University of Glasgow

Rozario, Zefran (2026) *Measurement of the $t\bar{t}H(b\bar{b})$ process using the ATLAS detector with an effective field theory interpretation*. PhD thesis.

https://theses.gla.ac.uk/85679/

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

School of Physics and Astronomy
University of Glasgow
Glasgow
G12 8QQ

PhD Thesis

Measurement of the $t\bar{t}H(b\bar{b})$ process using the ATLAS detector with an effective field theory interpretation

Zefran Rozario

January 5, 2026

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

Abstract

The associated production of a Higgs boson with a top-quark pair ($t\bar{t}H$) provides a direct probe of the top–Higgs Yukawa coupling, the largest coupling in the standard model. Constraining this process is essential for testing the standard model and for probing possible contributions from physics beyond the standard model. This thesis presents a measurement of $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel, using the full Run 2 dataset, 140 fb^{-1} of proton-proton collision data at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, collected by the ATLAS detector at the LHC. Included in this analysis are: novel machine learning methods, improved modelling and reconstruction techniques, dedicated high p_T uncertainty extrapolation and more. An excess of events over the non-$t\bar{t}H$ background is found with an observed (expected) significance of 4.6 (5.4) standard deviations. The $t\bar{t}H$ cross-section is measured to be $\sigma_{t\bar{t}H} = 411^{+101}_{-92}$ fb $= 411 \pm 54$ (stat.) $^{+85}_{-75}$ (syst.) fb, for a Higgs boson mass of 125.09 GeV, consistent with the standard model prediction of 507^{+35}_{-50} fb. The cross-section is also measured differentially in bins of the Higgs boson transverse momentum within the Simplified Template Cross-Section framework. This analysis is then re-interpreted in the Standard Model Effective Field Theory framework. The three main operators governing the ttH production in this framework, $\mathcal{O}_{tG}, \mathcal{O}_{t\phi}, \mathcal{O}_{\phi G}$, are investigated and constraints are applied to their respective Wilson coefficients $c_{tG}, c_{t\phi}, c_{\phi G}$. Improvements in precision of the Wilson coefficients were seen in individual quadratic fits compared to early Run-2 (2016) studies, particularly with $c_{t\phi} = 2.55^{+2.22}_{-2.15}$, whose contributions are unique to the $t\bar{t}H$ production mode, sees improvements in precision by a factor of 2.

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university.

Zefran Rozario

Acknowledgements

First and foremost I would like to thank the $t\bar{t}H$ team at Glasgow. Without whom, I would have floundered and accomplished nothing during my PhD. Tony, Mark, Bruno and Giuseppe, you have all taught me so much - Thank you.

After all these years I have spent in the Kelvin building I have never met a more welcoming group that the PPE department. Dima, Andy, Kenny, Katie, Jonathan, to you and everyone else in the department: Thank you for the help, support, squash, chats over a pint and escapades at conferences.

To my peers, Elliot, Anna, Rory, Tom, Iain, Jay, Dylan and everyone I have met along the way at Geneva, Glasgow, Oxford, Oviedo and everywhere else: You've all brought so much joy, love and late nights over Guinness into my life, here's to more.

There were small additions to the family household recently. Mac and Coco: A couple of kittens with noodles for brains. Bennie, I'm sorry you never got to meet them, you would have hated their lack of personal space and uneducated palate with respect to tuna.

Speaking of the family. Mum, Dad and Zoe, thank you for all the support over the years even when we were all begin needlessly difficult or when life decided to hand over a poorly timed lemon. I love you all.

Finally, to Hannah: I'm so lucky to have you in my life. You are my inspiration. I love you. Yes, I will put the shelves up soon.

Contents

"What happened to your enlightenment?
I don't know. It wore off."

— Don Draper and Roger Sterling, Mad Men.

Chapter 1

Introduction

"There are no bad ideas, Lemon. Only good ideas that go horribly wrong."
— Jack Donaghy to Liz lemon, 30 rock.

This thesis presents a measurement of the cross-section of the Higgs boson production in association with a pair of top quarks and decaying to a pair of bottom quarks, $t\bar{t}H(b\bar{b})$. Measuring this process is particularly challenging because of its relatively small cross-section and the presence of large irreducible $t\bar{t}$+jets backgrounds. Despite these difficulties, it offers unique direct access to the interaction between the Higgs boson and the top quark, referred to as the top Yukawa coupling. As the top quark is the heaviest known fermion, its coupling to the Higgs boson is the strongest Yukawa interaction in the Standard Model (SM). A precise measurement of this coupling provides a powerful probe of new physics, since any significant deviation from the SM expectation could signal contributions from yet-unknown particles or interactions. As such, the result of this analysis is then re-interpreted within the Standard Model Effective Field Theory (SMEFT) framework in order to test for deviations in the SM measurement. To arrive at the final results, that are the constrains of the operators governing the $t\bar{t}H$ production mode in this framework, this thesis is constructed as follows.

Chapter 2 provides an overview of the SM, laying out its construction and main features including the production of the Higgs boson and the $t\bar{t}H$ production mode. Then an overview of the SMEFT as it relates to $t\bar{t}H$ and describes how the main operators can affect the measurement of $t\bar{t}H$.

Chapter 3 begins by describing the Large Hadron Collider accelerator chain. The ATLAS detector and the detector subsystems are then discussed.

Chapter 4 investigates the Inner Tracker upgrade that will move the ATLAS detector into the High-Luminosity era. Outlined here is the qualification task completed in order to be a qualified ATLAS author. The task was to create a set of reporting tools for the global upgrade team to track production across institutes and regions.

Chapter 5 provides the definitions and reconstruction techniques used to create the physics objects in the $t\bar{t}H(b\bar{b})$ analysis. Effort is concentrated in the discussion of particle flow jets and flavour tagging algorithms as this provides the main background for the high p_T pseudo continuous b-tagging particle flow jet uncertainty extrapolation study detailed later.

Chapter 6 gives an overview of the Monte-Carlo Simulations and the techniques used in order to create appropriate simulated events to compare with data in a physics analysis.

Chapter 7 provides detail on my main contribution to the $t\bar{t}H(b\bar{b})$ analysis: the development and validation of the high p_T pseudo continuous b-tagging particle flow jet uncertainty extrapolation. This allowed for the proper systematic treatment of the uncertainties of the b, c and light flavour particle flow jets, produced above their calibration threshold, to be evaluated and folded into an analysis in a plug and play style. This had direct impact on the selected events in the high p_T resolved and boosted regions of the $t\bar{t}H(b\bar{b})$ analysis. Helping the high p_T regions also directly affected the statistical power in the context of a re-interpretation such as that used in the SMEFT analysis.

Chapter 8 describes the main analysis techniques used in the $t\bar{t}H(b\bar{b})$ analysis, this is split up into town main sections: one on the foundation of the statistical model, profile likelihoods, and one on the machine learning techniques, multi-headed attention based transformer, used in the $t\bar{t}H(b\bar{b})$ analysis multi variate analysis strategy.

Chapter 9 details the $t\bar{t}H(b\bar{b})$ analysis in its entirety. First discussing the previous $t\bar{t}H(b\bar{b})$ ATLAS result, published in 2022, then going thorough the major components of the analysis. The event selection, event modelling, multi variate analysis strategy for event classification and Higgs boson reconstruction, definitions of the analysis regions, the systematic uncertainties of the analysis, the construction of the statistical model and

fit studies conducted and finally the results are presented.

Chapter 10 presents the SMEFT re-interpretation of the $t\bar{t}H(b\bar{b})$ analysis. The chapter opens with a discussion on the current state of $t\bar{t}H$ BSM interpretation from the Run-2 data set. The analysis methodology is then laid out followed by results and the discussion of individual and combined operator fits for the linear SMEFT framework and then individual fits for the quadratic fits.

Chapter 11 opens by discussing the impact of the results obtained and their limiting factors. The thesis is then concluded with a discussion on the outlook of potential future $t\bar{t}H(b\bar{b})$ measurements and the power they have in the world of re-interpretation.

Chapter 2

Theory

"The discography of ABBA."

— Unknown, when asked about the greatest works of human history.

2.1 Introduction

The SM of particle physics is a fundamental description of modern day particle physics. It is the cumulation of over a half-century's intellectual efforts and describes, all the known, particle interactions. However, it is not without its shortfalls in order to understand these let us begin by outlining the theory.

The SM is a relativistic quantum field theory (QFT) and is made up of three generations of fundamental particles, shown in Figure 2.1. We can describe the SM as a gauge group, meaning it is invariant under non-local (gauge) transformations and has the group structure of:

$$\mathcal{G}_{\text{SM}} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y. \tag{2.1}$$

According to Noether's theorem [1]: for every symmetry there is a corresponding conserved quantity. Here, $SU(3)_C$ is the gauge group for quantum chromodynamics, where the conserved quantity is colour charge, C. $SU(2)_L$, encapsulates the weak interactions and its generators, here the conserved quantity is weak iso-spin, L. Finally we have $U(1)_Y$ which describes the electroweak interactions in the SM and the conserved quantity is

hypercharge, Y. The combination of gauge groups describing the weak interaction and electromagnetic sector, namely: $SU(2)_L \otimes U(1)_Y$ is known as the electroweak sector [2, 3, 4, 5].

2.2 $\mathcal{L}_{Gauge}$

To account for these gauge symmetries in a locally invariant manner, we introduce gauge fields associated with each component of the gauge group. These gauge fields are responsible for mediating the interactions and correspond to the bosonic force carriers. Each gauge group contributes its own field:

$$\begin{aligned} SU(3) : G^a_\mu,\ a = 1, 2, 3..., 8 \\ SU(2) \otimes U(1) : W^J_\mu,\ J = 1, 2, 3;\ B_\mu \end{aligned} \tag{2.2}$$

where G^a_μ denotes the 8 independent gluon fields and a the eight gluons. They are also the generators of the $SU(3)$ group. W^J_μ, where $J = 1, 2, 3$, represents the three gauge fields and B_μ is the gauge field associated with the $U(1)$ hypercharge symmetry. These gauge fields are of spin-1 which are said to facilitate the interactions of the SM. We can use these gauge fields to construct the Yang-Mills Lagrangian and thus capture the dynamics of these fields. The Yang-Mills Lagrangian [6] takes the form:

$$\mathcal{L}_{gauge} = -\frac{1}{4} F^{a\mu\nu} F^a_{\mu\nu}, \tag{2.3}$$

where $F^a_{\mu\nu}$ is the field strength tensor. Accepting the definitions of: $A_\mu = A^a_\mu T_a$ and $F_{\mu\nu} = F^a_{\mu\nu} T_a$, where T_a are the generators of the gauge group, we relate the field strength tensor to the commutator of the gauge covariant derivative D_μ by:

$$F_{\mu\nu} = \frac{i}{g}[D_\mu, D_\nu], \tag{2.4}$$

$$= \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu], \tag{2.5}$$

where $A_{\mu,\nu}$ are gauge fields and g is the coupling strength of the interaction facilitated by the gauge group. Where D_μ is defined as:

$$D_\mu = \partial_\mu - igA_\mu, \tag{2.6}$$

where, g is gauge coupling constant such that the gauge covariant derivative, allows the gauge field to interact with matter fields in a way that respects gauge symmetry, ensuring consistent transformation properties under local gauge transformations. Both are foundational in constructing gauge-invariant Lagrangians, which are essential for describing interactions in the SM. Extending this general definition of the covariant derivative to all the gauge fields from our three gauge symmetries we have the SM covariant derivative:

$$D_\mu = \partial - ig_1 \frac{Y}{2} B_\mu - ig_2 \frac{\tau_j}{2} W_\mu^i, \tag{2.7}$$

where, g_1 and g_2 are the hypercharge and weak iso-spin coupling constant and τ_j is a Pauli matrix.

2.3 $\mathcal{L}_{\text{Fermion}}$

As stated earlier, the covariant derivative allows for the interactions of gauge fields and matter fields. Consider a set of matter fields, $\Psi = \{\psi\}$ that transform under a representation of the gauge groups. These represent the fermionic fields that house the quarks and leptons that matter is made of. The SM fermionic fields can be summarised by,

$$L_i = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L,\ e^-_{Ri},\ Q_i = \begin{pmatrix} u \\ d \end{pmatrix}_L,\ u_{Ri},\ d_{Ri},\ i = \{1, 2, 3\}, \tag{2.8}$$

where i represents the three generations of matter. We let these fields act on the covariant derivative such that,

$$\mathcal{L}_{\text{kinetic}} = i\bar{\Psi} D_\mu \Psi. \tag{2.9}$$

2.4 $\mathcal{L}_{Higgs}$

The Higgs mechanism [7, 8, 9] was theorised to answer the simple yet puzzling question of how do elementary particles acquire their mass? This question posed particular trouble in the early days of the unification of the weak force and the electromagnetic force (the

electroweak sector). Early gauge invariant theories contained symmetry requirements that required the W and Z gauge bosons to have zero mass. This, was in contradiction to the experimental results where they had non-zero masses. The introduction of the Higgs mechanism solves this problem. The Higgs sector consists of a doublet of complex scalar fields with hypercharge $Y_\phi = +1$

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{2.10}$$

where the invariant scalar potential is:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2. \tag{2.11}$$

Solving for the minimum of V, we see that

$$V'(|\phi|)^2 = 0 \Rightarrow |\phi|^2 = -\frac{\mu^2}{2\lambda}. \tag{2.12}$$

Let us now consider the quartic coupling to be positive ($\lambda > 0$) and the mass term squared to be negative ($\mu^2 < 0$). In this case we see that the fields develops a non-zero vacuum expectation value (VEV).

$$\phi \to \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix},$$

$$\nu = \sqrt{\frac{-\mu^2}{\lambda}} = \frac{1}{(\sqrt{2} G_F)^{\frac{1}{2}}} = 246\text{GeV}, \tag{2.13}$$

where G_F is the Fermi constant and the remaining degree of freedom, h, is the Higgs boson. We see from the Weinberg-Salam model [10] that the VEV is directly related to the masses of the W and Z bosons:

$$M_W = \frac{1}{2}\nu g, \; M_Z = \frac{1}{2}\nu\sqrt{g_1^2 + g_2^2}, \tag{2.14}$$

where g and g' are the $SU(2)_L$ and $U(1)_Y$ coupling constants, respectively. These are all related by the electroweak mixing angle θ_w:

$$\sin^2(\theta_w) = \frac{g_1^2}{g_1^2 + g_2^2} = 1 - \frac{M_W^2}{M_Z^2}. \tag{2.15}$$

It is worth mentioning that in order to generate these masses the electroweak symmetry spontaneously breaks (EWSB):

$$SU(2)_L \times U(1)_Y \Rightarrow U(1)_{EM}.$$

The Higgs mechanism is also responsible for generating the intrinsic masses of the elementary particles.

2.5 $\mathcal{L}_{Yukawa}$

Yukawa interactions [11] describe the interaction between fermions and a scalar field. Adapting this to the idea of the SM we can generate mass terms for fermions (with both left and right handed chirality) as they interact with the Higgs field (a scalar field) via Yukawa interactions. Let us now consider the full unabridged Yukawa lagrangian for the Higgs field and first generation fermions:

$$\mathcal{L}_{Yukawa} = -\lambda_e \bar{L}\phi e_R - \lambda_d \bar{Q}\phi d_R - \lambda_u \bar{Q}\tilde{\phi} u_R + h.c. \tag{2.16}$$

Where $\bar{L}$ and $\bar{Q}$ are the Dirac ajoint of the left handed lepton and quark doublets from Equation 2.8, respectively. These are defined as $\bar{L} = L^\dagger\gamma^0$ and $\bar{Q} = Q^\dagger\gamma^0$ and $\tilde{\phi}$ is the conjugate of the Higgs doublet. The Higgs effective Lagrangian will now contain the fermionic mass terms. Thus, the full effective Lagrangian is:

$$\mathcal{L}_{Higgs} = g_{hWW} h W_\mu^+ W^{-\mu} + g_{hZZ} h Z_\mu Z^\mu + \sum_f g_{hff} h \bar{f}_L f_R. \tag{2.17}$$

The Higgs interaction strength increases with the mass of the particle it is interacting with. Thus, the couplings to the gauge bosons and fermions[i] can be written as:

$$g_{hff} = \frac{m_f}{\nu},\ g_{hWW} = \frac{2M_W^2}{\nu},\ g_{hZZ} = \frac{M_Z^2}{\nu}. \tag{2.18}$$

[i]The masses of the second and third generation of fermions are generated in the same way.

2.6 $\mathcal{L}_{SM}$

Now that we have the full description of all the various sectors of the SM, we can put them together to complete the SM Lagrangian, given by,

$$\begin{aligned}\mathcal{L}_{SM} &= \mathcal{L}_{Gauge} + \mathcal{L}_{Fermion} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \\ &= -\frac{1}{4}G^{a}_{\mu\nu}G^{a\mu\nu} + i\bar{\Psi}D_{\mu}\Psi + |D_{\mu}\phi|^2 - V(\phi) + (\psi_i y_{ij}\psi_j\phi + h.c),\end{aligned} \tag{2.19}$$

the above equation summarises the interactions of the gauge bosons with themselves, the interactions of the Higgs field with itself, the gauge bosons and fermions and finally the fermions with gauge bosons. The interactions that include gauge bosons are described as the forces of the SM. The interactions with the Higgs field result in these particles acquiring masses post EWSB. In the SM we have three generations of matter, as to why exactly three we are yet to understand. Each generation of matter consists of a pair of quarks, a lepton and a left handed (massless) neutrino, this is summarised in Figure 2.1.

