.191803. arXiv: 1503.07589 [hep-ex].

[9] ATLAS Collaboration. "Measurement of the Higgs boson mass in the $H \to ZZ^* \to 4\ell$ and $H \to \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector". In: *Physics Letters B* 784 (2018), pp. 345–366. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2018.07.050. URL: http://www.sciencedirect.com/science/article/pii/S0370269318305884.

[10] CMS collaboration. "Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV". In: *Journal of High Energy Physics* 2017.11 (2017), p. 47. ISSN: 1029-8479. DOI: 10.1007/JHEP11(2017)047.

[11] ATLAS Collaboration. "Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector". In: *Eur. Phys. J.* C75.10 (2015), p. 476. DOI: 10.1140/epjc/s10052-015-3685-1. arXiv: 1506.05669 [hep-ex]. Erratum. ATLAS Collaboration [165].

[12] ATLAS Collaboration. "Test of CP invariance in vector-boson fusion production of the Higgs boson using the Optimal Observable method in the ditau decay channel with the ATLAS detector". In: *The European Physical Journal C* 76.12 (2016), p. 658. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4499-5.

[13] CMS Collaboration. "Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV". In: *Phys. Rev.* D92.1 (2015), p. 012004. DOI: 10.1103/PhysRevD.92.012004. arXiv: 1411.3441 [hep-ex].

[14] CMS Collaboration. "Combined search for anomalous pseudoscalar HVV couplings in $VH(H \to b\bar{b})$ production and $H \to VV$ decay". In: *Physics Letters B* 759 (2016), pp. 672–696. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2016.06.004. URL: http://www.sciencedirect.com/science/article/pii/S0370269316302556.

[15] CMS Collaboration. "Constraints on anomalous Higgs boson couplings using production and decay information in the four-lepton final state". In: *Physics Letters B* 775 (2017), pp. 1–24. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2017.10.021. URL: http://www.sciencedirect.com/science/article/pii/S037026931730833X.

[16] ATLAS Collaboration 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". In: *JHEP* 08 (2016), p. 045. DOI: 10.1007/JHEP08(2016)045. arXiv: 1606.02266 [hep-ex].

[17] G. 't Hooft and M. Veltman. "Regularization and renormalization of gauge fields". In: *Nuclear Physics B* 44.1 (1972), pp. 189 –213. ISSN: 0550-3213. DOI: 10.1016/0550-3213(72)90279-9. URL: http://www.sciencedirect.com/science/article/pii/0550321372902799.

[18] M. Tanabashi et al. (Particle Data Group). "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[19] N. Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (12 1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[20] M. Kobayashi and T. Maskawa. "CP Violation in the Renormalizable Theory of Weak Interaction". In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.

[21] Super-Kamiokande Collaboration. "Measurements of the solar neutrino flux from Super-Kamiokande's first 300 days". In: *Phys. Rev. Lett.* 81 (6 1998), pp. 1158–1162. DOI: 10.1103/PhysRevLett.81.1158. arXiv: hep-ex/9805021 [hep-ex]. Erratum. Super-Kamiokande Collaboration [166].

[22] Peter Minkowski. "$\mu \to e\gamma$ at a Rate of One Out of 10^9 Muon Decays?" In: *Phys. Lett.* 67B (1977), pp. 421–428. DOI: 10.1016/0370-2693(77)90435-X.

[23] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[24] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[25] P. W. Higgs. "Spontaneous Symmetry Breakdown without Massless Bosons". In: *Phys. Rev.* 145 (4 1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156. URL: https://link.aps.org/doi/10.1103/PhysRev.145.1156.

[26] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. "Global Conservation Laws and Massless Particles". In: *Phys. Rev. Lett.* 13 (20 1964), pp. 585–587. DOI: 10.1103/PhysRevLett.13.585. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[27] J. Goldstone, A. Salam, and S. Weinberg. "Broken Symmetries". In: *Phys. Rev.* 127 (3 1962), pp. 965–970. DOI: 10.1103/PhysRev.127.965. URL: https://link.aps.org/doi/10.1103/PhysRev.127.965.