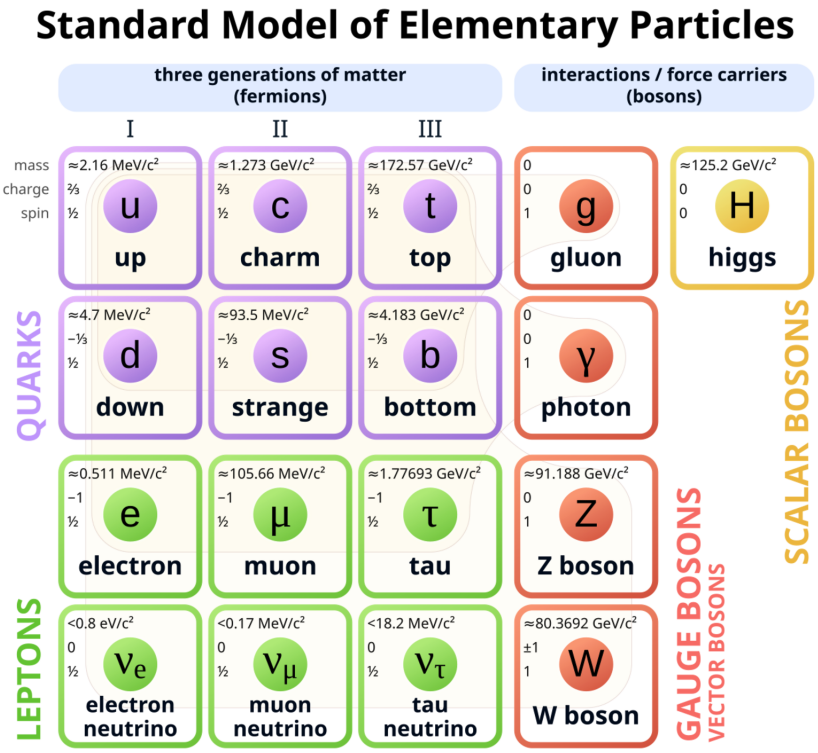
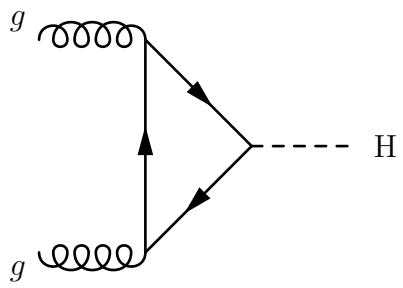


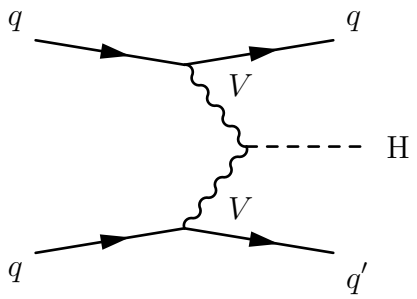
Figure 2.1 The fundamental particles of the SM organised into three generations of matter and the bosons and the force carriers [12].

2.7 The Higgs Boson

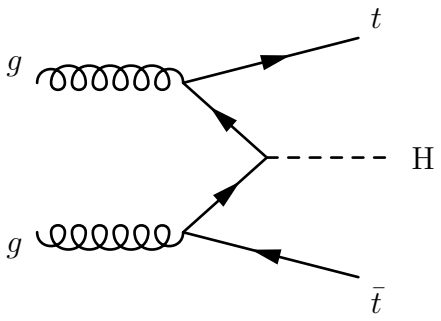
The crown jewel of the SM is the theorisation and discovery of the Higgs boson. Since its discovery in 2012, the main goal of particle physicists has been to study the Higgs boson in all of its, production and decay modes at proton-proton colliders. The four main production modes of the Higgs boson are: gluon-gluon fusion (ggF), vector boson fusion (VBF), associated production with a vector boson (VH) and associated production with a top quark pair ($t\bar{t}H$). The Feynman diagrams for these production modes can be seen in Figure 2.2. The Higgs boson production cross section, for these production modes, as a function of centre of mass energy can be seen in Figure 2.3.



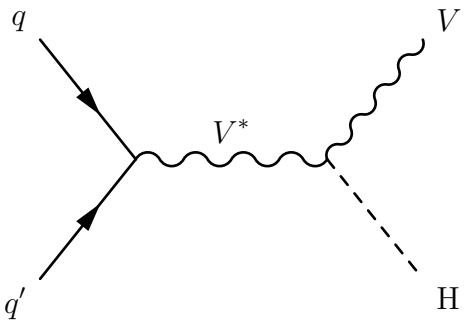
(a) Gluon-gluon fusion (ggF)



(b) Vector boson fusion (VBF)



(c) Top quark associated production ($t\bar{t}H$)



(d) Higgs-strahlung

Figure 2.2 Main leading order Feynman diagrams of the Higgs boson production. (a) ggF, where gluons form a loop of virtual top or bottom quarks resulting in the production of a Higgs. (b) VBF, where quarks, q, radiate off a massive vector boson which then fuse to create a Higgs boson. (c) $t\bar{t}H$, shows gluons decaying into top and anti-top pair, a top or anti-top from each gluon then fuses to create a Higgs boson leaving the newly formed Higgs boson and a top and anti-top pair. (d) Shows a quark creating vector boson which radiates off a Higgs bosons.

The main decay modes of the Higgs boson and their corresponding fractional branching ratio are as follows: $H \to b\bar{b}$ ($\sim 58\%$), $H \to WW^*$ ($\sim 21\%$), $H \to \tau\tau$ ($\sim 6.3\%$), $H \to ZZ$

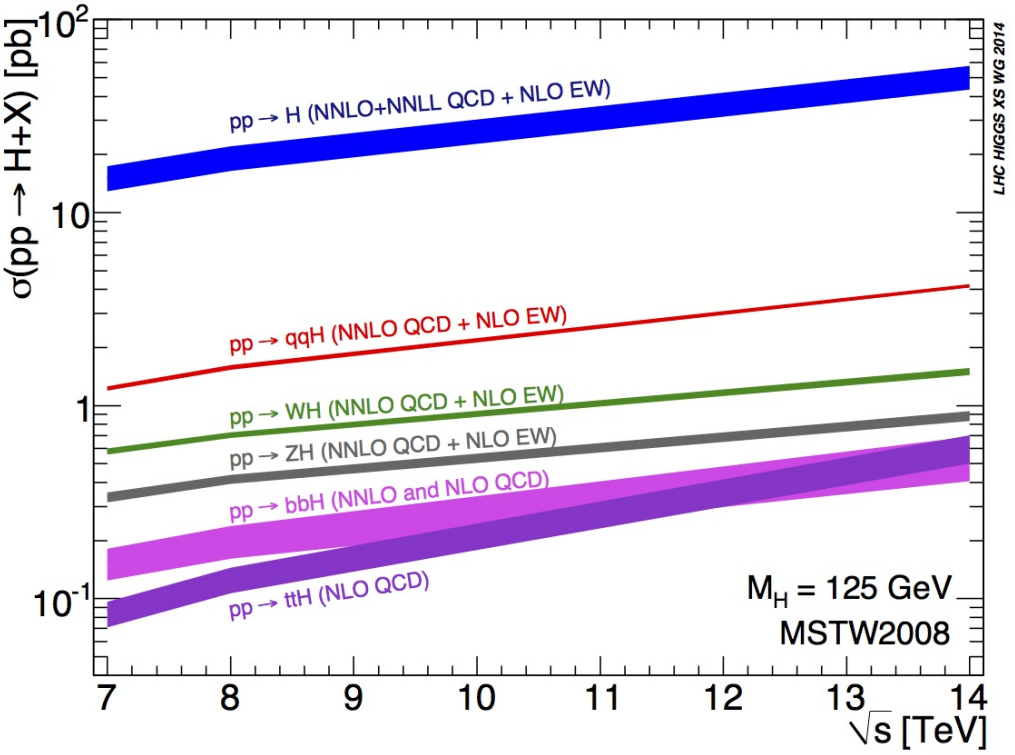


Figure 2.3 Higgs boson production cross-section, as a function of centre of mass energy $\sqrt{s}$ at a proton-proton collider for a Higgs mass, $M_H = 125$ GeV [13].

($\sim 2.6\%$), $H \to \gamma\gamma$ ($\sim 2.3\%$), $H \to Z\gamma$ ($\sim 1.5\%$) and finally $H \to \mu\mu$ ($\sim 0.2\%$). The Higgs boson decay branching ratios as a function of Higgs mass is shown in Figure 2.4.

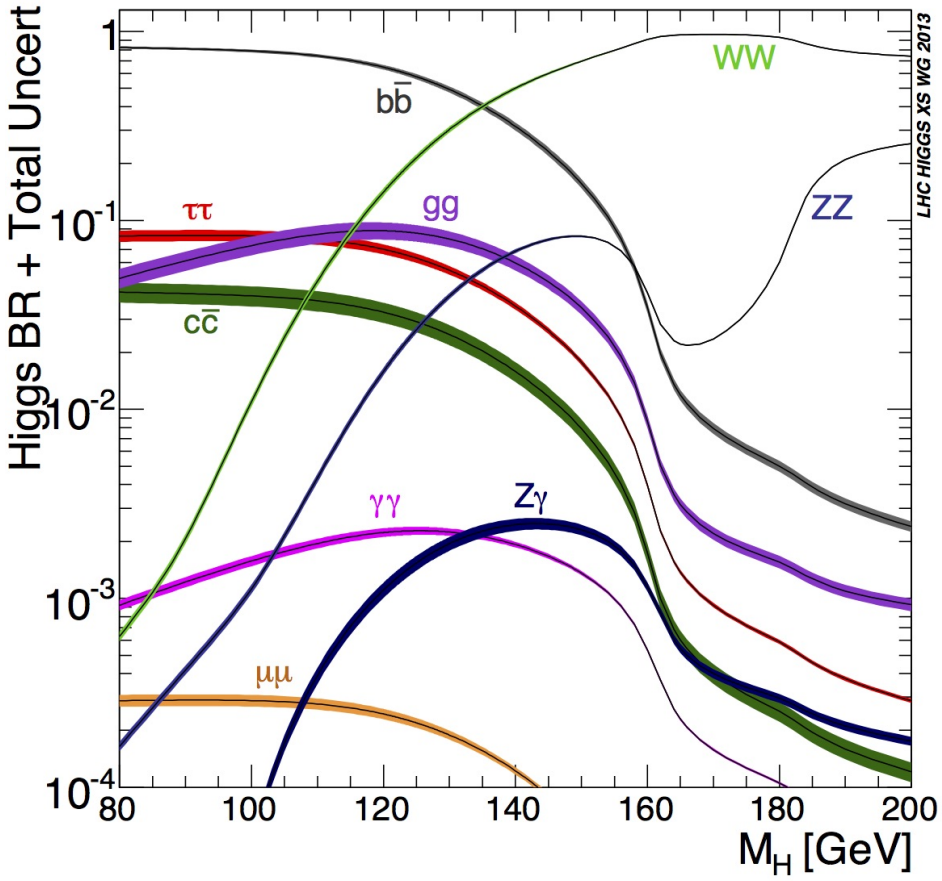


Figure 2.4 The SM Higgs boson decay branching ratios and widths as a function of Higgs mass [14]

2.7.1 Higgs-Top Yukawa Coupling, y_t

As the top quark is the heaviest of all the known fundamental particles, measuring the strength of its Yukawa coupling, y_t, is of particular interest as it may shed some light on the hierarchy problem, discussed below.

We define the measured mass of the Higgs boson as:

$$m^2_{H,obs} = \tilde{m}^2_{H,bare} + \Delta m^2_H, \tag{2.20}$$

where $\tilde{m}^2_{H,bare}$ encapsulates the mass contributions of the theory parameters and Δm^2_H are the quantum corrections from loop diagrams as shown in Figure 2.5. The top quark

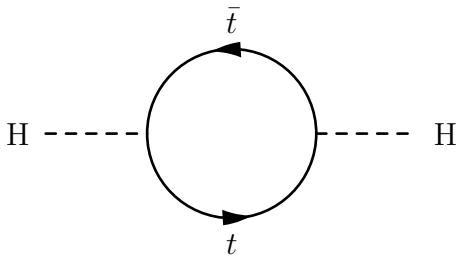


Figure 2.5 Diagram representing the one-loop correction to the Higgs mass due to a top quark loop.

contributions of the quantum corrections to the Higgs boson mass are:

$$\Delta m^2_H \approx -\frac{|y_t|^2}{16\pi^2}\Lambda^2, \tag{2.21}$$

where Λ is the energy scale at which new physics (NP) sits at. If we set $\Lambda \sim 10^{19}$GeV (the Planck scale) we see that in order for $m^2_{H,obs} = 125$ GeV we must fine tune $\tilde{m}^2_{H,bare}$ in order to cancel out the large quantum corrections. This coincidence is called the naturalness problem [15]. We see that y_t plays a key role in this conundrum, as the top quark is the heaviest known particle and its contribution would be the dominant factor in the value of Δm^2_H. So a precise measurement of y_t would help estimate the scale of NP and identity where potential beyond the standard model (BSM) effects could take place. For instance, if we find larger than expected values of y_t, this would imply larger corrections to the Higgs boson mass potentially requiring NP close to the weak scale ($\mathcal{O}(100)$GeV).

2.7.2 $t\bar{t}H$

The $t\bar{t}H$ process, as seen in Figure 2.2c, plays a special role in the SM as it allows for direct tree-level measurement of the top quark Yukawa coupling, y_t. We see that from Figure 2.3 that the $t\bar{t}H$ process is the least favoured production mode and so we expect the rate at which we observe the $t\bar{t}H$ process to be much lower than the other production modes. Even with this limitation, the $t\bar{t}H$ production was observed by ATLAS [16] and CMS [17] in 2018 by combining various decay channels. As we see from Figure 2.4, $H \to b\bar{b}$ is the dominant decay mode of the Higgs. Most likely that if a $t\bar{t}H$ event occurred, the Higgs boson would probably decay into a $b\bar{b}$ pair and so in order to have a

statistical advantage in observing $t\bar{t}H$ one would only require to look here [ii]. As simple as this plan is, it is fraught with difficulty, namely $H \to b\bar{b}$ does not have a clear signal to trigger on and so makes signal extraction, from its almost identical background, extremely difficult. There have been a few attempts at a $t\bar{t}H(b\bar{b})$ analysis in the past by both ATLAS [16] and CMS [18], these analyses were later combined to reach the 5σ threshold for discovery, and were based on partial Run-2 data sets. Later, ATLAS released a full Run-2 analysis [19] however there has since been a reinterpretation of the full Run-2 ATLAS data set, in a so-called $t\bar{t}H(b\bar{b})$ legacy analysis. This is where most of my thesis lays and I lay these out in Chapter 9.

2.8 Beyond the Standard Model: Effective Field Theories

As we begin our journey in physics, we are told that: Isaac Newton *discovered* gravity or Christiaan Huygens *discovered* that light is a wave. Regarding the idea of discovery with the same verve as that of a curious child spotting insects under a rock for the first time. Later in our journey, once we are shown vector calculus for possibly the third time in what seems to be unrelated university modules, it clicks: We as physicists we only describe the world around us and hope to do it in a tidy enough package that our work can be considered fundamental.

Returning to Issac and his apple, a correct statement would be: Newtons laws of motion and universal law of gravity correctly *describe* bodies in motion. We know that, although brilliant, Newtons works doesn't describe everything, such as: Black-holes, Gravitational lensing, the behaviour of fast moving objects, the list goes on. The solution to all of this is of course the work of Einstein and his general theory of relativity. This of course doesn't make Newtons laws obsolete, they still *effectively* describe the motion of bodies up to a certain point.

These two theories are a good example of how we can use different dynamics for different energy scales, we can also recover Newtons laws (the effective theory) from general relativity (the full theory) by removing heavy degrees of freedom while still *effectively* describing the dynamics at play (provided you're not moving near the speed of light).

[ii] The best laid schemes o' mice an' men / Gang aft agley, / An' lea'e us nought but grief an' pain

I raise the analogy above as we can define the SM as an effective theory which works at a lower energy scale than a full theory which sits at a higher energy scale. In a Standard Model Effective Field Theory[iii] (SMEFT) [21] we can extend the SM with higher dimension operators constructed from SM parameters, $\mathcal{O}$ and Wilson coefficients, c_i that modify the contributions of the operators. This enables us to quantify the possible effects of new physics that sit at a higher energy scale, Λ. The SMEFT allows for a model independent approach, we can define the SMEFT lagrangain as:

$$\mathcal{L}_{SMEFT} = \mathcal{L}_{SM} + \sum \frac{c_i}{\Lambda^2}\mathcal{O}_i^{d=6} + \frac{c_i}{\Lambda^4}\mathcal{O}_i^{d=8} \dots \tag{2.22}$$

where the scale of the new physics suppress the BSM effects. We can express the SMEFT in terms of physical observables, for now let us consider cross-section, σ:

$$\sigma_{SMEFT} = \sigma_{SM} + \frac{c_i}{\Lambda^2}\sigma_i^{\text{interf.}} + \frac{c_i^2}{\Lambda^4}\sigma_i^{\text{quad.}}. \tag{2.23}$$

Where $\sigma_i^{\text{interf.}}$ is the interference term between the SM and the BSM, and $\sigma_i^{\text{quad.}}$ is a term that is purely dependent on BSM effects. Using Equation 2.23 we can see how powerful a reinterpretation of particle physics analysis using this framework could be, allowing us to gain hints of potential BSM effects in our SM measurements, as opposed to feeling around in the dark or lifting up a rock, surprised to not find any insects.

2.8.1 $t\bar{t}H$ SMEFT

Following the framework in Section 2.8, there are three main dimension-6 operators that govern the $t\bar{t}H$ process, as shown in Table 2.1. Dimension-6 operators are considered because they provide the leading, model-independent corrections to the Standard Model in the SMEFT expansion: dimension-5 operators are either forbidden by SM gauge symmetries or violate lepton number, while dimension-8 operators are suppressed by additional powers of the new-physics scale, Λ, and therefore contribute less so at LHC energies, making dimension-6 the lowest-order operators with observable impact in precision measurements. The definition of the operators are made from components of the SM and so have their usual SM definitions. Our convention is such that for each operator

[iii] There are of course other effective field theories that are floating around in modern day particle physics such as Heavy Quark Effective Field Theory [20] designed to limit the dynamics of heavy quark interactions, for the scope of this work we will focus on the SMEFT.

we will add its Hermitian conjugate, even if it is already Hermitian. We want to conserve CP-symmetry and so consider $c_{t\phi}$ and c_{tG} to be real. $c_{\phi G}$ is always real as the operator is, $\mathcal{O}_{\phi G}$, is Hermitian.

$\mathcal{O}_{t\phi}$ rescales the top Yukawa coupling in the SM, while also giving rise to a new $ttHH$ coupling. Explicitly, this operator is written as $y_t^3(\phi^\dagger\phi)(\bar{Q}t)\tilde{\phi}$, involving the Higgs doublet ϕ, the left-handed third-generation quark doublet $\bar{Q}$, and the right-handed top quark t. After EWSB, the Higgs field generates a shift in the effective $t\bar{t}h$ vertex proportional to y_t^3 and a new $t\bar{t}hh$ contact term. These new interactions affect single-Higgs production by modifying the top Yukawa strength and contribute to Higgs pair production via diagrams where two Higgs bosons couple to a top loop or are emitted directly through the contact interaction.

$\mathcal{O}_{tG}$ is the chromomagnetic dipole operator, introducing a dipole-type interaction between the top quark and gluons. It is expressed as $y_t g_s(\bar{Q}\sigma^{\mu\nu}T^A t)\tilde{\phi}G^a_{\mu\nu}$. After EWSB, the Higgs field's vacuum expectation value provides a constant term and physical Higgs insertions, leading to multiple interaction structures. At leading order, the operator modifies the gtt vertex in the SM, adding a momentum-dependent dipole form factor. It also introduces new higher-point vertices such as $ggtt$ and $gttH$, as well as a five-point $ggttH$ vertex. These additional couplings arise because the gluon field strength contains non-linear terms in the gluon fields, and the Higgs doublet expansion adds Higgs insertions. The momentum dependence of these interactions enhances their contribution in kinematic regions with large transverse momentum, making them especially relevant in boosted top or boosted Higgs searches.

$\mathcal{O}_{\phi G}$ is the Higgs–gluon contact term, given by $y_t^2(\phi^\dagger\phi)G^a_{\mu\nu}G^{a\mu\nu}$. This operator couples the Higgs bilinear directly to the gluon field strength tensor. Upon EWSB, the Higgs bilinear expands as $\phi^\dagger\phi \to \frac{1}{2}(v^2+2vh+h^2)$, generating a constant term that renormalizes the gluon kinetic term, a tree-level hgg coupling, and a contact $hhgg$ interaction. In the SM, these couplings appear at loop level via top-quark loops, but here they are generated directly at dimension six. As a result, $\mathcal{O}_{\phi G}$ can significantly modify the Higgs production rate in gluon fusion and enhance Higgs pair production by introducing a direct two-Higgs–two-gluon vertex.

$\mathcal{O}_{t\phi}$ is the main operator of interest as its uniquely constrained by $t\bar{t}H$ production and so only an analysis, such as that detailed in Chapter 9, and a SMEFT re-interpretation

of the analysis, detailed in Chapter 10, can provide insight to the current strength of the operator. The other two operators are better constrained in global fits as they are influenced by many Higgs production modes.

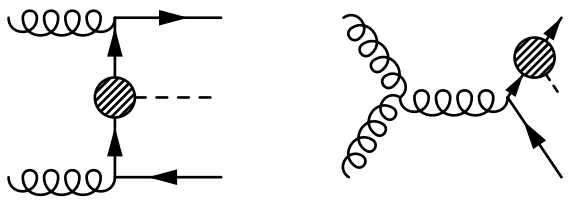
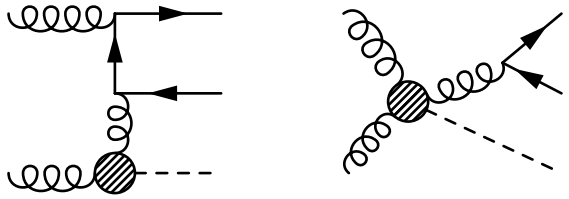
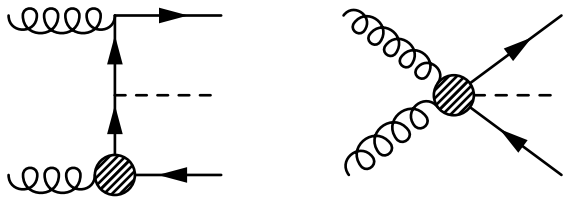
Operator	**Diagrams**
$\mathcal{O}_{t\phi} = y_t^3(\phi^\dagger\phi)(\bar{Q}t)\tilde{\phi}$	
$\mathcal{O}_{\phi G} = y_t^2(\phi^\dagger\phi)G^a_{\mu\nu}G^{a\mu\nu}$	
$\mathcal{O}_{tG} = y_t g_s(\bar{Q}\sigma^{\mu\nu}T^A t)\tilde{\phi}G^a_{\mu\nu}$	

Table 2.1 The three main SMEFT operators for the $t\bar{t}H$ production, their definitions and Feynman diagrams showing which vertex they modify.

Chapter 3

The Large Hadron Collider and ATLAS experiment

"Bigger..."

— Everyone when I ask them about the FCC

3.1 The Large Hadron Collider

In order to probe the fundamentals of the universe we have to build a powerful particle accelerator in-order to reach high enough energies, and, recreate the conditions moments after the big bang. Enter the Large Hadron Collider (LHC). The LHC straddles the Swiss-French border, near the city of Geneva. It is a hadron accelerator, delivering proton-proton collisions and also supporting collisions of protons with heavy-ions or heavy-ion heavy-ion collisions. The LHC sits in a tunnel with a circumference of 27 km, the tunnel varies in depth sitting between 40 m to 120 m below the surface. The LHC is designed to operate at a centre of mass energy of $\sqrt{s} = 14$ TeV with an instantaneous luminosity of $L = 10^{34}\text{cm}^{-2}\text{s}^{-1}$. The LHC collides beams of accelerated charged particles travelling in opposite directions inside two separate rings, which each contain a vacuum. The beams are deflected by super conducting magnets, keeping them centred in the rings. The beams collide at four focus points around the LHC ring where the major experiments are positioned, these experiments are: ATLAS [22], CMS [23], LHCb [24] and ALICE [25].

3.1.1 The CERN Accelerator Chain

The LHC is fed by a series of particle accelerators designed to take ionized hydrogen gas and accelerate the protons before they reach the LHC. The chain of accelerators are as follows: First we begin with a bottle of hydrogen gas, the electrons are then stripped from the hydrogen, ionizing them resulting in raw protons. These protons are then accelerated up to an energy of 50 MeV in the LINear ACcelerator 2 (LINAC2) [26, 27]. They are then fed into the Proton Synchrotron Booster (PSB) [28, 29] and accelerated to 1.4GeV and subsequently to 25 GeV by the Proton Synchrotron (PS) [30]. Finally, the beam enters the Super Proton Synchrotron (SPS) [31]. At this point the proton beam is split into two and enter the LHC in a clockwise and anti-clock wise direction. Once in the LHC, the proton beams should ideally reach an energy of 7 TeV each, resulting in an ideal centre of mass energy of $\sqrt{s} = 14$ TeV. A detailed layout of the accelerator network can be found in Figure 3.2.

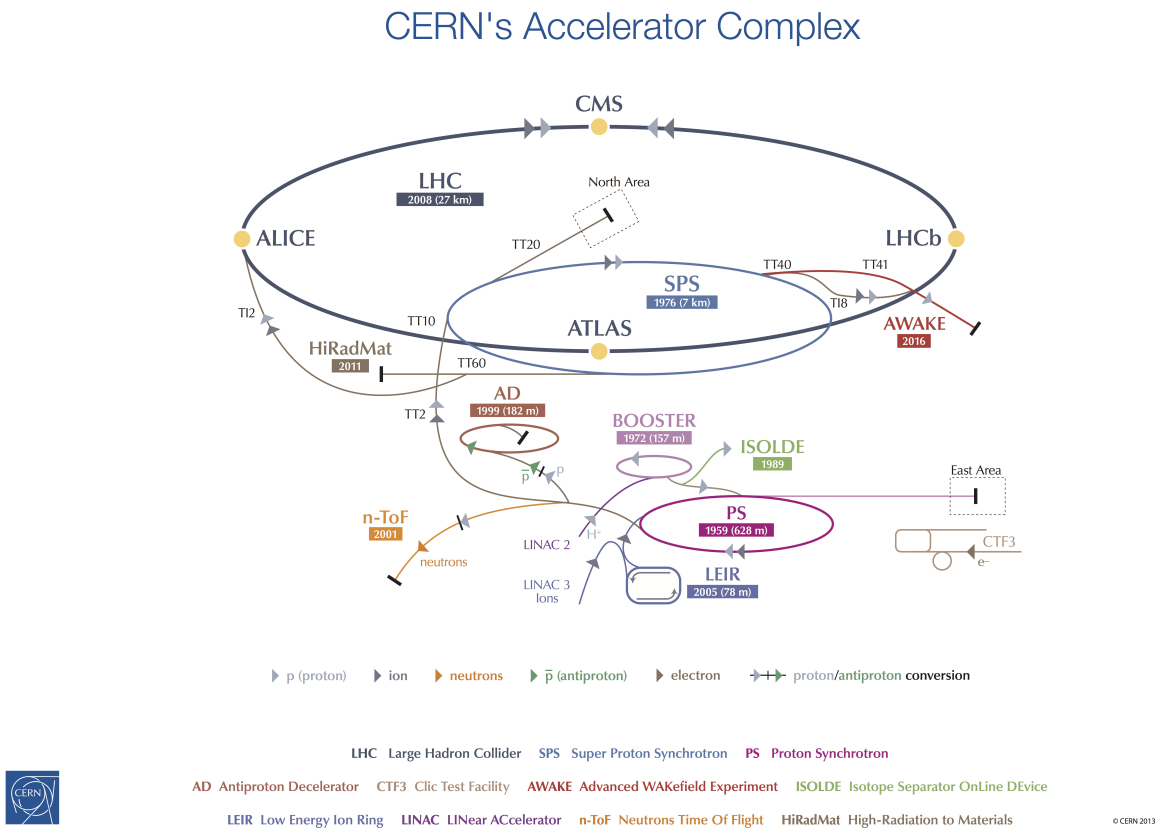


Figure 3.1 A summary diagram of the CERN accelerator network [32].

This thesis will focus on the data taken during 2015-2018, or Run-2. During this period the centre of mass energy stabilised at $\sqrt{s} = 13$ TeV. After the completion of Run-2, the LHC accelerator chain has seen a number of upgrades, the main of which is the introduction of LINAC4 [33], which replaces the now decommissioned LINAC2 and the increase of LHC centre of mass energy to $\sqrt{s} = 13.6$ TeV. LINAC4 is designed to accelerate negative hydrogen ions then stripping the ions of all electrons upon its exit and into the next accelerator in the chain.

3.1.2 Luminosity and Pile-up

The LHC mainly accelerates protons and thus produces proton-proton collisions which we call collision events. The number of events, N_{events}, produced by the LHC as a function of instantaneous luminosity and a function of time, t, is given by [34]:

$$N_{\text{events}} = \sigma_{\text{event}} \int L dt = \sigma_{\text{event}} \mathcal{L}, \tag{3.1}$$

where σ_{event} is the cross-section of your given event and $\mathcal{L}$ is the integrated luminosity. We define the instantaneous luminosity as:

$$L = f \frac{n_1 n_2}{4\pi \sigma_x \sigma_y} N_b. \tag{3.2}$$

Where the nominal spacing between bunches is 25 ns, with a collision frequency of $f = 40$MHz. The number of protons per bunch is denoted by n_1, n_2 (one for each direction of travel of the beams) and N_b is the number of bunches. The height and width of the Gaussian beam profile is given by σ_x and σ_y, respectively. The integrated luminosity is used to quantify the volume of data that is collected over a given period of time.

Due to the large number of protons per bunch, more than one hard scattering interaction takes place per bunch crossing. Interactions beside the interaction of interest are called pile-up events, there are two kinds of pile-up: in-time and out-time pile-up. We define in-time pile-up as when more than one radiation event arrives at a detector simultaneously or very close together in time. This is usually caused by bunch crossings at the central point of collisions. Bunch crossings that result in an extra interaction but occur past the central crossing point are called out-time pile-up.

3.2 The ATLAS Detector

The ATLAS Detector is a general purpose detector, meaning that it conducts a broad range of high energy particle physics analyses, including measurements of SM properties and searches for hints of BSM physics. The ATLAS detector is also the largest of the major particle physics detector experiments based at CERN. It measures 44 m in length with a 25 m diameter and weighs in at 7,000 tonnes. The LHC beam pipe travels along the centre of the cylindrical detector with proton-proton collisions occurring at the, geometric centre of the detector. The detector is designed in layers of sub-detectors, with each layer designed to detect and record different types of particles.

The first layers begins 3 cm away from the beam pipe, and is designed to log the tracks of the charged particles produced immediately after the proton-proton collision. These charged particles then hit layers designed to measure the electromagnetic and hadronic showers produced using dedicated calorimeters for each. Finally the outermost layer of the ATLAS detector is designed to make precision measurements of the muons produced after the proton-proton collisions. There are also two systems of super conducting magnets, to provide magnetic fields for the tracking sub-detectors. The overall detector subsystems are split up into a barrel and two end cap regions providing almost full 4π coverage.

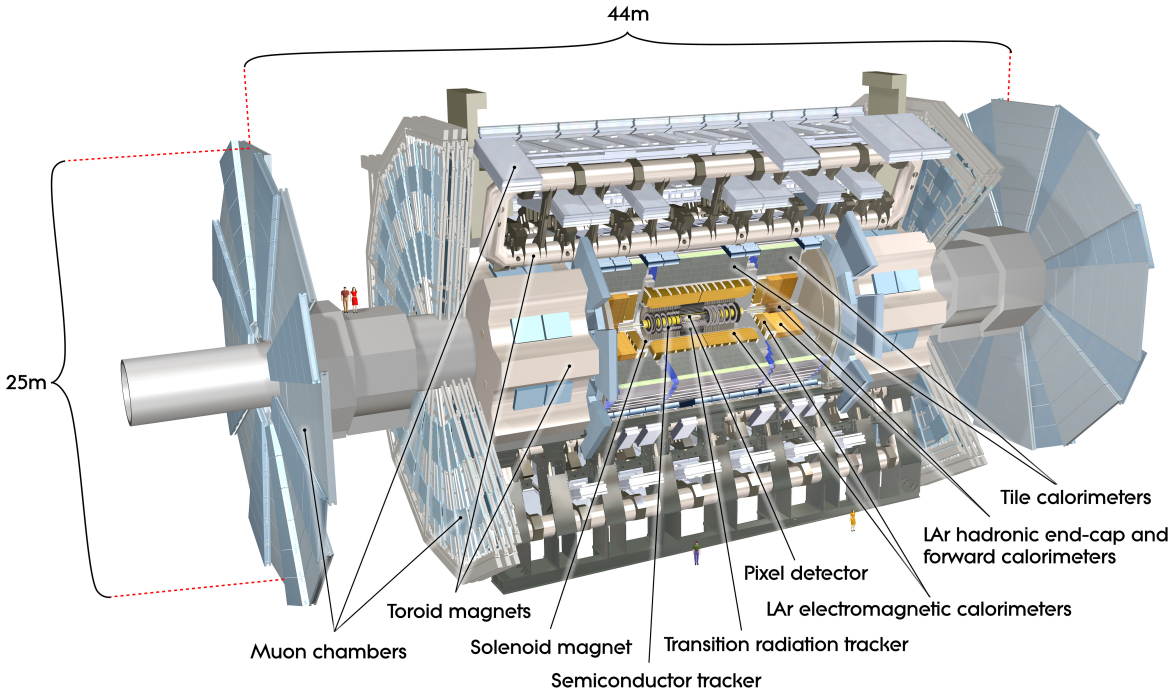


Figure 3.2 Computer generated image of the whole ATLAS detector, identifying different parts of the detector and two people are shown for scale [35].

3.2.1 Coordinate Systems

Given the cylindrical shape of the ATLAS detector, we use both, a right-handed Cartesian coordinate (x, y, z) and a cylindrical polar coordinate (r, θ, ϕ). The interaction point, which is defined as the geometrical centre of the detector and the origin of the coordinate systems. The positive x-direction of the defined Cartesian coordinates, points towards the centre of the LHC ring. The positive y-direction points upwards. The z-direction is a tangent to the LHC beam pipe and is positive in the counter-clockwise direction of the beam pipe.

Using the cylindrical polar coordinates, r is the distance from the beam line. The azimuthal angle, ϕ is the angle around the z-axis and has a $[\pi, -\pi]$ range. The polar angle θ is measured in the (x, z) plane and is the angle from the positive z-axis.

As we collide particles in the z-direction we use variables which are Lorentz invariant in this direction. We define the transverse momentum of a particle as $p_T = \sqrt{p_x^2 + p_y^2}$, which is an invariant quantity under boosts along the z-axis. The longitudinal momentum,

p_z, however, changes under such boosts and is therefore replaced with rapidity, which is defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{3.3}$$

where E is the energy of a particle. Rapidity itself is not Lorentz invariant but the differences in rapidity are invariant under boosts along the z-direction. if we consider a highly energetic massless particles, $E \gg m$, rapidity reduces to pseudorapidity, which gives us a Lorentz invariant quantity measure in terms of θ:

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \tag{3.4}$$

Pseudorapidity approaches infinity as the polar angle decreases to zero. Using this set of coordinates and variables we can define the distance between two objects in the (η, ϕ) plane as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{3.5}$$

3.2.2 Magnet System

The ATLAS detector has two superconducting magnet systems [36]: the first system is a central solenoid magnet [37] which creates a 2T magnetic field for the inner tracking detector. The second system uses a series of toriodal magnets [38]to create magnetic fields for the outer muon tracker (1T), one for each endcap region, and the barrel region (0.5T). The two magnet systems are shown in Figure 3.3.

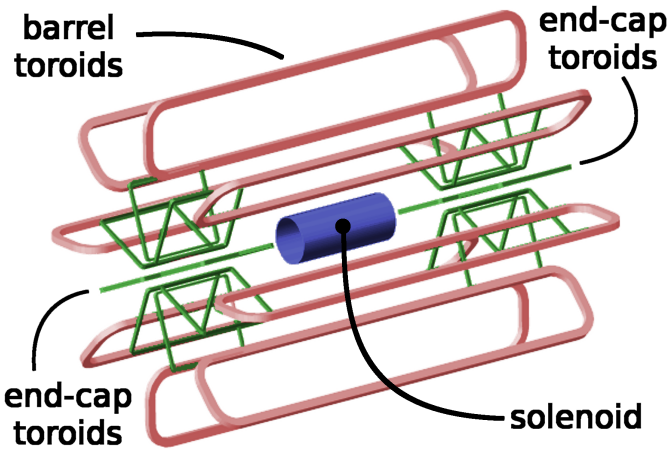


Figure 3.3 Computer generated image of the ATLAS magnet system [39].