[28] F. Bezrukov and M. Shaposhnikov. "Why should we care about the top quark Yukawa coupling?" In: *Journal of Experimental and Theoretical Physics* 120.3 (2015), pp. 335–343. ISSN: 1090-6509. DOI: 10.1134/S1063776115030152.

[29] J. R. Espinosa and M. Quiros. "Improved metastability bounds on the standard model Higgs mass". In: *Phys. Lett.* B353 (1995), pp. 257–266. DOI: 10.1016/0370-2693(95)00572-3. arXiv: hep-ph/9504241 [hep-ph].

[30] G. Isidori, G. Ridolfi, and A. Strumia. "On the metastability of the Standard Model vacuum". In: *Nuclear Physics B* 609.3 (2001), pp. 387–409. ISSN: 0550-3213. DOI: 10.1016/S0550-3213(01)00302-9. URL: http://www.sciencedirect.com/science/article/pii/S0550321301003029.

[31] D. de Florian et al. "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector". In: (2016). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph].

[32] ATLAS Collaboration. "Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector". In: *Eur. Phys. J.* C75.7 (2015), p. 349. DOI: 10.1140/epjc/s10052-015-3543-1. arXiv: 1503.05066 [hep-ex].

[33] ATLAS Collaboration. "Search for the associated production of the Higgs boson with a top quark pair in multilepton final states with the ATLAS detector". In: *Phys. Lett.* B749 (2015), pp. 519–541. DOI: 10.1016/j.physletb.2015.07.079. arXiv: 1506.05988 [hep-ex].

[34] ATLAS Collaboration. "Search for $H \to \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector". In: *Phys. Lett.* B740 (2015), pp. 222–242. DOI: 10.1016/j.physletb.2014.11.049. arXiv: 1409.3122 [hep-ex].

[35] ATLAS Collaboration. "Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 05 (2016), p. 160. DOI: 10.1007/JHEP05(2016)160. arXiv: 1604.03812 [hep-ex].

[36] CMS Collaboration. "Search for a Standard Model Higgs Boson Produced in Association with a Top-Quark Pair and Decaying to Bottom Quarks Using a Matrix Element Method". In: *Eur. Phys. J.* C75.6 (2015), p. 251. DOI: 10.1140/epjc/s10052-015-3454-1. arXiv: 1502.02485 [hep-ex].

[37] CMS Collaboration. "Search for the associated production of the Higgs boson with a top-quark pair". In: *JHEP* 09 (2014). [Erratum: JHEP10,106(2014)], p. 087. DOI: 10.1007/JHEP09(2014)087, 10.1007/JHEP10(2014)106. arXiv: 1408.1682 [hep-ex].

[38] Jean-Luc Caron. *The LHC injection complex*. LHC-PHO-1993-008. 1993. URL: https://cds.cern.ch/record/841568.

[39] *ATLAS luminosity public results Run 2*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. Accessed 07/08/2018.

[40] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[41] ATLAS Collaboration. *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV*. Tech. rep. ATL-PHYS-PUB-2015-018. 2015. URL: http://cds.cern.ch/record/2037683.

[42] ATLAS Collaboration. "Study of the material of the ATLAS inner detector for Run 2 of the LHC". In: *JINST* 12.CERN-EP-2017-081 (2017), P12009. 71 p. URL: https://cds.cern.ch/record/2273894.

[43] M. Capeans et al. *ATLAS Insertable B-Layer Technical Design Report*. Tech. rep. CERN-LHCC-2010-013. ATLAS-TDR-19. 2010. URL: https://cds.cern.ch/record/1291633.

[44] ATLAS Collaboration. "Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties". In: *Journal of Instrumentation* 12.05 (2017), P05002–P05002. DOI: 10.1088/1748-0221/12/05/p05002.

[45] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data". In: *Eur. Phys. J.* C74.10 (2014), p. 3071. DOI: 10.1140/epjc/s10052-014-3071-4. arXiv: 1407.5063 [hep-ex].

[46] ATLAS Collaboration. "Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector". In: *Eur. Phys. J.* C73.3 (2013), p. 2306. DOI: 10.1140/epjc/s10052-013-2306-0. arXiv: 1210.6210 [hep-ex].

[47] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur. Phys. J.* C76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[48] *ATLAS approved plots DAQ*.
https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsDAQ.
Accessed 26/07/2018.

[49] J. M. Miguéns on behalf of the ATLAS Collaboration. "The ATLAS Run-2 Trigger: Design, Menu, Performance and Operational Aspects". In: *PoS* ICHEP2016 (2016), p. 244. DOI: 10.22323/1.282.0244.

[50] B. Andersson et al. "Parton Fragmentation and String Dynamics". In: *Phys. Rept.* 97 (1983), pp. 31–145. DOI: 10.1016/0370-1573(83)90080-7.

[51] B. R. Webber. "A QCD Model for Jet Fragmentation Including Soft Gluon Interference". In: *Nucl. Phys.* B238 (1984), pp. 492–528. DOI: 10.1016/0550-3213(84)90333-X.

[52] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[53] E. Richter-Was, D. Froidevaux, and L. Poggioli. *ATLFAST 2.0 a fast simulation package for ATLAS*. Tech. rep. ATL-PHYS-98-131. 1998. URL: https://cds.cern.ch/record/683751.

[54] ATLAS Collaboration. *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*. Tech. rep. ATL-PHYS-PUB-2010-013. 2010. URL: https://cds.cern.ch/record/1300517.

[55] T. Cornelissen et al. "The new ATLAS track reconstruction (NEWT)". In: *J. Phys. Conf. Ser.* 119 (2008), p. 032014. DOI: 10.1088/1742-6596/119/3/032014.

[56] ATLAS Collaboration. "Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2". In: *Eur. Phys. J.* C77.10 (2017), p. 673. DOI: 10.1140/epjc/s10052-017-5225-7. arXiv: 1704.07983 [hep-ex].

[57] R. Fruhwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nucl. Instrum. Meth.* A262 (1987), pp. 444–450. DOI: 10.1016/0168-9002(87)90887-4.

[58] ATLAS Collaboration. "Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC". In: *Eur. Phys. J.* C77.5 (2017), p. 332. DOI: 10.1140/epjc/s10052-017-4887-5. arXiv: 1611.10235 [physics.ins-det].

[59] G. Piacquadio, K. Prokofiev, and A. Wildauer. "Primary vertex reconstruction in the ATLAS experiment at LHC". In: *J. Phys. Conf. Ser.* 119 (2008), p. 032033. DOI: 10.1088/1742-6596/119/3/032033.

[60] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*. Tech. rep. ATLAS-CONF-2016-024. 2016. URL: https://cds.cern.ch/record/2157687.

[61] ATLAS Collaboration. "Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data". In: *Eur. Phys. J.* C77.3 (2017), p. 195. DOI: 10.1140/epjc/s10052-017-4756-2. arXiv: 1612.01456 [hep-ex].

[62] ATLAS Collaboration. *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. ATLAS-CONF-2012-047. 2012. URL: http://cds.cern.ch/record/1449796.

[63] J. Illingworth and J. Kittler. "A Survey of the Hough Transform". In: *Computer Vision, Graphics, and Image Processing* 44.1 (1988), pp. 87 –116. ISSN: 0734-189X. DOI: 10.1016/S0734-189X(88)80033-1. URL: http://www.sciencedirect.com/science/article/pii/S0734189X88800331.

[64] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J.* C77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[65] M. Cacciari, G. P. Salam, and G. Soyez. "The anti-k_t jet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (2008), p. 063. URL: http://stacks.iop.org/1126-6708/2008/i=04/a=063.

[66] S. Catani et al. "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions". In: *Nucl. Phys.* B406 (1993), pp. 187–224. DOI: 10.1016/0550-3213(93)90166-M.