3.2.3 Inner Detector

The Inner Detector [40, 41] is the section of the detector closet to the beam pipe. It is designed to register a series of points, or *hits*, along the path of a passing charged particle. These hits left by the particles are used to reconstruct the path, or track, of the particle. Using the magnetic field of the central solenoid magnet, the ID is able to record the direction and curvature of each track allowing it to determine the momentum and charge of the particle. The ID is split up into three main regions: The pixel detector, which sits at the innermost layer of the ID, the semiconductor tracker, and finally, the Transition radiation tracker. A cross section of the ID, showing the layout of the sub systems, is shown in Figure 3.4.

Pixel Detector

The Pixel Detector [42, 43] is the first point of contact for particles post collision. It is made up of four layers of silicon pixel sensors in the barrel region and three layers in the endcap regions. When the charged particles are incident on the silicon, electro-hole pairs are created which generate an electronic signal, this is then read out and recorded. The layer closest to the beam pipe is the Insertable B-Layer (IBL) [44, 45, 46], this is designed to improve the identification of secondary vertices associated with B-hadron decays, improving the quality of data from the remaining pixel layers. The IBL has the finest pixels with a size of $50 \times 250\,\mu m^2$ while the remaining pixel layers have a size of $50 \times 400\,\mu m^2$. There are a total of 92 million pixels which cover a total of 1.9 m^2, a full

azimuthal angle and $|\eta| < 3$. It achieves a resolution of 10 µm in the orthogonal direction of the beam pipe and a resolution of 115 µm parallel to the beam pipe.

Semi-Conductor Tracker

The Semi-Conductor Tracker (SCT) [47] is the main layer of the ID and surrounds the pixel detector with modules of long silicon strip sensors. The SCT modules are arranged in four layers in the barrel region ($|\eta| < 1.4$) and nine planar layers discs in the endcap regions ($1.4 < |\eta| < 2.5$). Each module contains two silicon strip sensors placed at a slight offset, in order to improve the spatial resolution. In total, the SCT contains 4,088 modules of double sided silicon strips resulting of 60m^2 of silicon. The SCT has a resolution of 17 µm in the $(r - \phi)$ plane and 580 µm in the z-direction.

Transition Radiation Tracker

The Transition Radiation Tracker (TRT) [48] is the final and largest layer of the ID, its used for particle tracking and pion-electron separation. The TRT is constructed using two main components: drift tubes with a diameter of 4 mm, filled with a mixture of Xe, O_2 and CO_2 gases, and gold-plated tungsten wire with a diameter of 30 µm which is stretched along the length of the tube. As a charged particle enters the drift tube, the gas is then ionised releasing electrons from the gas. These electrons then drift towards the anode wire creating a signal. The drift tubes are arranged in 73 barrel layers and 122 layers in each endcap giving an average of 30–40 position measurements per track. The TRT is also able to track particle in the $|\eta| < 2$ region with a spatial resolution of 130 µm in the $(r - \phi)$ plane which is notably worse than the silicon-based counterparts. However, because of the unique construction of the TRT, it is able to provide bespoke information in aiding to the discrimination of pions and electrons due to the differences in the amount of transition radiation generated by these particles. The TRT also improves the tracking performance of the ID, particularly at high transverse momentum due to many position measurements [49].

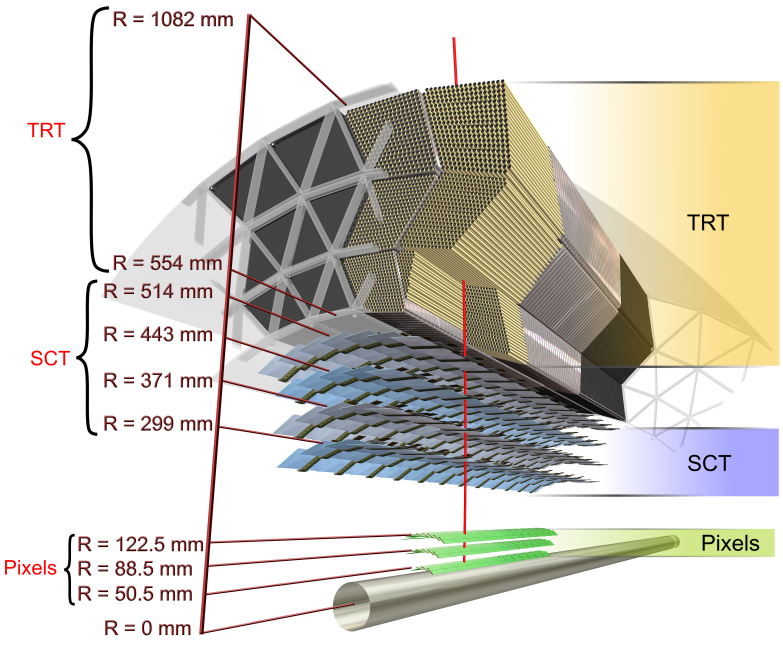


Figure 3.4 Computer generated image of the ATLAS Inner Detector [50].

3.2.4 Calorimeter

After the ID, the next main set of layers in the ATLAS detector are the calorimeters. The ATLAS calorimeters [51] are designed to measure the energy of charged and neutral particles, namely, electrons, photons and hadrons. These are called sampling calorimeters and are built in alternating layers of absorbing and active materials, noting that the energy is only measured in the active layers. As high energy particles interact with these layers, they lose energy by producing cascades of secondary lower-energy particles, called showers. Electrons and photons lose energy with the material via electromagnetic interactions. Electrons above 10 MeV lose energy via bremsstrahlung, while photons undergo electron-positron production. The formation of hadronic showers is similar to that of electromagnetic cascades but involving more complex cascades of hadronic (and electromagnetic) showers.

There are two properties that are fundamental to the design of the calorimeter systems: the radiation length, X_0, which is a characteristic of a material, related to the energy loss of high energy particles electromagnetically interacting with it. And for

hadronic interactions we consider the nuclear interaction length, λ_I, of a material. For a normal calorimeter material choice $\lambda_I >> X_0$. This means that electrons and photons interact much closer to the beam pipe than hadrons. So as a result, the ATLAS detector uses two calorimeters: an electromagnetic calorimeter and a hadronic calorimeter. A cross section showing the layout of the calorimeters is shown in Figure 3.5

Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECal) [52] uses liquid Argon (LAr) in its active layers and lead (Pb) in its absorbing layers. The lead layers are designed to cause showering, the resulting particles then enter a LAr layer where ionisation occurs, resulting in electronic signals. The ECal is designed to provide a consistent amount of material in all directions, corresponding to an interaction length of $22X_0$, covering a pseudo rapidity range of $|\eta| < 3.2$ and the full range of ϕ. The ECal is also split into the barrel region and endcap region. The barrel region starts with an active layer, called the pre-sampler. This is designed to measure the energy lost in the ID, after this there are then three finely granulated layers. In the endcap region there are only two, coarser, layers. The resolution of the ECal is:

$$\frac{\sigma_E}{E} = 10\% \frac{1}{\sqrt{E}} \oplus 0.7\%, \tag{3.6}$$

where the first term is the stochastic term which dominates at low energies and the second term is the constant term which dominates at high energies.

Hadronic Calorimeter

Following the ECal we have the Hadronic Calorimeter (HCal). The HCal uses steel in the barrel region and copper in the endcap region, as the absorber layers, and scintillating tiles [53] in the barrel region and LAr in the endcap for the active layers. The resolution of the HCal is:

$$\frac{\sigma_E}{E} = 50\% \frac{1}{\sqrt{E}} \oplus 3\%, \tag{3.7}$$

Forward Calorimeter

The Forward Calorimeter (FCal) is used to increase the precision of measuring missing transverse energy and naturally covers a higher range of pseudo-rapidity, $3.2 < |\eta| < 4.9$, compared to the barrel or endcap regions. The sampling is done using LAr for the active layers and copper and tungsten for absorbing layers. The FCal has three layers the first is designed to sample electromagnetic showers and the other tow are designed to sample hadronic showers. The resolution of the FCal is:

$$\frac{\sigma_E}{E} = 100\% \frac{1}{\sqrt{E}} \oplus 10\%. \tag{3.8}$$

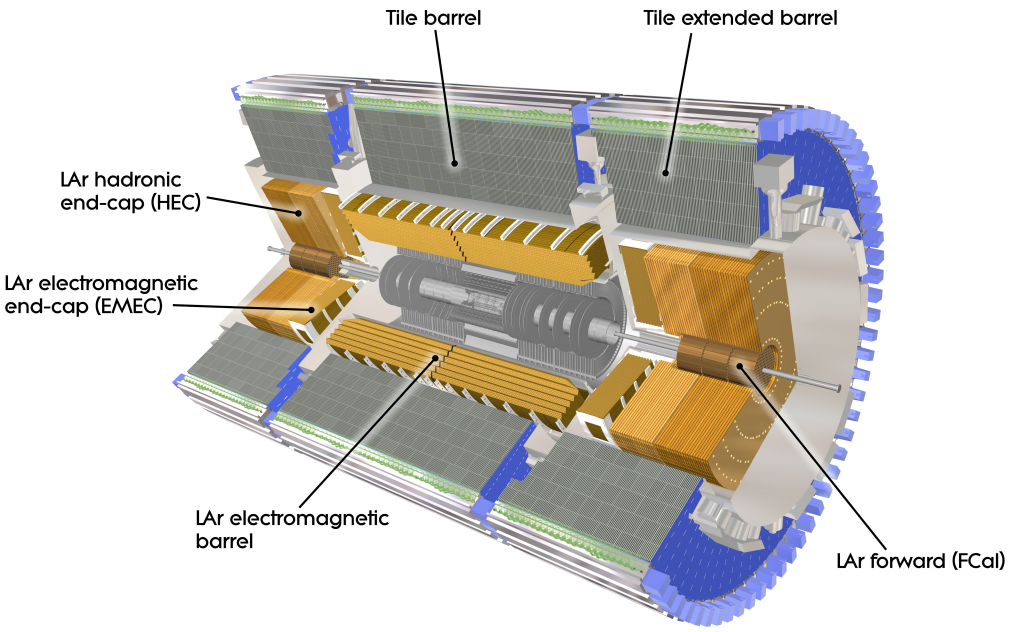


Figure 3.5 Computer generated image of the ATLAS calorimeters [54].

3.2.5 Muon Spectrometer

The final, outer-most, sub-detector system of the ATLAS detector is the Muon Spectrometer (MS) [55], as muons take the longest time to decay and don't interact with the calorimeters. The MS is a tracking detector designed solely for the purpose of measuring the tracks of muons. The design of MS relies on the notion that the path of the muon

bends in a magnetic field, and so the MS is embedded in a series of barrel and end cap torodial magnets, described in Section 3.2.2. The MS consists of four main systems: The Monitored Drift tubes (MDTs), the Cathode Strip Chambers (CSCs), Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs). The combination of these technologies allows for this allows for the precision measurement of muon tracks over a wide range of p_T and serve as a hardware level trigger. Figure 3.6 shows the MS in the ATLAS detector and Figure 3.7 is a schematic picture showing a quarter-section of the MS in a plane containing the beam axis, with MDTs, CSCs chambers for momentum determination and RPCs and thin gap TGCs chambers for triggering.

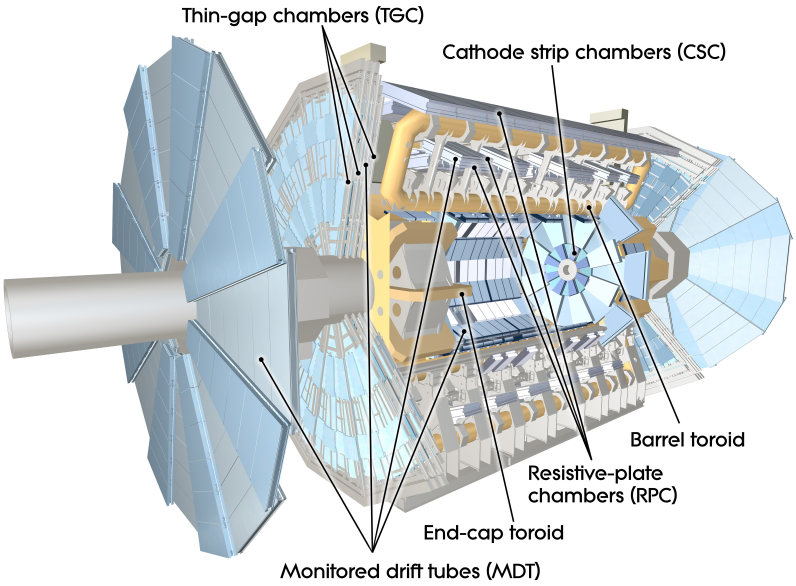


Figure 3.6 Computer generated image of the ATLAS muon spectrometer [56].

Monitored Drift Tubes

The MDTs [57] are cylindrical tubes filled with a mixture of Ar and CO_2 gases. The tubes contain a a singular tungsten wire running down its centre. As a muon enters a tube it ionises the gaseous mixture resulting in electrons drifting towards the anode wire and the ionised gas drifting to the cathode tube resulting in a signal. The MDTs are used in both the barrel and endcap region, with each region containing either three or four layers of MDTs each, and covers a range of $|\eta| < 2.7$. The MDTs have excellent spatial resolution, combined with the CSCs they allow for the precision muon tracking and momentum determination.

Cathode Strip Chambers

The CSCs [58] are series of cathode strips sandwiching a series of anode wires which exist in a chamber of Ar and CO_2. The orientation of the layers of the cathode strips are flipped as you descend through the layers. Starting from the outer layer, the cathode strip plate is parallel to the wires to provide η measurements and the subsequent layer is perpendicular to the anode wires and provides ϕ information. There are four layers in total. The CSCs rely on the same ionisation mechanisms as the MDTs to provide readout information.

The CSCs are located in the at the beginning of the endcap region of the MS and cover the high pseudorapidity range of $2.0 < |\eta| < 2.7$. The CSCs can handle a much higher sustained flux compared to the MDTs, 1000 Hz/cm^2 vs 150 Hz/cm^2 respectively. Hence the CSCs, compliment the information provided by the MDTs.

Resistive-Plate Chambers

The RPCs [59] are gaseous detectors that act as the one of the main components in the MS hardware trigger response. The RPCs are constructed by filling the gap of two parallel plates with a mixture of $C_2H_2F_4$, Iso-C_4H_1O and SF_6 gases. Similarly to the MDTs the incident muon ionises the gas, but here an electron avalanche is caused, providing a large unmistakable signal and an immediate response. The first two layers are used for low p_T muons triggers, while the third is used for high p_T muons. The RCPs cover a range of $|\eta| < 1.05$. The fast response of the RCPs allows for a trigger response of < 3 ns but a spatial resolution of a 10 mm.

Thin-Gap Chambers

The TGCs [60] are mainly employed in the endcap region and perform a similar trigger job to the RCPs. The TGCs have a similar construction to the CSCs, with more wires per chamber. The TGCs allow for the MS trigger functionality to be extended to $|\eta| = 2.4$ as it provides a similar trigger response time with improved spatial resolution.

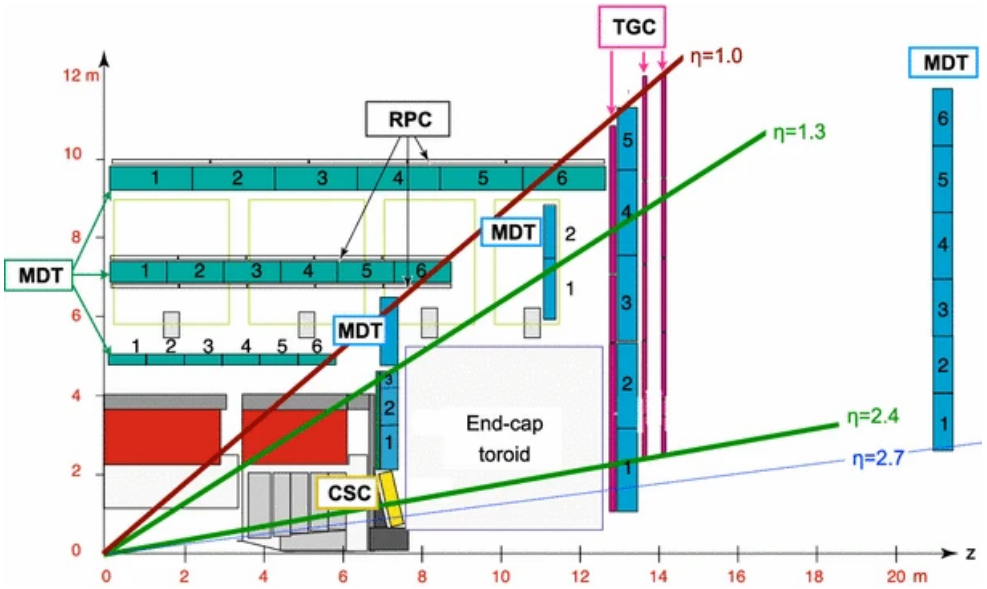


Figure 3.7 A schematic picture showing a quarter of a cross-section of the muon system within the ATLAS detector [61].

3.3 Trigger and Data Acquisition System

If one were to capture all the information in the events provided by the LHC (a bunch crossing every 25 ns resulting in an event rate of 40 MHz), one would need a way of handling 60 TB/s of data. This is too much data to handle, most of which is probably not very interesting. Alongside the ATLAS detector systems, software systems are needed to manage the volume of data being generated. The Trigger and Data Acquisition (TDAQ) system was designed to solve this problem, which employs a two-level data filtering system, one a hardware level trigger which feeds into a software level trigger, reducing the overall data read out to a more manageable 2GB/s. The resulting data is stored for physics analysis.

3.3.1 Level-1 Trigger

The Level-1 (L1) trigger [62] uses the information from the calorimeter and the MS's RPCs and TGCs to perform a coarse analysis of the event and process it in less than 2.5 µs. At this stage we do not identify the particle only that the event contains a high-p_T lepton, jet or high missing transverse energy, which would make the event potentially interesting and so is stored for later analysis by the software trigger. The region where these objects are located are called regions of interest (ROIs). This step reduces our event

rate to 100 kHz. The L1 trigger runs in real time with the collision and is considered ónline'.

3.3.2 High-Level Trigger

The High-Level Trigger (HLT) [63] is a software based trigger. It looks into the stored events and ROIs passed to it by the L1 trigger, and uses a series of algorithms to provide initial rejection and physics object reconstruction, further reducing the event rate to 1 kHz. The HLT runs after the collisions have occurred and so runs 'offline'.