[67] Y. L. Dokshitzer et al. "Better jet clustering algorithms". In: *Journal of High Energy Physics* 1997.08 (1997), p. 001. URL: http://stacks.iop.org/1126-6708/1997/i=08/a=001.

[68] M. Wobisch and T. Wengler. "Hadronization corrections to jet cross-sections in deep inelastic scattering". In: *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*. 1998, pp. 270–279. arXiv: hep-ph/9907280 [hep-ph]. URL: https://inspirehep.net/record/484872/files/arXiv:hep-ph_9907280.pdf.

[69] M. Cacciari, G. P. Salam, and G. Soyez. "FastJet user manual". In: *The European Physical Journal C* 72.3 (2012), p. 1896. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-012-1896-2.

[70] ATLAS Collaboration. "Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. D* 96.7 (2017), p. 072002. DOI: 10.1103/PhysRevD.96.072002. arXiv: 1703.09665 [hep-ex].

[71] ATLAS Collaboration. "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector". In: *Eur. Phys. J.* C76.11 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex].

[72] ATLAS Collaboration. *Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector*. Tech. rep. ATLAS-CONF-2015-029. 2015. URL: http://cds.cern.ch/record/2037702.

[73] ATLAS Collaboration. *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*. Tech. rep. ATL-PHYS-PUB-2016-012. 2016. URL: https://cds.cern.ch/record/2160731.

[74] ATLAS Collaboration. "Performance of b-jet identification in the ATLAS experiment". In: *JINST* 11 (2016), P04008. DOI: 10.1088/1748-0221/11/04/P04008. arXiv: 1512.01094 [hep-ex].

[75] ATLAS Collaboration. *Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data*. Tech. rep. ATLAS-CONF-2010-042. 2010. URL: https://cds.cern.ch/record/1277682.

[76] ATLAS Collaboration. *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*. Tech. rep. ATLAS-CONF-2018-001. 2018. URL: http://cds.cern.ch/record/2306649.

[77] ATLAS Collaboration. *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*. Tech. rep. ATLAS-CONF-2017-029. 2017. URL: http://cds.cern.ch/record/2261772.

[78] ATLAS Collaboration. *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*. Tech. rep. ATL-PHYS-PUB-2015-045. 2015. URL: http://cds.cern.ch/record/2064383.

[79] ATLAS Collaboration. *E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p-p collisions*. Tech. rep. ATLAS-CONF-2018-023. 2018. URL: https://cds.cern.ch/record/2625233.

[80] C. Kilby on behalf of the ATLAS Collaboration. "The design and performance of the ATLAS Inner Detector trigger for Run 2 LHC Collisions at $\sqrt{s} = 13$ TeV". In: *J. Phys. Conf. Ser.* 762.1 (2016), p. 012029. DOI: 10.1088/1742-6596/762/1/012029.

[81] C. Kilby on behalf of the ATLAS collaboration. "The design and performance of the atlas inner detector trigger in high pileup collisions at 13 TeV at the Large Hadron Collider". In: CEUR Workshop Proceedings Vol-2023 (Dec. 2017). Ed. by Vladimir Korenkov and Andrey Nechaevskiy, p. 188.

[82] ATLAS Collaboration. "Performance of the ATLAS Trigger System in 2015". In: *Eur. Phys. J.* C77.5 (2017), p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661 [hep-ex].

[83] ATLAS Collaboration. *The ATLAS Tau Trigger in Run 2*. Tech. rep. ATLAS-CONF-2017-061. 2017. URL: https://cds.cern.ch/record/2274201.

[84] T. Sjöstrand, S. Mrenna, and P. Z. Skands. "A brief introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), p. 852. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[85] P. Nason. "A new method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.

[86] S. Frixione, P. Nason, and C. Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[87] S. Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[88] J. M. Campbell et al. "Top-pair production and decay at NLO matched with parton showers". In: *JHEP* 04 (2015), p. 114. DOI: 10.1007/JHEP04(2015)114. arXiv: 1412.1828 [hep-ph].