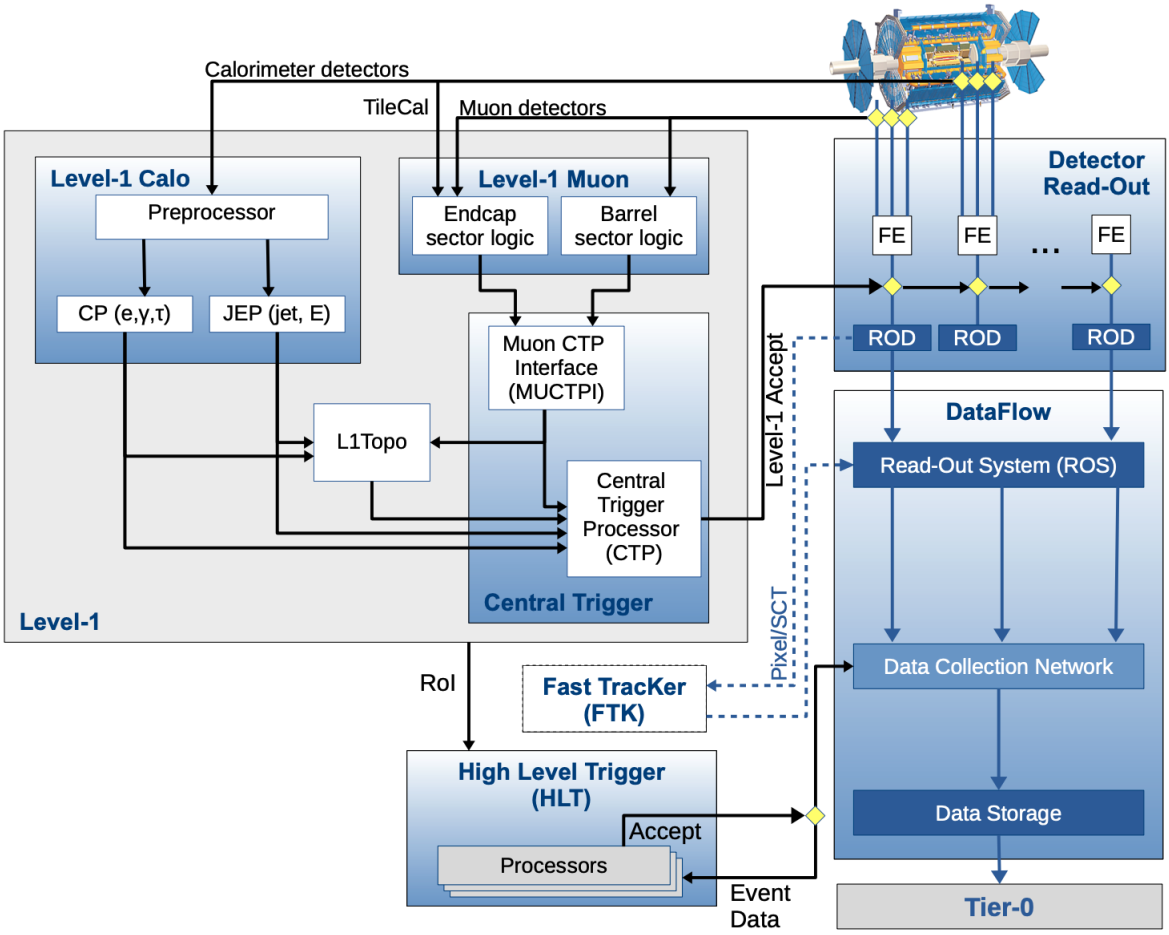


Figure 3.8 : The ATLAS TDAQ system in Run 2 showing the components relevant for triggering as well as the detector read-out and data flow [64].

3.4 Run-2 Data Taking Conditions

Figure 3.9 shows the cumulative integrated luminosity delivered by the LHC and recorded by the ATLAS detector during Run-2 (2015-2018) at $\sqrt{s} = 13$ TeV, together with the fraction of data certified as good for physics analyses. Over the full Run-2 period, a total of 156 fb^{-1} was delivered by the LHC, of which 147 fb^{-1} was successfully recorded by ATLAS and 139 fb^{-1} passed all detector and data-quality requirements. The close agreement between the recorded and good-for-physics luminosities demonstrates the high operational efficiency and stability of the ATLAS detector throughout Run-2, despite the increasingly challenging data-taking conditions associated with rising instantaneous luminosity and pile-up. The good-for-physics luminosity defines the dataset used in this analysis.

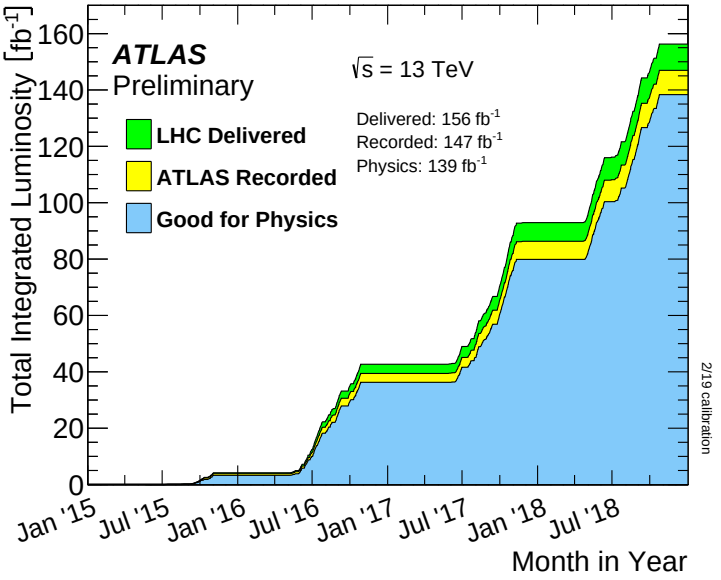


Figure 3.9 Cumulative integrated luminosity as a function of time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified as good for physics analyses (blue) during stable-beam periods for *pp* collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV in 2015-2018. The delivered luminosity corresponds to the luminosity provided by the LHC from the start of stable beams until ATLAS is requested to enter a safe standby mode for beam dumps or beam studies. The recorded luminosity accounts for data-acquisition inefficiencies, including the so-called warm-start period during which detector high voltages and front-end electronics are ramped up after the stable-beam flag is raised. The good-for-physics luminosity reflects the application of the ATLAS *All Good* data-quality criteria, which require all relevant detector subsystems and reconstructed physics objects to be fully operational.

Chapter 4

The ITk Upgrade and Qualification Task

"Nah man, objects are stupid. Ya just need hunners of scripts."

— Dr. Kenneth Wraight, a rebel without a cause so he rebels against object orientation.

The LHC, and the ATLAS detector, is currently in its Run-3 data taking stage which will provide more data for our measurements and analyse and providing more statistics for combination measurements across the data taking stages. The delivered luminosity and mean crossings per bunch are higher than Run-2, these numbers are dwarfed when compared to Run-4, where the LHC will enter its High-Luminosity era (HL-LHC). It is anticipated to deliver an integrated luminosity of 3000 fb^{-1} over ten years of operation (2029-2039). With the increased luminosity, there will also be increased pile-up, readout rates and radiation. This all means that we will be operating in a noisier environment. Major reforms to the detector need to take place in order to benefit fully from the HL-LHC era. The main upgrade for this time will be a new inner detector system, which will be replaced by an all silicon inner tracker (ITk).

4.1 The ITk

The ITk is designed to replace the current ID for the HL-LHC era in Run-4. It will be an all silicon solution, that comprises a pixel and a strip subsystem [65]. The pixel

subsystem covers a pseudorapidity range of $|\eta| < 4.0$ and consists of five flat barrel layers and five layers of inclined or vertical rings for forward region coverage. The strip subsystem spans $|\eta| < 2.7$ and includes four strip layers in the barrel region and six disks in the endcaps, all using double-sided modules. The layout of these subsystems can be shown in Figure 4.1. The resulting upgrade is expected to increase the trigger rate from 100 kHz to 1 MHz [66]. Improved radiation hardness of the subsystems will allow for the ITk to perform better in the high radiation environment of the HL-LHC. The higher granularity of the subsystems will improve the spatial resolution over the current ID and improve the vertex reconstruction, enhancing the b-tagging and pile-up rejection.

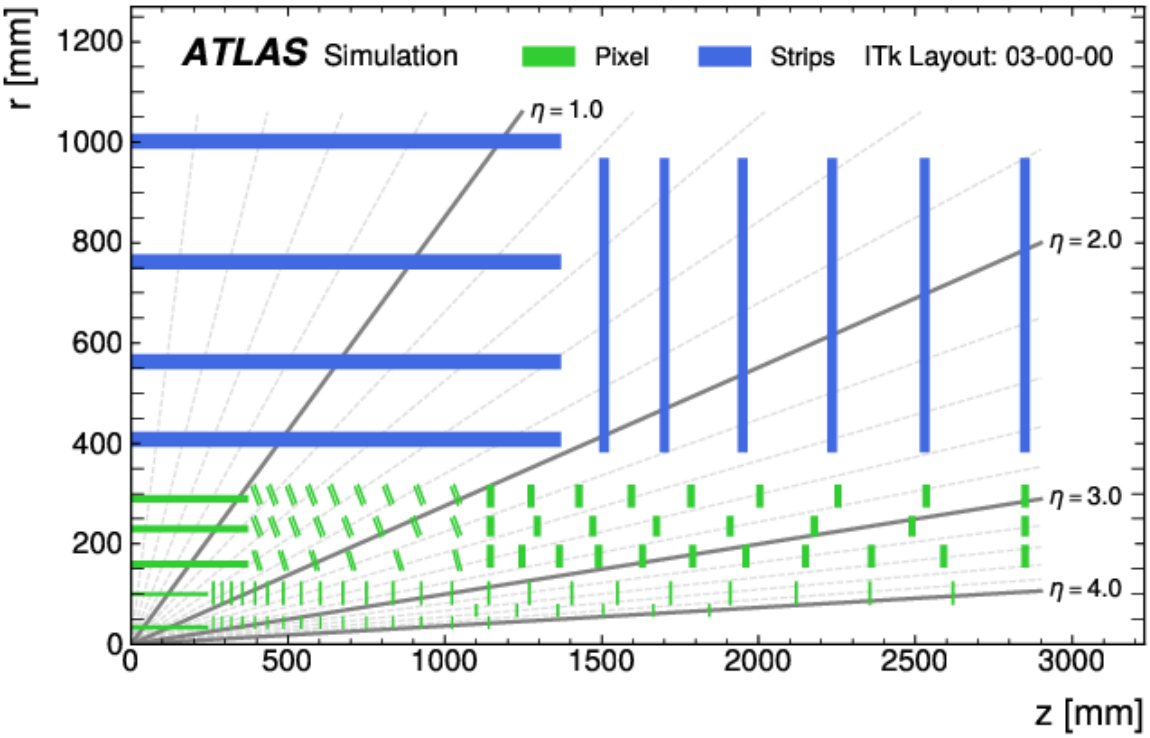


Figure 4.1 A Schematic view of the ITk detector where the strip detectors are shown in blue and the pixel detector is shown in green [65].

4.1.1 Pixels

The ITk pixel subsystem upgrade will consist of around 9400 modules with 5 billion pixels covering an area of around 13 m^2. The Pixel subsystem layout is shown in Figure 4.2 and will be separated into three main zones the Inner System, Outer Rings and Outer Barrel. The Inner System consists of the Barrel with flat staves and Ring layers. The Pixel subsystem will use two main sensor technologies: Planar and 3D.

Planar technology is currently used in the pixel subsystem of the ID. For the ITk they have been redesigned in order to make them thinner and use n-in-p doping. There are many advantages for making the planar sensors thinner: we improve the radiation hardness, reduce the leakage current which helps with noise suppression, Improved response time, better track resolution, lower power consumption and reduce the amount of material (number of radiation lengths) in front of the calorimeter [67]. Planar sensors are incorporated in the design and will be used in the outer rings and outer barrel regions.

3D sensors are a type of silicon detector where electrodes (columns) are etched into the silicon vertically, instead of being placed only on the surface like in planar sensors. This design allows for charge collection to happen in three dimensions, hence the name. As a result of the design, the 3D sensors have some unique benefits: the spacing of the electrodes are quite small resulting in shorter drift distances which reduces charge trapping. The charge collection time is also very short, improving time resolution and helping with pile up rejection [68]. The 3D sensors have improved radiation hardness compared to planar sensors making them ideal candidates for the inner-most part of the ITk, the Inner System.

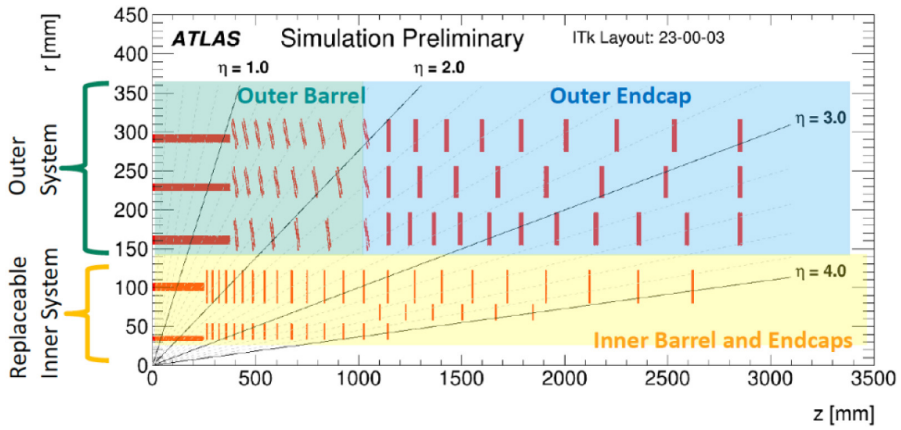


Figure 4.2 The layout of the pixel detector. The horizontal axis is along the beam line with zero being the nominal interaction point. The vertical axis is the radius measured from the interaction point [65].

4.1.2 Strips

Surrounding the Pixel subsystem is the Strips subsystem, it is split into barrel and endcap regions. The barrel is made from four concentric layers, where the two layers closest to the collision point are made from short strip segments while the outer two layers are made from long strip sensors.

4.2 Dual Use

The main components for the ITk: ASICs and FPGAs and all assemblies with these components in them are classified as Dual Use, meaning that they have both civilian and military use. After the invasion of Ukraine in 2022, FPGAs manufactured for US companies were found in Russian missiles [69]. Following this discovery the Biden administration imposed strict regulations on FPGAs and other high performance chips in order to stop this from happening. As a result of these restrictions every component or assembly must be exported from CERN and returned to CERN for final assembly or to be destroyed even if the component fails a test. Given the high stakes of FPGAs and the like going 'missing' we need to make sure all components are accounted for. This raises a new problem: who has what and what did they do with it? Since there are 50+ institutes, which are broken up into several regions, all contributing to this upgrade, the flow of components form one institute to an other starts to balloon in complexity. Take for instance the UK-CHINA cluster shown in Figure 4.3. The flow of components isn't a

straight forward assembly line and is in fact a spider's web of possibilities for components to get lost in production as so we need a robust reporting tool in order to keep track of production rates and spot when and where components might go missing.

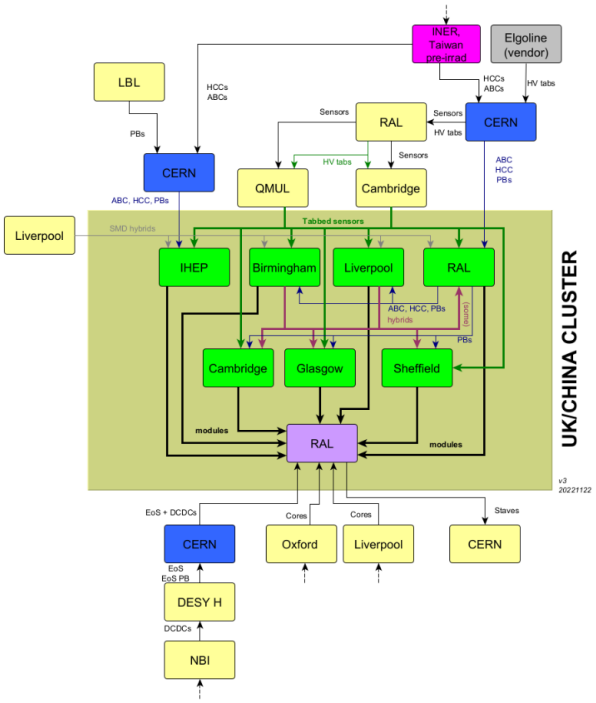


Figure 4.3 A production flow diagram for the UK-CHINA cluster for the manufacture, testing and assembly of parts for the ITk upgrade. All components are given to CERN from the manufacturing vendors, they are then added to inventory for bookkeeping. The relevant components are then give to RAL where they are distributed amongst the UK institutes contributing to the upgrade for testing and assembly. The components, if assembled or deemed unfit, they are then returned to RAL where they are assembled into larger components and returned to CERN for final assembly be destroyed.

4.3 ITk Web-App

For my Qualification Task (QT) I was tasked with creating a tool in order to track the production rates of various components for the ITk upgrade, the components will undergo various tests before they are assembled into larger modules to make it into the ITk. The official description of my task was as follows:

> The task will perform a regular query of the ATLAS ITk Production Database to understand the parts (e.g. modules) in different sites and compare to the production model. This information will then be used to guide the activity coordinators and the production managers in the decisions around part flow. For example, we will be able to see the number of modules in the different production sites and the need for modules at the different loading sites and compare to plan and act upon variance. A proposed weekly (frequency to be defined) job will perform the query, tabulate the data, and generate standard plots for the users. Access will be via a web-app. The tabulated data will be stored so that a history file will be created to allow time dependent analysis to be performed to understand groups production rates as well as part count on a given week. The tool and its use will be presented at the DB meeting. This task is useful both for Pixel and Strip.

For the remainder of this chapter I shall focus on the UK-CHINA cluster and strips production. However, the web-app frame-work can be used for global (several clusters), individual cluster and local institution monitoring for both strips and pixels.

4.3.1 Data Pipelines

Information about the location of the components and at which stage of testing or assembly they are in is contained in the ITk Production Database (PD). The PD however does not provide timestamping, which is a crucial piece of information in order to produce the production timelines. In order to combat this short-fall we need to timestamp these components ourselves. Timestamping the individual components would result in a lot of data points to manage as there could be thousands of the same components at the same testing or assembly stage. To reduce the number of data points, we consider the population of the components per stage. This will allow us to visualise the time-line of a batch of components going through the various stages of testing or assembly, and if any components fail the assembly or testing stages, thus deduce the yield and rate of production.