[89] ATLAS Collaboration. *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*. ATLAS-CONF-2016-080. 2016. URL: http://cds.cern.ch/record/2206255.

[90] ATLAS Collaboration. "Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} =$ 13 TeV with the ATLAS detector". In: *Phys. Rev. D* 97 (7 Apr. 2018), p. 072016. DOI: 10.1103/PhysRevD.97.072016. URL: https://link.aps.org/doi/10.1103/PhysRevD.97.072016.

[91] Byron P. Roe et al. "Boosted decision trees, an alternative to artificial neural networks". In: *Nucl. Instrum. Meth.* A543.2-3 (2005), pp. 577–584. DOI: 10.1016/j.nima.2004.12.018. arXiv: physics/0408124 [physics].

[92] ATLAS Collaboration. *Interal ATLAS communication: Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $\boldsymbol{b\bar{b}}$ in $\boldsymbol{pp}$ collisions at $\sqrt{\mathbf{s}} = \mathbf{13}$ TeV with the ATLAS detector*. Tech. rep. ATL-COM-PHYS-2016-116. 2016.

[93] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth. A* 462 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.

[94] P. Artoisenet et al. "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations". In: *JHEP* 03 (2013), p. 015. DOI: 10.1007/JHEP03(2013)015. arXiv: 1212.3460 [hep-ph].

[95] R. Raitio and W. W. Wada. "Higgs-boson production at large transverse momentum in quantum chromodynamics". In: *Phys. Rev. D* 19 (1979), p. 941. DOI: 10.1103/PhysRevD.19.941.

[96] W. Beenakker et al. "NLO QCD corrections to $t\bar{t}H$ production in hadron collisions". In: *Nucl. Phys. B* 653 (2003), pp. 151–203. DOI: 10.1016/S0550-3213(03)00044-0. arXiv: hep-ph/0211352.

[97] S. Dawson et al. "Associated Higgs boson production with top quarks at the CERN Large Hadron Collider: NLO QCD corrections". In: *Phys. Rev. D* 68 (2003), p. 034022. DOI: 10.1103/PhysRevD.68.034022. arXiv: hep-ph/0305087.

[98] Y. Zhang et al. "QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider". In: *Phys. Lett. B* 738 (2014), pp. 1–5. DOI: 10.1016/j.physletb.2014.09.022. arXiv: 1407.1110 [hep-ph].

[99] S. Frixione et al. "Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons". In: *JHEP* 06 (2015), p. 184. DOI: 10.1007/JHEP06(2015)184. arXiv: 1504.03446 [hep-ph].

[100] A. Djouadi, J. Kalinowski, and M. Spira. "HDECAY: a program for Higgs boson decays in the Standard Model and its supersymmetric extension". In: *Comput. Phys. Commun.* 108 (1998), pp. 56–74. DOI: 10.1016/S0010-4655(97)00123-9. arXiv: hep-ph/9704448.

[101] ATLAS Collaboration. *Studies on top-quark Monte Carlo modelling for Top2016*. ATL-PHYS-PUB-2016-020. 2016. URL: https://cds.cern.ch/record/2216168.

[102] M. Czakon and A. Mitov. "Top++: a program for the calculation of the top-pair cross-section at hadron colliders". In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph].

[103] 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". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[104] R. D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[105] ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data*. ATL-PHYS-PUB-2014-021. 2014. URL: https://cds.cern.ch/record/1966419.

[106] M. Cacciari et al. "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation". In: *Phys. Lett. B* 710 (2012), p. 612. DOI: 10.1016/j.physletb.2012.03.013. arXiv: 1111.5869 [hep-ph].

[107] M. Czakon and A. Mitov. "NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels". In: *JHEP* 12 (2012), p. 054. DOI: 10.1007/JHEP12(2012)054. arXiv: 1207.0236 [hep-ph].

[108] M. Czakon and A. Mitov. "NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction". In: *JHEP* 01 (2013), p. 080. DOI: 10.1007/JHEP01(2013)080. arXiv: 1210.6832 [hep-ph].

[109] M. Czakon, P. Fiedler, and A. Mitov. "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$". In: *Phys. Rev. Lett.* 110 (2013), p. 252004. DOI: 10.1103/PhysRevLett.110.252004. arXiv: 1303.6254 [hep-ph].

[110] M. Guzzi et al. "CT10 parton distributions and other developments in the global QCD analysis". In: (2011). arXiv: 1101.0561 [hep-ph].

[111] Jun Gao et al. "CT10 next-to-next-to-leading order global analysis of QCD". In: *Phys. Rev. D* 89 (2014), p. 033009. DOI: 10.1103/PhysRevD.89.033009. arXiv: 1302.6246 [hep-ph].

[112] T. Sjöstrand, S. Mrenna, and P. Z. Skands. "PYTHIA 6.4 physics and manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175.

[113] P. Z. Skands. "Tuning Monte Carlo generators: The Perugia tunes". In: *Phys. Rev. D* 82 (2010), p. 074018. DOI: 10.1103/PhysRevD.82.074018. arXiv: 1005.3457 [hep-ph].

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

[115] ATLAS Collaboration. *Measurement of $W^{\pm}$ and Z Boson Production Cross Sections in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*. ATLAS-CONF-2015-039. 2015. URL: https://cds.cern.ch/record/2045487.

[116] S. Schumann and F. Krauss. "A Parton shower algorithm based on Catani-Seymour dipole factorisation". In: *JHEP* 03 (2008), p. 038. DOI: 10.1088/1126-6708/2008/03/038. arXiv: 0709.1027 [hep-ph].

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

[118] N. Kidonakis. "NNLL resummation for s-channel single top quark production". In: *Phys. Rev. D* 81 (2010), p. 054028. DOI: 10.1103/PhysRevD.81.054028. arXiv: 1001.5034 [hep-ph].

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

[120] ATLAS Collaboration. *Multi-boson simulation for* 13 *TeV ATLAS analyses*. ATL-PHYS-PUB-2016-002. 2016. URL: https://cds.cern.ch/record/2119986.

[121] P. M. Nadolsky et al. "Implications of CTEQ global analysis for collider observables". In: *Phys. Rev. D* 78 (2008), p. 013004. DOI: 10.1103/PhysRevD.78.013004. arXiv: 0802.0007 [hep-ph].

[122] M. Bahr et al. "Herwig++ physics and manual". In: *Eur. Phys. J. C* 58 (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[123] F. Cascioli et al. "NLO matching for $t\bar{t}b\bar{b}$ production with massive b-quarks". In: *Phys. Lett. B* 734 (2014), p. 210. DOI: 10.1016/j.physletb.2014.05.040. arXiv: 1309.5912 [hep-ph].

[124] F. Cascioli, P. Maierhofer, and S. Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph].

[125] S. Frixione et al. "Single-top hadroproduction in association with a W boson". In: *JHEP* 07 (2008), p. 029. DOI: 10.1088/1126-6708/2008/07/029. arXiv: 0805.3067 [hep-ph].

[126] T. Gleisberg and S. Höche. "Comix, a new matrix element generator". In: *JHEP* 12 (2008), p. 039. DOI: 10.1088/1126-6708/2008/12/039. arXiv: 0808.3674 [hep-ph].

[127] S. Höche et al. "QCD matrix elements + parton showers: The NLO case". In: *JHEP* 04 (2013), p. 027. DOI: 10.1007/JHEP04(2013)027. arXiv: 1207.5030 [hep-ph].

[128] ATLAS Collaboration. *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at $\sqrt{s}$ =8 TeV with the ATLAS detector*. ATLAS-CONF-2014-058. 2014. URL: https://cdsweb.cern.ch/record/1951336.

[129] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J.* C76.12 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].