Influxdb [70] was used to create timestamps of component stage populations. Influxdb is a third party timestamping data-base (db), as data is fed into the db, it is then timestamped at the time of entry. Once the data is time-stamped it can be readout using the inbuilt API. From here we can transform the raw timestamped data into our desired

format and then plot. For plotting, we used Altair a python based graphing package that offered the interactive functionality we needed to make a user friendly experience [71].

For this stage of development various versions of the interactive plot objects were prototyped and refined with feedback from institutes in the UK-China cluster. This refinement allowed for key performance indicators to be visualised within the plots. The finalised plot objects were then embedded in a web-app hosted at CERN, allowing for all the institutes in the UK-CHINA cluster to access not only their own production time lines but also each others, meaning that all the production timeline information for each type of strips related component at all stages in all institutes were concatenated in one place. RedHat was used for hosting and resource management, which pulled a series of Docker configs for containerisation. The Docker configs were stored securely on CERN's servers. Once the container was spun up, the web-app is able to be hosted and the landing page can accessed by anyone on the web. However, the information can only be accessed by users of the PD as they save passwords are required. A flowchart depicting the data pipeline for the web-app is show in Figure 4.4.

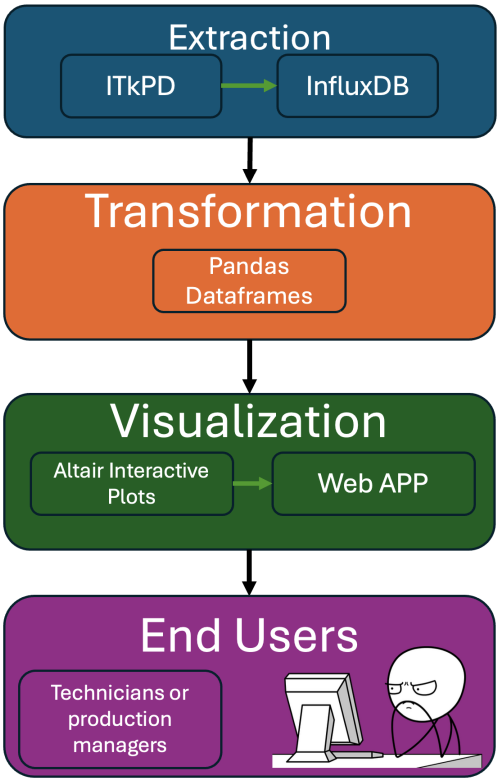


Figure 4.4 A simplified diagram to show the flow of data from the ITkPD to the end user. Data is extracted from the ITkPD then time stamped in the influxDB. The time stamped data is then transformed with a series of functions in order to extract useful insights, this is then sent to the front end where the interactive plots are made and embedded in the web-app where the end user can access them.

4.3.2 Example Interactive Plots

In Figure 4.5 one of the early interactive plot prototypes is shown. At the beginning of this project the ITk production was still in its relative infancy so only very technical institutes could contribute to the production and testing of components. One of these institutes was the Rutherford Appleton Laboratory (RAL). The large volume of components at RAL allowed for it to provide an ideal data-set for the visualisation of components populations and the stage migration of these components. We see that in Figure 4.5 we are only plotting modules and the components associated with them. This plot was made with an early build of the framework described in Figure 4.4, even still we see that we managed to capture and visualise some batches of components transitioning through stages. For example we see various interplay between Glued (red) and Bonded (purple) as once the modules are glued to the stave they can be considered bonded.

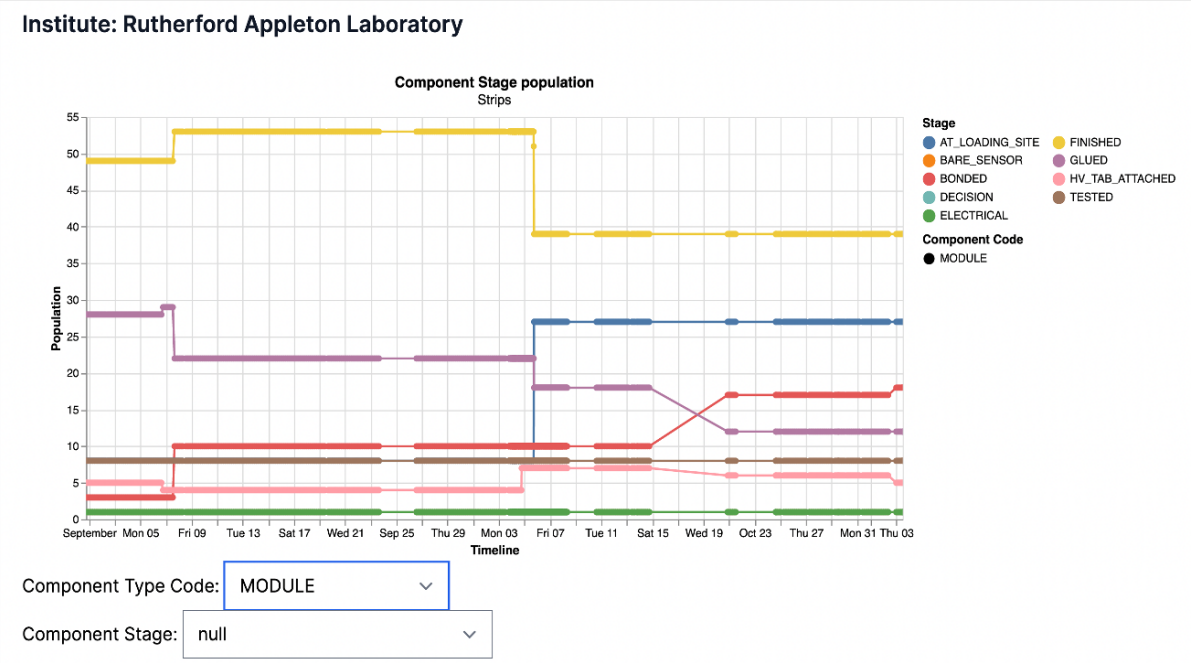


Figure 4.5 A prototype interactive plot depicting the population of components in each stage of testing or assembly at the Rutherford Appleton Laboratory.

This plot was then modified to include more component types, their stages and if the components had been trashed or not, as showed in Figure 4.6.

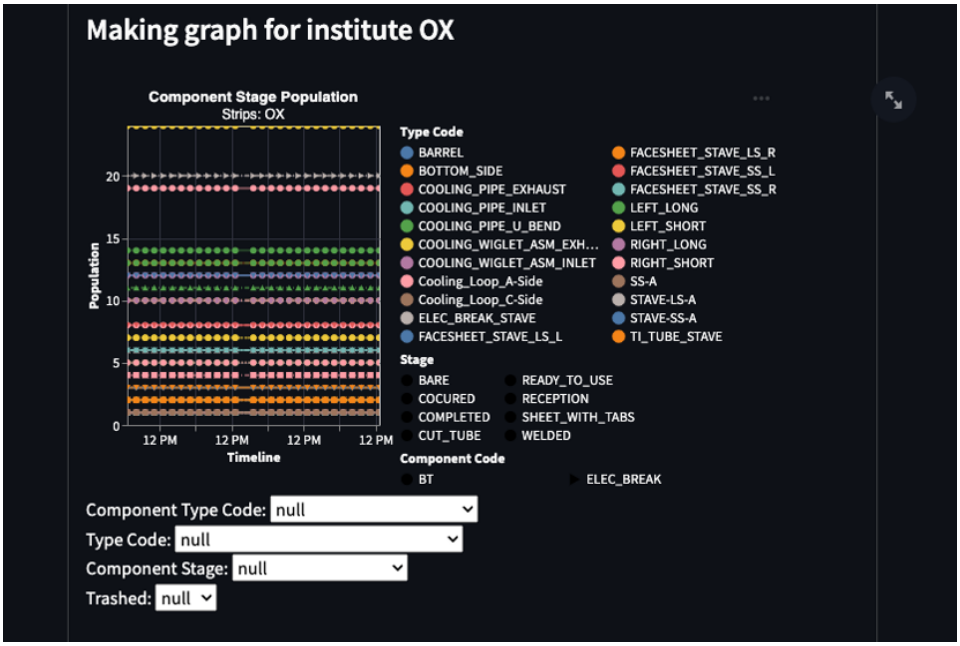


Figure 4.6 The finalised front end of the web-app showing a final version interactive plot with various selection criteria for Oxford.

4.4 Discussion

With the tool now deployed and the web-app is available to anyone in the ITk upgrade group so they can see the production rates of all components across any institute. As mentioned earlier in this chapter this tool will help spot missing components but can also help spot bottle necks in production allowing production managers to address these issues and ensure components make it to their next stage or institute at a timely manner, thus making sure the ITk keeps to its production deadline.

Smaller versions of this tool have been retrofitted to provide, local, in-institute monitoring namely in TRIUMF labs. With the completion of this task I became a qualified ATLAS author.

Chapter 5

Object Definitions

"Electrons are green."

— Dr. Jonathan Jamieson, after years of research.

Following the trigger and data acquisition stages, the collected data is used to construct physics objects. From the side of the experiment we only see detector signals created by the passing of particles through various sub-detector systems. We use these signals to reconstruct and identify the parent particle, thus deducing the physics at play in the event. The sub-detector systems can be split up into two groups: tracking and calorimeter information. Tracking information is provided by the ID and MS, while the calorimeter information is provided by the ATLAS calorimeter system. The raw data these systems provide is used in various dedicated reconstruction algorithms, which aim to construct physics objects. The main physics objects that we will discuss in this chapter are: particle tracks, Primary Vertex (PV), jets and E_T^{miss}.

5.1 Tracks, Clusters and Vertices

Nearly all of the other reconstruction algorithms require track reconstruction as an prerequisite as they rely on this process as an input. The main tracking algorithm, illustrated in Figure 5.2 uses information from all sections of the ID, starting from the PD and SCT then using the TRT [72, 73]. As a charged particle moves through the layers of the ID it deposits energy, these deposits (hits) are then organised into clusters using a Connected Component Analysis [74]. The clusters are then used to determine

three dimension space points of where the particle pass through the detector layers. This is done in two ways: for the PD each cluster corresponds to one space point, while in the SCT a combination of both sides of the strip layer are used to determine space points, which are then used to create track seeds. Track candidates are extrapolated from these track seeds using a combinatorial Kalman filter [75].

Using the spatial data from the remaining layers of the pixel and silicon detectors, provided they align with the initial track trajectory, results in a notably higher reconstruction efficiency for the primary particle tracks, eliminating a substantial number of tracks formed solely from random groupings of space points. The track reconstruction efficiencies as a function of η and p_T are shown in Figure 5.1. Once all realistic combinations of space-points have been identified, multiple track candidates can occur due to overlapping or incorrectly assigned space-points. This ambiguity is resolved by ranking the track candidates using a so-called track score [76]. After passing these criteria, hits

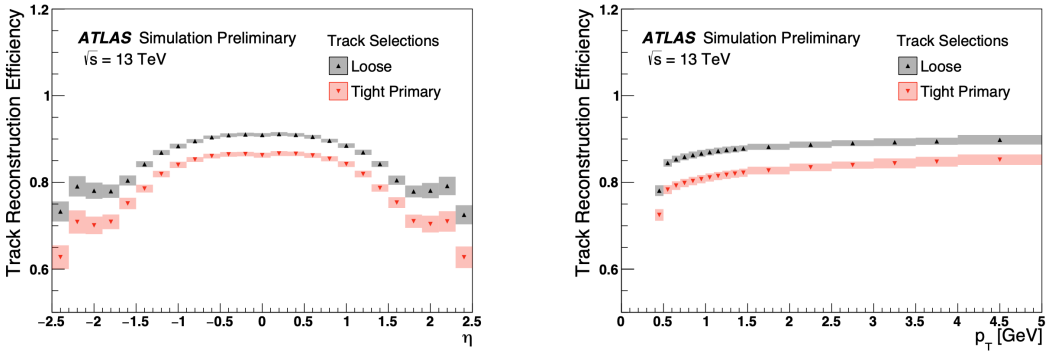


Figure 5.1 Track selection efficiency as a function of truth-level η and p_T for the Loose and Tight primary track selection criteria, evaluated using simulated samples selected with a minimum bias trigger. The shaded band indicates the total uncertainty on the efficiency estimate, with the dominant contribution arising from uncertainties in the known material distribution within the Inner Detector.

in the TRT are identified to further extend the track. The final stage of this inside-out approach is a global χ^2 fit to accurately determine the track parameters $(d_0, z_0, \phi, \theta, q/p)$ and the proton-proton interaction reference point. Here d_0 is the transverse impact parameter and defines the distance of closest approach to the beamline in the $(r - \phi)$ plane. z_0 is the longitudinal impact parameter and is the distance between the closest point of approach and the reference point, with ϕ and θ being the azimuthal angle and polar angle, respectively. Finally, q/p is the ratio of charge and momentum. In addition to inside out tracking we also employ an outside-in method. Here, we start from the

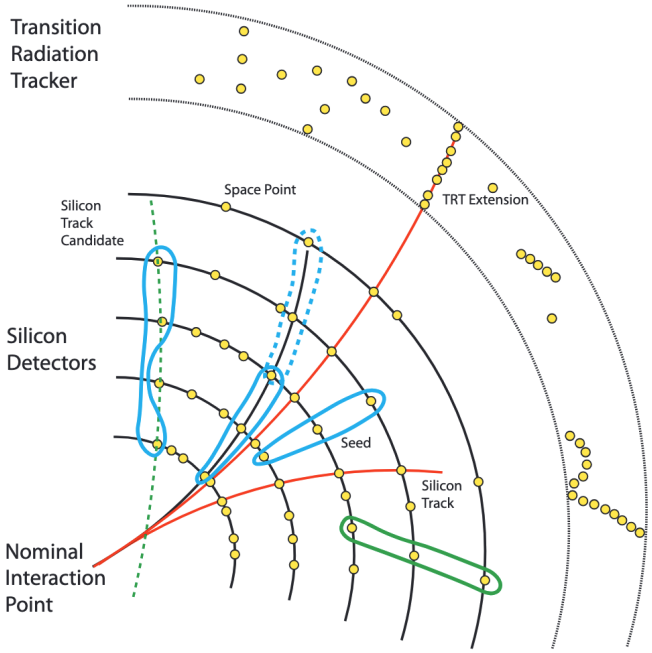


Figure 5.2 This illustration shows the track reconstruction process using pattern recognition. Track seeds (blue) are formed from triplets of space points (yellow), with each point coming from a different detector layer in the ID. Seeds associated with the same track candidate are connected by a dashed blue line. Green-coloured seeds represent those that fail to meet the requirement of originating from a charged particle at the primary interaction vertex. Tracks rejected for the same reason are indicated with dashed green lines. A fully reconstructed track using only silicon hits, without a TRT extension, is shown in black. A track that includes a TRT extension is depicted in red [77].

TRT and work inwards towards the PD and SCT. This method is particularly effective in identifying tracks from secondary decays. The TRT tracks are disregarded if they do not have extensions to the PD or SCT.

PVs are now reconstructed using the tracks [78], there are two main steps for this reconstruction of vertices: finding and fitting. Once the full set of tracks are identified we then project inwards to find a shared vertex point, if the tracks don't fit this point they are used to find another vertex point. with all the tracks for a particular vertex in hand an iterative χ^2 minimisation fit is applied to accurately determine the vertex point's position. We assume that charged particles that are the result of hard scattering interactions have larger transverse momentum than those that are classified as pile up and so the the primary vertex is the vertex which has the highest sum of squared transverse momentum From all the associated tracks.

5.2 Electrons

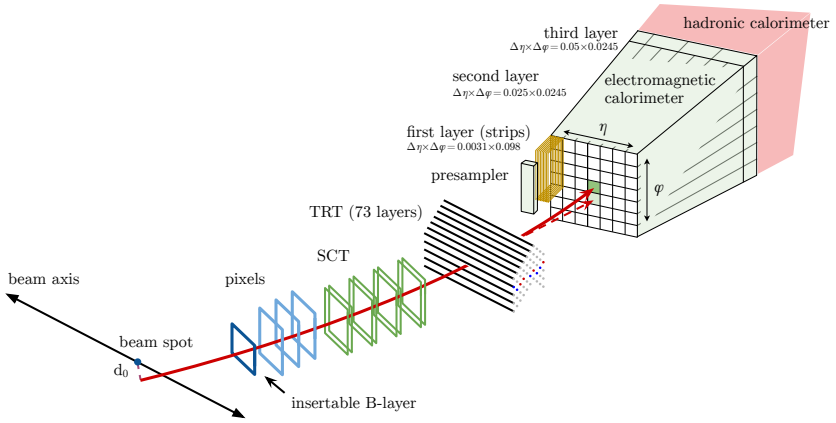


Figure 5.3 The path of an electron through the detector. The electron passes through the tracking system of the inner detector before depositing its energy in the EM calorimeter [79].