[130] ATLAS Collaboration. "Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC". In: *Phys. Rev. Lett.* 117 (18 2016), p. 182002. DOI: 10.1103/PhysRevLett.117.182002. URL: https://link.aps.org/doi/10.1103/PhysRevLett.117.182002.

[131] T. Ježo et al. "New NLOPS predictions for $t\bar{t}$ + b-jet production at the LHC". In: *Eur. Phys. J.* C78.6 (2018), p. 502. DOI: 10.1140/epjc/s10052-018-5956-0. arXiv: 1802.00426 [hep-ph].

[132] J. Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[133] ATLAS Collaboration. *Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@NLO*. ATL-PHYS-PUB-2017-007. 2017. URL: https://cds.cern.ch/record/2261938.

[134] ATLAS Collaboration. *Studies of $t\bar{t}$ + $c\bar{c}$ production with MadGraph5_aMC@NLO and Herwig++ for the ATLAS experiment*. ATL-PHYS-PUB-2016-011. 2016. URL: https://cds.cern.ch/record/2153876.

[135] A. D. Martin et al. "Parton distributions for the LHC". In: *Eur. Phys. J. C* 63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].

[136] Richard D. Ball et al. "Parton distributions with LHC data". In: *Nucl. Phys.* B867 (2013), pp. 244–289. DOI: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph].

[137] John M. Campbell and R. Keith Ellis. "$t\bar{t}W^{+-}$ production and decay at NLO". In: *JHEP* 07 (2012), p. 052. DOI: 10.1007/JHEP07(2012)052. arXiv: 1204.5678 [hep-ph].

[138] V. N. Gribov and L. N. Lipatov. "Deep inelastic e p scattering in perturbation theory". In: *Sov. J. Nucl. Phys.* 15 (1972). [Yad. Fiz.15,781(1972)], pp. 438–450.

[139] G. Altarelli and G. Parisi. "Asymptotic Freedom in Parton Language". In: *Nucl. Phys.* B126 (1977), pp. 298–318. DOI: 10.1016/0550-3213(77)90384-4.

[140] Y. L. Dokshitzer. "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics." In: *Sov. Phys. JETP* 46 (1977). [Zh. Eksp. Teor. Fiz.73,1216(1977)], pp. 641–653.

[141] J. Butterworth et al. "PDF4LHC recommendations for LHC Run II". In: *J. Phys.* G43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].

[142] S. Dulat et al. "New parton distribution functions from a global analysis of quantum chromodynamics". In: *Phys. Rev.* D93.3 (2016), p. 033006. DOI: 10.1103/PhysRevD.93.033006. arXiv: 1506.07443 [hep-ph].

[143] L. A. Harland-Lang et al. "Parton distributions in the LHC era: MMHT 2014 PDFs". In: *Eur. Phys. J.* C75.5 (2015), p. 204. DOI: 10.1140/epjc/s10052-015-3397-6. arXiv: 1412.3989 [hep-ph].

[144] J. Gao and P. Nadolsky. "A meta-analysis of parton distribution functions". In: *JHEP* 07 (2014), p. 035. DOI: 10.1007/JHEP07(2014)035. arXiv: 1401.0013 [hep-ph].

[145] A. Buckley et al. "LHAPDF6: parton density access in the LHC precision era". In: *Eur. Phys. J.* C75 (2015), p. 132. DOI: 10.1140/epjc/s10052-015-3318-8. arXiv: 1412.7420 [hep-ph].

[146] G. Cowan. *Statistical Data Analysis*. Oxford University Press, 1998. ISBN: 9780198501558.

[147] F. James and M. Roos. "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations". In: *Comput. Phys. Commun.* 10 (1975), pp. 343–367. DOI: 10.1016/0010-4655(75)90039-9.

[148] *Minuit2 Users Guide*. https://root.cern.ch/root/htmldoc/guides/minuit2/Minuit2.html. Accessed 21/08/2018.

[149] *Interpretation of the errors on parameters as given by MINUIT*. http://seal.web.cern.ch/seal/documents/minuit/mnerror.pdf. Accessed 21/08/2018.

[150] K. Cranmer et al. "HistFactory: A tool for creating statistical models for use with RooFit and RooStats". In: (2012).

[151] K. Cranmer. "Practical Statistics for the LHC". In: *Proceedings, 7th CERN - Latin-American School of High-Energy Physics (CLASHEP2013): Arequipa, Peru, March 6-19, 2013*. 2015, pp. 247–289. DOI: 10.5170/CERN-2015-001.247. arXiv: 1503.07622 [physics.data-an].

[152] G. Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J.* C71 (2011), p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an]. Erratum. G. Cowan et al. [167].

[153] A. L. Read. "Presentation of search results: the CL_s technique". In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (2002), p. 2693. URL: http://stacks.iop.org/0954-3899/28/i=10/a=313.

[154] T. Junk. "Confidence level computation for combining searches with small statistics". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 434.2 (1999), pp. 435–443. ISSN: 0168-9002. DOI: 10.1016/S0168-9002(99)00498-2. URL: http://www.sciencedirect.com/science/article/pii/S0168900299004982.

[155] ATLAS Collaboration. "Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev.* D98 (2018), p. 052005. DOI: 10.1103/PhysRevD.98.052005. arXiv: 1802.04146 [hep-ex].

[156] ATLAS Collaboration. "Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4\ell$ decay channel at $\sqrt{s}$ = 13 TeV with the ATLAS detector". In: *JHEP* 03 (2018), p. 095. DOI: 10.1007/JHEP03(2018)095. arXiv: 1712.02304 [hep-ex].

[157] ATLAS Collaboration. "Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector". In: *Phys. Rev.* D97.7 (2018), p. 072003. DOI: 10.1103/PhysRevD.97.072003. arXiv: 1712.08891 [hep-ex].

[158] ATLAS Collaboration. "Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector". In: *Phys. Lett.* B784 (2018), pp. 173–191. DOI: 10.1016/j.physletb.2018.07.035. arXiv: 1806.00425 [hep-ex].

[159] 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*. ATLAS-CONF-2018-031. 2018. URL: https://cds.cern.ch/record/2629412.

[160] CMS Collaboration. *Search for ttH production in the H-to-bb decay channel with leptonic tt decays in proton-proton collisions at sqrt(s) = 13 TeV with the CMS detector*. CMS-PAS-HIG-17-026. 2018. URL: https://cds.cern.ch/record/2308267.

[161] CMS Collaboration. "Observation of $t\bar{t}H$ production". In: *Phys. Rev. Lett.* 120.23 (2018), p. 231801. DOI: 10.1103/PhysRevLett.120.231801. arXiv: 1804.02610 [hep-ex].

[162] CMS Collaboration. "Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: *Submitted to: Eur. Phys. J.* (2018). arXiv: 1809.10733 [hep-ex].

[163] G Apollinari et al. *High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report*. CERN Yellow Reports: Monographs. 2015. URL: http://cds.cern.ch/record/2116337.

[164] ATLAS Collaboration. *Projections for measurements of Higgs boson cross sections, branching ratios, coupling parameters and mass with the ATLAS detector at the HL-LHC*. ATL-PHYS-PUB-2018-054. 2018. URL: http://cds.cern.ch/record/2652762.

[165] ATLAS Collaboration. "Erratum to: Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector". In: *The European Physical Journal C* 76.3 (2016), p. 152. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-3934-y.

[166] Super-Kamiokande Collaboration. "Erratum: Measurements of the Solar Neutrino Flux from Super-Kamiokande's First 300 Days [Phys. Rev. Lett. 81, 1158 (1998)]". In: *Phys. Rev. Lett.* 81 (19 1998), pp. 4279–4279. DOI: 10.1103/PhysRevLett.81.4279. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.4279.

[167] Glen Cowan et al. "Erratum to: Asymptotic formulae for likelihood-based tests of new physics". In: *The European Physical Journal C* 73.7 (2013), p. 2501. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-013-2501-z.