Electrons are reconstructed using information from the ECal and ID. Electromagnetic showers in the ECal are reconstructed using a topological clustering algorithm [80]. The algorithm builds variable-size clusters dynamically and is able to recover the low energy photons emitted from bremsstrahlung ($\mathcal{O}(100)$MeV). The clustering algorithm utilises a cell significance variable:

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{|E|_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}, \tag{5.1}$$

where $|E|_{\text{cell}}^{\text{EM}}$ is the absolute cell energy at the EM scale and $\sigma_{\text{noise,cell}}^{\text{EM}}$ is the expected cell noise. The goal of this stage is to form topological clusters (topo-clusters). Initially cells with $\zeta_{\text{cell}}^{\text{EM}} \geq 4$ are selected to seed proto-clusters. These proto-clusters then grow by progressively collecting neighbouring cells with $\zeta_{\text{cell}}^{\text{EM}} \geq 2$ until no neighbouring candidates remain. The cell significance condition is then relaxed to $\zeta_{\text{cell}}^{\text{EM}} \geq 0$ to collect all remaining cells with energy deposits. The outputs of this procedure are the topo-clusters. For the group of topo-clusters, the cluster with the highest p_T, with a minimum 1 GeV energy and with matching ID tracks, with at least four hits, is then selected as a seed in a super-cluster. The remaining topo-clusters are examined for classification as a satellite cluster of the super cluster, which can be see in Figure 5.4. Each super-cluster with matching tracks are considered candidates for electrons, where the energy of the electron is given by the cluster energy. Following the reconstruction of the electron candidates, some energy calibration is required to improve the energy resolution of the

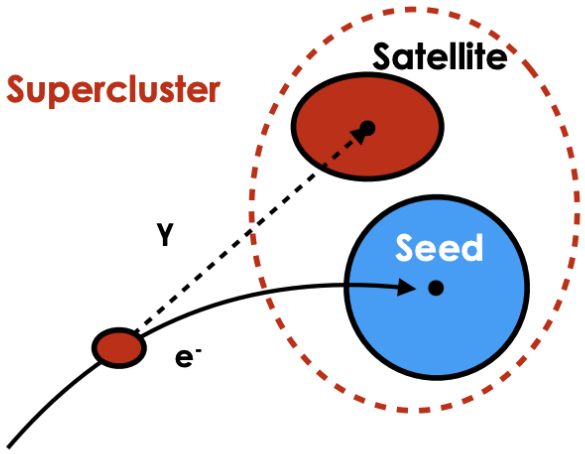


Figure 5.4 Diagram of an example supercluster showing a seed electron cluster and a satellite photon cluster [81].

candidates and correct for differences in the candidates' energy scale in data and MC. This calibration accounts for energy loss in the tracker material, energy deposition in neighbouring cells relative to the energy cluster, and general energy loss outside the ECal. The final correction is that of the energy scale of the candidate, where a data correction is derived using $Z \to e^+e^-$ and $J/\psi \to e^+e^-$. The calibrated electrons are then subject to a a series of cut-offs in their p_T, η, and the longitudinal and transverse impact parameters, in order to refine the set of candidates further. The following conditions must be met for a candidates to pass selection: $p_T > 7$ GeV, $|\eta_{\text{clust}}| < 2.47$ excluding the transition region $1.37 < |\eta_{\text{clust}}| < 1.52$, $|d_0/\sigma(d_0)| < 5.0$, and $|z_0 \sin\theta| < 0.5$ mm. Electrons are further classified using the identification quality criteria. The final step of classification involves a using a likelihood-based discriminant built from a fit to a combination of calorimeter and tracking information. Backgrounds to the identification consist of photon conversions, jets with electron-like features, and non-prompt electrons from hadron decays. Fixed values of the likelihood discriminant are used to define different Working Points (WP), each consisting of differing signal selection and background rejection efficiencies. These are aptly termed Loose, Medium and Tight. More details can be found in [79].

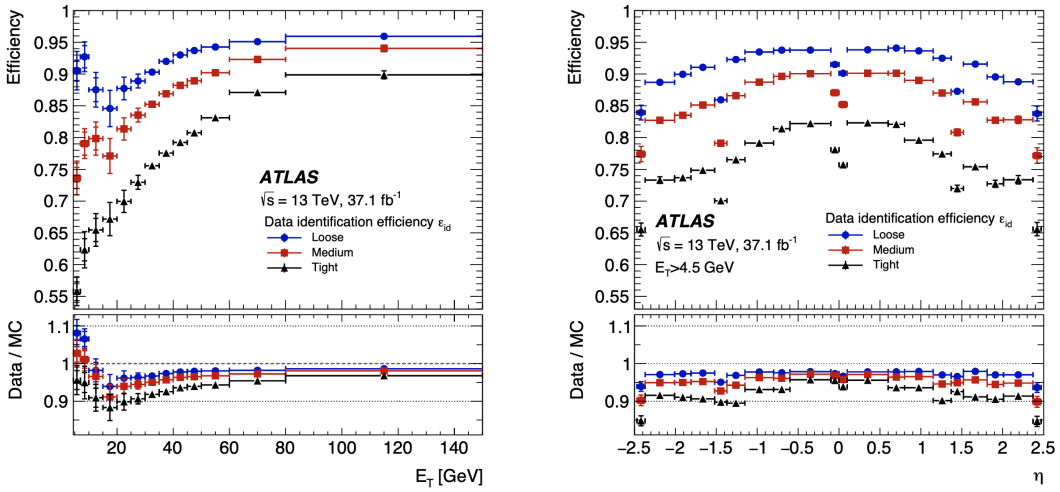


Figure 5.5 Electron identification efficiencies using the LH method are shown for the Loose (blue circles), Medium (red squares), and Tight (black triangles) operating points, measured in $Z \to e^+e^-$ events. The efficiencies are plotted as a function of η (right) and E_T (left). Vertical error bars indicate the statistical uncertainties (inner bars) and total uncertainties (outer bars) [79].

5.3 Muons

In the ATLAS detector, muons are unique because they can penetrate the outer layer of the calorimeter system while still being detectable, as they are minimum-ionizing particles. This distinctive property results in clear tracks within the ID and MS, along with characteristic energy deposits in the calorimeter. Muon reconstruction begins by using the hits in the individual layers of the MS. These hits are used to create track segments which are identified using a Hough transform and are used to build our initial track candidates [82]. The process leverages precision measurements of the track's bending radius and information from the trigger detectors. The process intends to find a straight line fit along the hits in a single layer of the MS.

A combinatorial search is used to build segments in the CSC in the $(\eta - \phi)$ plane. With the segments identified some refinement is required. A χ^2 fit is applied to the trajectory of the muon in the tarodial magnetic field in order to account for material interactions in the MS and possible effects of misalignments. Any outlier hits along the track trajectory are removed, and the fit is recalculated. If ambiguities occur they are resolved by removing tracks that share hits with higher-quality tracks. These final muon tracks are then refitted, taking into account the possible energy loss in the calorimeter

system. With this process in hand my we identify five different types of muons: combined, inside-out combined, segment-tagged, calorimeter-tagged, and MS extrapolated.

Combined muons are constructed by matching the tracks in the ID and MS with a fit based on their underlying hits in their respective sub-detector. The fit also takes into account potential energy loss from the calorimeter. Inside-out muons are reconstructed by extrapolating tracks from the ID into the MS. If at least three MS hits can be aligned with the track, the track is reconstructed as a muon. Segment-tagged muons are identified using a track in the ID that tightly matches in angular distance with at least one MS segment. Calorimeter-tagged muons do not rely on the MS. They are identified using an ID track with an energy deposition in the calorimeter system consistent with a minimum-ionising particle. Finally, a track is reconstructed as an MS-extrapolated muon if the track in the MS cannot be matched to an ID track. At this stage, similarly to electrons, the muon candidates are split up into WP points of tight medium and loose [83].

5.3.1 Prompt leptons and isolation

Prompt leptons are those produced directly in the primary hard-scattering interaction, typically from decays of heavy bosons like W, Z, or from top quarks. These leptons are generally clean and isolated from other particle activity. In contrast, non-prompt leptons arise from the decays of hadrons, especially those containing heavy-flavour quarks (such as b or c quarks) and are often embedded within jets. Fake leptons are not real leptons but are instead hadrons or other particles (like photons or jets) that are misidentified as leptons by the detector due to overlapping signatures or reconstruction errors. To distinguish prompt leptons from non-prompt and fake leptons, ATLAS uses lepton isolation techniques [84]. This involves measuring the amount of additional activity, such as tracks or calorimeter energy deposits in a cone around the lepton's direction. Prompt leptons, being isolated, are expected to have little surrounding activity, whereas non-prompt and fake leptons are usually accompanied by nearby particles from jets or hadronic decays. By applying isolation requirements either based on tracks, calorimeter deposits, or both leptons with too much nearby activity can be rejected.

5.4 Jets

Compared to leptons, partons require effort in order to identify as they rely upon a definition and aren't directly detectable. As the partons are created they begin to decay in a cascade of secondary charged and neutral particles. This decay is cone-like in shape and we call them jets. These jets are reconstructed using dedicated clustering algorithms, this clustering algorithm is required to be infrared and collinear safe (IRC). The result of the algorithm (number and direction of jets) should not change with the addition of soft gluons to the event (infrared safe) and the outcome should also be the same if the parent jet splits into two or more collinear child jets (collinear safe). The anti-k_T jet clustering algorithm satisfies the IRC prerequisites and is the main jet clustering algorithm used by ATLAS [85]. The algorithm groups clusters of energy deposits left by hadronic showers based on a distance between two entities i and j with distance between these two entities, d_{ij}, give by:

$$d_{ij} = \min\left(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,j}^2}\right)\frac{\Delta R_{ij}^2}{R^2}, \tag{5.2}$$

additionally with the distance between the beam and entity i, d_{iB}:

$$d_{iB} = \frac{1}{p_{T,i}}, \tag{5.3}$$

Where $\Delta R_{ij}^2 = (\phi_i - \phi_j)^2 + (y_i - y_j)^2$, where ϕ_i and ϕ_j are the azimuthal angles, and y_i, y_j are the rapidities of i and j. The transverse momenta of the particles, denoted $p_{T,i}$ and $p_{T,j}$, are also used in defining the clustering criteria. In practical applications such as the anti-k_t algorithm, an adjustable fixed-radius parameter R is introduced, which controls the typical size of jets. The anti-k_T algorithm reconstructs jets using topologically clustered energy deposits in the hadronic calorimeter. This topological clustering follows the same procedure as that used for electrons, as described in Section 5.2. A cluster is identified as a jet if its distance to the beam, d_{iB}, is smaller than all pairwise distances, d_{ij}. If this condition is not met, the algorithm merges cluster i with the cluster or pseudo-jet j that has the smallest value of d_{ij}. This merging process is repeated iteratively, combining topo-clusters and pseudo-jets, until all constituents are assembled into jets.

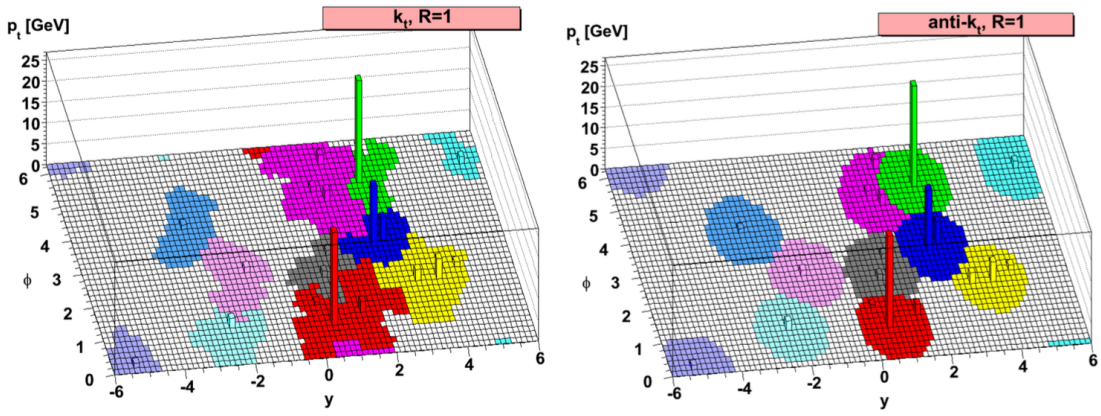


Figure 5.6 An example parton-level event generated with HERWIG [86] in the $y - \phi$ plane is shown. To this event, many random soft particles, or "ghosts," are added. The calorimeter cells and the deposited energy are depicted. Jets are reconstructed using the k_t algorithm (left) and the anti-k_t algorithm (right). Using a radius parameter of $R = 1$, one can observe the difference in jet shapes between the two algorithms. While the jet shape produced by the k_t algorithm is influenced by the soft particles, the anti-k_t algorithm yields jets whose shapes are largely determined by the hard particles [85].

5.4.1 Particle Flow (PFlow) jets

Traditional calorimeter only jet reconstruction often led to an overestimation of the jet energy. The Particle Flow (PFlow) algorithm and resulting object aims to reduce this overestimation by using track and calorimeter information in the object reconstruction process [87].

The process starts with the fact that as a charged particle passes through the detector it leaves tracks in the ID and topo-clusters in the calorimeter. The PFlow algorithm aims to match the tracks with topo-clusters, utilising the positional information of the calorimeter and the momentum resolution of the ID. The expected energy deposition of the same particle that generated the track in the calorimeter is estimated. Noting that a single particle can often create more than one topo-cluster, the algorithm checks if more than one topo-cluster needs to be added to the track topo-cluster pair.

With the track topo-clusters in hand the energy deposit in the topo-cluster is then estimated with the track information and then subtracted from the topo-cluster on a cell by cell basis, leaving behind the neutral energy (energy from neutral particles in the shower). The last step of the algorithm is to check if the remaining neutral

energy is consistent with what we expect of a single particle at that momentum. This method leverages the tracker's superior momentum resolution relative to the calorimeter's energy resolution, particularly for low-energy charged particles. It also enhances particle acceptance within the detector, as the track reconstruction threshold lies below the calorimeter's noise level required for topological cluster seeding. Additionally, it reduces the likelihood of missing jet constituents caused by significant track curvature in the magnetic field, which can lead to mismatches in calorimeter-based reconstruction.

Finally, pile-up contamination is naturally suppressed through more effective identification of the primary interaction vertex and improved rejection of contributions from secondary (pile-up) vertices. This works well for low-p_T jets but struggles at high-p_T where jet groupings are tighter and there is more energy overlap between the topo-clusters. In order to combat this the updated PFlow algorithm improves performance in dense, high-energy jets by introducing a gradual transition in how it subtracts calorimeter energy from matched clusters. For high-p_T tracks, less energy is subtracted, recognizing the increased precision of calorimeter measurements at those energies. This refinement minimizes errors in crowded jet cores and ensures a more accurate division between track-based and calorimeter-based energy measurements. The resulting PFlow objects are used to make PFlow jets, the main jet type used in the $t\bar{t}H$ analysis we will discuss Chapter 9 and the high p_T uncertainty extrapolation study in Chapter 7.

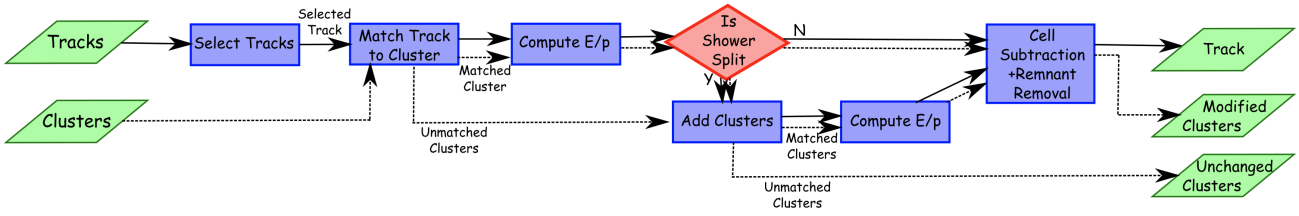


Figure 5.7 A flow chart of how the particle flow algorithm proceeds, starting with track selection and continuing until the energy associated with the selected tracks has been removed from the calorimeter.

5.4.2 Jet Collections

A jet collection is a set of reconstructed jets in an event, built using a specific clustering algorithm (anti-k_t), input objects (PFlow) and parameter settings (R). As the anti-k_t and Pflow algorithms are pre-defined we only have one parameter to play with: R. For the

analysis in Chapter 9 there are two main jet collections used: Small-R and Large-R jets. Small-R jets, or more simply jets, are used to model quark and gluon jets with $R = 0.4$. ATLAS calls the jet collection the anti-k_t, PFlow with $R = 0.4$ `Antikt4EMPFlowJets`.

Large-R jets are slightly more complicated. They are defined with a $R = 1.0$ and are used to model boosted hadronic decay products of heavy particles such as Higgs bosons and top quarks. In boosted topologies the decay products of the parent particle become more collimated and so the individual jets become harder to resolve individually. Instead a re-clustering algorithm in employed in order to for the decay products to be reconstructed as a single large-radius jet object to ensure the full hadronic decay is captured [88].

5.4.3 Jet Calibration

As the jets are reconstructed the energy associated with them will differ from the truth-particle jets, in order to adjust for this discrepancy several stages of calibration takes place, which are shown in Figure 7.2. In jet reconstruction, topological clusters from the calorimeter are initially calibrated using the electromagnetic (EM) scale, which correctly reflects energy deposits from electromagnetic showers such as those from electrons or photons. However, this scale underestimates the energy of hadrons. PFlow jets, the reconstruction includes both calorimeter clusters and charged-particle tracks. Since tracks provide more precise measurements of charged particles' momenta, the initial calibration of PFlow jets is closer to the true particle-level energy. To ensure consistency between data and simulation, a final in situ calibration step is applied, aligning the jet energy response in both data and Monte Carlo (MC) simulations.

The first step in the jet calibration is to mitigate the effects of pile-up interaction on the jet kinematics. To account for this we estimate the pile-up energy density in the event and subtract the product of this density and the jet area as a bigger jet picks up more contamination. Further pile-up corrections are applied based on the average number of interactions per bunch crossing and the number of reconstructed primary vertices in the event.

After correcting for pile-up effects, the next step is the absolute jet energy scale (JES) correction. This adjustment ensures that the energy of reconstructed jets in the detector

matches the energy of the corresponding truth-level (simulated particle-level) jets. In addition, a correction is applied to the jet's direction of η to account for biases introduced by the detector geometry, such as in the transition regions between different parts of the calorimeter, where jet reconstruction can be less accurate.

Next, a Global Sequential Calibration (GSC) is applied to further refine the jet energy measurement and reduce its dependence on the type of particle or process that produced the jet. Different types of partons (i.e. quarks and gluons) create jets with different internal colour structures affecting the fragmentation which can affect how the calorimeter responds. For example, gluon-initiated jets tend to have more soft, low-energy particles, leading to a lower calorimeter response compared to quark jets. This variation introduces unwanted biases. The GSC is applied in a step-by-step (sequential) manner, using additional information from: tracks, the MS, and calorimeter shower shapes. These inputs help correct the jet response without changing the total jet energy, but reduce fluctuations and improve the jet energy resolution (JER).

The final step in the calibration is to account for any residual difference between the MC and data. It relies on well-understood reference objects, such as photons or Z bosons, whose energies are precisely measured. By comparing the p_T of a jet to that of the recoiling reference object in events like γ+jet or Z+jet, any imbalance reveals a residual miscalibration in the jet energy and the JER. The associated PFlow jet calibration uncertainties are reliable only up until a certain p_T. At this point we run out of statistical power to perform in-situ calibration and so must extrapolate in order to extend to higher energies. This step is discussed further in Chapter 7

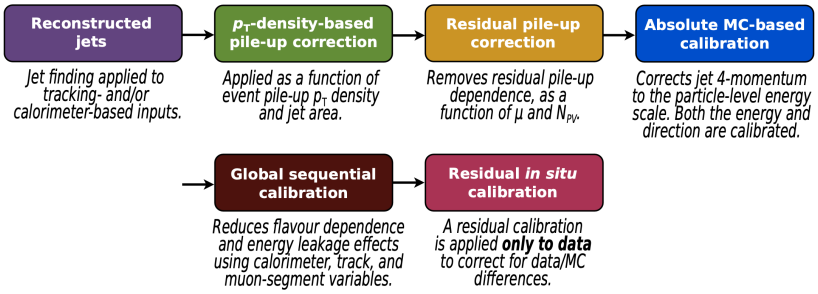


Figure 5.8 The stages of the jet energy scale calibration procedure for reconstructed jets. All steps up until the final in-situ calibration use MC based corrections [89].

5.5 Tau Leptons

Tau leptons decay either leptonically or hadronically. Leptonic decays are reconstructed as electrons or muons, while hadronic decays (τ_{had}) are reconstructed using a procedure similar to that for small-R jets, with dominant decay modes involving one or three charged pions. The primary background to one-prong decays is electron misidentification, which is mitigated using a dedicated electron identification working point, replacing the earlier explicit veto. To distinguish τ_{had} from quark or gluon-initiated jets, a multivariate discriminant is used, incorporating track multiplicity, kinematics, jet substructure, and collimation [90]. The τ_{had} identification employs several WP, with the medium WP used in the $t\bar{t}H(b\bar{b})$ analysis. τ_{had} candidates are vetoed to ensure orthogonality with other $t\bar{t}H$ decay channel analyses, facilitating future combinations with minimal overlap.

5.6 Flavour Tagging

Flavour tagging is one of the more central points of the work in this thesis. It corresponds to the identification of the flavour of the hadron that created the jet in a given hard scattering event. This is a key part in the $t\bar{t}H(b\bar{b})$ analysis later on and also the backbone for the high p_T flavour tagging uncertainty extrapolation study described in Chapter 7, which in turn helps the overall systematic treatment of the high p_T of the regions of the $t\bar{t}H(b\bar{b})$ analysis.

Flavour tagging is done using specialised algorithms [91]. Where jets containing a b-hadron are called b-jets and jets containing a c-hadron are called c-jets. If the jet contains neither then it is designated a light-jet. The algorithms use the significant mean lifetime, mass and distinct fragmentation patterns of b-hadrons in order to distinguish between the flavours. As the energetic b-hadrons decay after $\approx$ 1.5 ps they can be identified by the presence of a displaced or secondary vertex shortly after the collision point where the initial hard scattering event occurred. This is shown in Figure 5.9, where the b-jet has a secondary vertex and the light jet does not. The track impact parameter information is used to reconstruct these vertices.

The current king of the hill in the flavour tagging algorithm world is the `DL1` group of algorithms. They incorporate a two-stage approach, the first stage (low-level algorithms) utilises low-level track features to build and reconstruct the tell-tail secondary vertex of

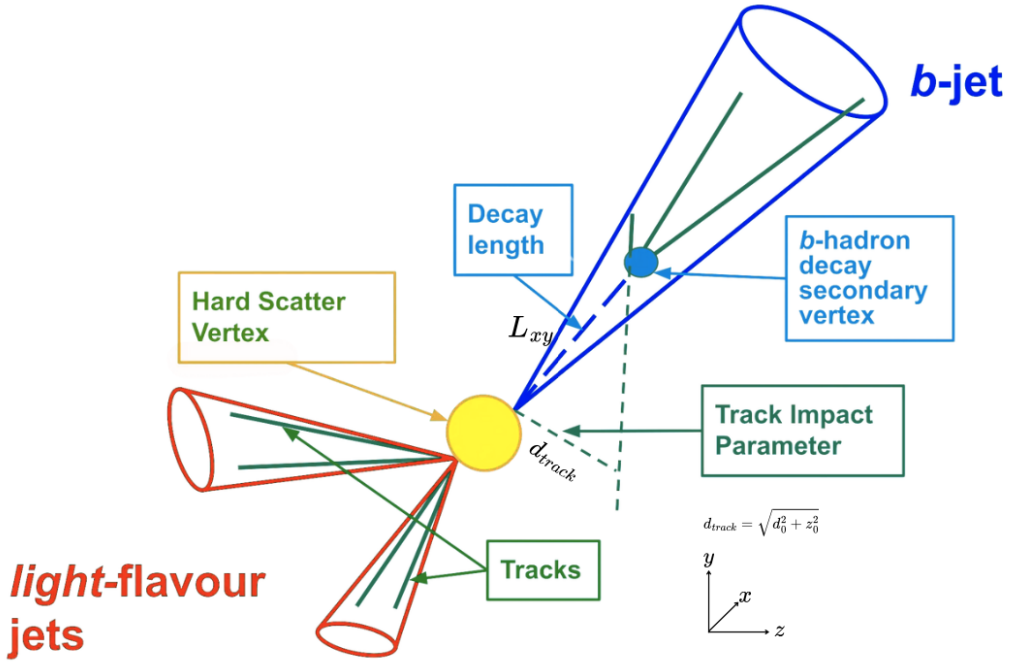


Figure 5.9 An illustration depicting the signature b-hadron decay with a b-jet containing a secondary vertex after a decay length from an initial hard scattering event, compared to a light-jet with no secondary vertex or decay length post hard scattering event [92].

a b-hadron jet structure. The second stage uses deep-learning methods to deduce the optimum combination of the low-level algorithm outputs.

5.6.1 Low-level Algorithms

The main inputs for the low-level algorithm are: charged-particle tracks, the primary-vertex, and hadronic jets. Tracks are matched to jets via the angular separation of the track momentum and jet axis, ΔR, this is dependent on p_T of the jet as higher momentum decay products tend to be more collimated and so will have a smaller ΔR.

The three main low-level algorithms are `IP2D`, `IP3D` and `RNNIP` [93, 94]. `IP2D` and `IP3D` build likelihood based discriminants using the impact parameter significance of tracks (impact parameter divided by its uncertainty). IP2D uses just the transverse impact parameter, while IP3D uses both transverse and longitudinal components. Each track's information is evaluated under different flavour hypotheses (b-jet, c-jet, light jet), and a likelihood score is built from the sum of individual track probabilities. They also use expected ID hit patterns to refine the likelihood. `RNNIP` s a machine learning approach using a Recurrent Neural Network (RNN), Unlike the previous methods, RNNIP captures correlations between tracks. Including, how multiple decay products from a b-hadron behave together. The main input features are: impact parameter significances, track momentum fractions relative to the jet p_T, hit multiplicities in the inner detector and ΔR. The RNNIP model outputs class probabilities (i.e., how likely a jet is to be a b-jet, c-jet, or light-flavour jet), and these are combined into a b-tagging discriminant. This approach outperforms the previous algorithms as it does not assume independence of per-track features in the jet.

The reconstruction of the secondary vertices is done using the `SV1` algorithm in an iterative approach as it examines all possible two-track vertex combinations using a χ^2 fit to evaluate all possible combinations [95]. The final step is applying a multi-vertex finding algorithm, `JetFitter` [96], which uses a Kalman filter to reconstruct multiple decay vertices inside a jet, capturing the full decay chain of b and c-hadrons. This improves the determination of the decay topology and enhances the accuracy of b-jet identification.

5.6.2 High-level Algorithms

The high-level algorithms use the outputs of the low-level ones as inputs into a multi-class feed-forward deep neural network. There are three main b-tagging algorithms worth mentioning `DL1`, `DL1r` (which uses the same inputs as `DL1` in addition to the `RNNIP`

algorithm [97]. and `MV2c10` (an older tagging algorithm which is used in the previous version of the $t\bar{t}H(b\bar{b})$ analysis and uses a boosted decision tree (BDT) approach) [98]. The algorithms are trained on an enhanced $t\bar{t}$ sample set which incorporates a large number of high mass $Z' \to q\bar{q}$ events in order to boost the statistics of high p_T jets [99]. Similar samples are used in the the high p_T uncertainty extrapolation study in Chapter 7.

In order to further discriminate between jet flavours, `Dl1r` uses the output probabilities to create a discriminate:

$$D_{\text{DL1r}} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\text{light}}}\right). \tag{5.4}$$

Here p_b, p_c and p_{light} represent the jet-flavour probabilities and f_c is the fraction of c-jets. f_c is was specifically tuned to optimise performance at the level of an analysis, with many analyses taken into consideration including $t\bar{t}H(b\bar{b})$. Figure 5.10 displays the distribution of the b-tagging discriminant, with annotations indicating the tagging score bins used to define the pseudo-continuous b-tagging (PCBT).

Before we proceed with the discussion of the taggers lets briefly explain PCBT. Traditionally fixed cut WPs were employed for b-tagging meaning that anything below the cut off threshold was thrown out. This meant, there were jets slightly below the threshold that with a relatively high b-tagged score that were thrown out, reducing the overall statistics. For analyses that are statistically limited this is a suboptimal approach. PCBT aims to use the whole the full output of a b-tagging algorithm, typically

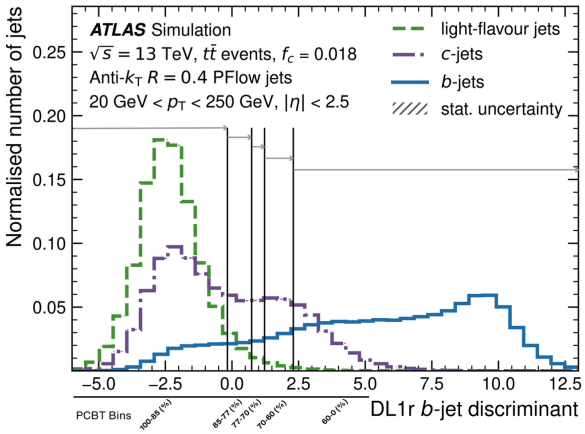


Figure 5.10 Distribution of the DL1r discriminant output for b, c, and light-flavour jets, assuming a charm-jet fraction of $f_c = 0.018$ in the sample.

a continuous discriminant score ranging from 0 to 1, to characterize how likely a jet is to originate from a b-quark. Instead of applying a single threshold to classify jets as b-tagged or not, the score is divided into multiple bins, often defined by standard fixed WP thresholds (e.g. 60%, 70%, 77%, 85% b-tagging efficiency). This approach retains more information about the jet's flavour probability, allowing analyses to include jets with varying degrees of b-likeness. These bins can then be used in template fits or multivariate classifiers, improving sensitivity to processes involving b-jets by making better use of the available statistics and exploiting differences in flavour composition across the score range.

Returning to the performance evaluations of the three b-tagging algorithms, which can be seen in Figure 5.11. The performance evaluation was conducted using a c-jet fraction of $f_c = 0.018$, and focuses on two key metrics: the b-jet tagging efficiency (ϵ_b)—the probability of correctly identifying a b-jet, and the background rejection factor ($1/\epsilon_j$), where j represents either light-flavour or charm jets. This rejection factor quantifies the inverse of the mistag rate for non-b-jets. Figure 5.11 presents the rejection performance of light and charm-jet backgrounds as a function of b-jet efficiency, comparing three algorithms: `DL1`, `DL1r`, and `MV2c10`. The lower panels of the figure provide a direct comparison of the rejection achieved by `DL1` and `DL1r` relative to the previously used `MV2c10`, illustrating significant improvements in flavour rejection.

The `DL1r` continued to outperform the other two algorithms when it came to using PFlow jets. Figure 5.12 presents the performance of the Run 2 flavour-tagging algorithms, evaluated as a function of jet transverse momentum p_T, using simulated $t\bar{t}$ and Z' events for the 77% b-tagging efficiency working point.

The results use PFlow jets in two p_T intervals: 20 GeV $< p_T <$ 250 GeV and 250 GeV $< p_T <$ 3000 GeV, with pseudorapidity restricted to $|\eta| < 2.5$. The tagging performance is shown in terms of light and c-jet rejection rates, with the `DL1r` algorithm indicated in blue.

The top left and right plots show light-jet rejection, where a degradation in performance is observed at high p_T, particularly evident in the Z' events on the right. This decline is attributed to increased jet fragmentation at high momentum, resulting in more reconstructed fake secondary vertices, as well as genuine secondary vertices in boosted light-flavour jets that can mimic heavy-flavour decays.

A more pronounced drop in light-jet rejection, visible around $p_T \sim 175$ GeV $\sim m_{\text{top}}$ in

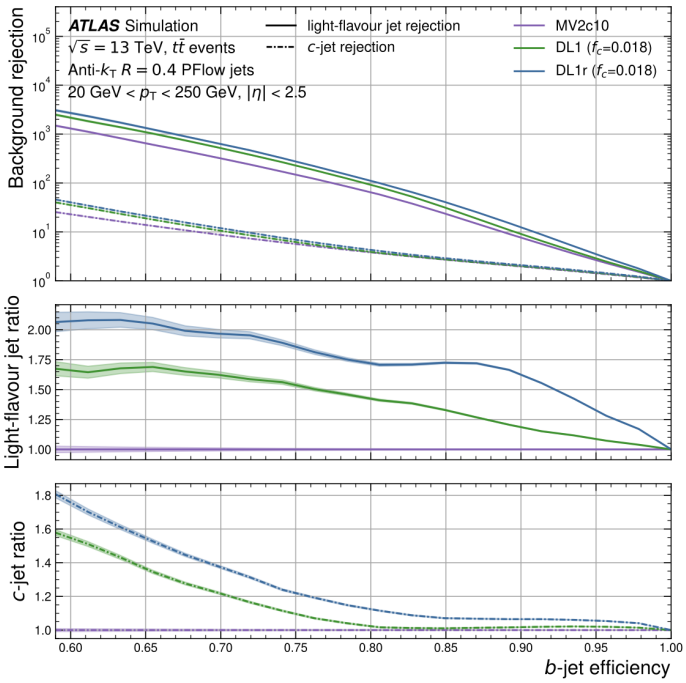


Figure 5.11 Comparison of light and c-jet mis-tag rates as a function of b-jet efficiency for the `MV2c10`, `DL1`, and `DL1r` taggers. The lower ratio panels show the light- and c-jet rejection relative to the `MV2c10` tagger [98].

the top-left plot, is linked to overlapping decay products in boosted $t\bar{t}$ events, where these effects become especially prominent.

Charm-jet rejection performance is shown in the lower plots, with Z' events on the right. Unlike the light-jet case, the ability to distinguish b-jets from charm-jets remains relatively stable across the p_T range. In Z' events, the `DL1r` algorithm's performance improves with increasing jet p_T, due to higher track multiplicity and more displaced vertex topologies in boosted jets. These features enhance the algorithm's ability to separate c from b initiated jets.

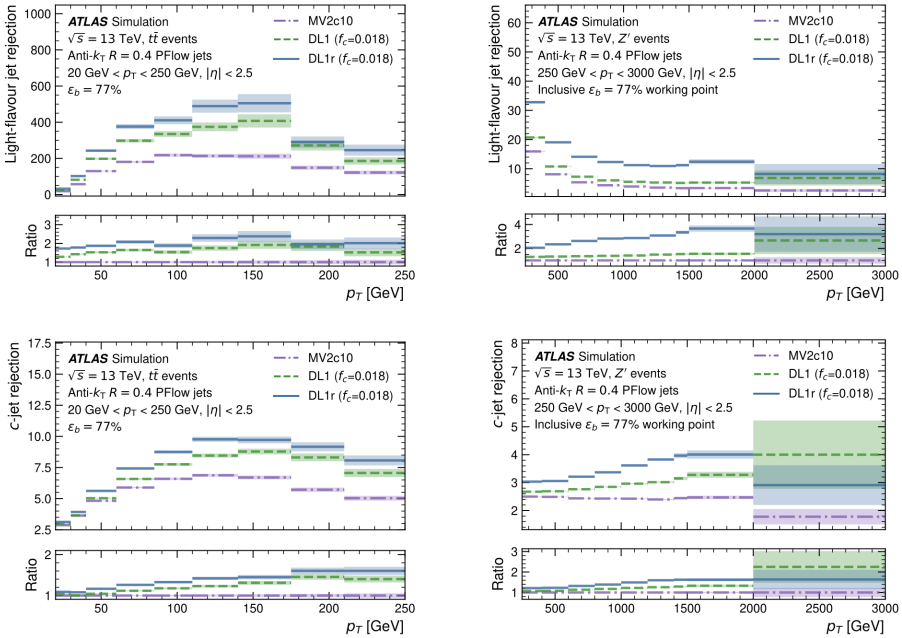


Figure 5.12 Performance of the Run 2 jet flavour-tagging algorithms as a function of jet transverse momentum (p_T) and pseudorapidity (η), illustrated through light-jet and c-jet rejection factors .

5.7 Overlap Removal

When an electron's energy deposits are used in jet reconstruction, occasionally the signal detector response of the calorimeter may misconstrue the deposits as more than one physics object per event. To avoid this double counting, overlap removal procedures are implemented. The overlap removal algorithm proceeds in a sequential manner, where at each stage only the objects that have not been rejected by earlier steps are considered. This approach ensures that each detector signature is attributed to only one reconstructed object, such as a lepton or a jet. The process begins by identifying cases where electrons and muons share a common inner detector track; in such instances, the electron is discarded. Next, jets located within a distance $\Delta R < 0.2$ of an electron are removed, as they are likely to be electromagnetic in origin. If an additional jet lies within $\Delta R < 0.4$ of an electron, the electron is instead removed to avoid ambiguities in reconstruction. Muon candidates are excluded if they are found within $\Delta R < 0.4$ of a jet, in order to suppress muons produced from in-jet decays of heavy-flavour hadrons. However, if the associated jet has fewer than three tracks, the muon is retained and the jet is removed—this exception accounts for energetic muons depositing significant

energy in the calorimeter. Hadronic tau candidates (τ_{had}) are eliminated if they lie within $\Delta R < 0.2$ of a selected electron or muon, while no overlap removal is performed between jets and tau candidates.

5.8 Missing Transverse Momentum

As the particles collide along the beam axis, we expect the total transverse momentum of all calibrated event objects to sum to zero. Any significant imbalance suggests the presence of weakly interacting particles (neutrinos) or it could be attributed to detector imperfections. Either way, we need to account for this imbalance and so we introduce a new "observable", missing transverse momentum p_T^{miss} [100]. We can deduce two main quantities from this: the first is missing transverse energy, E_T^{miss}, which is defined as the magnitude of the negative vector sum of all detected final-state object transverse momenta, including an additional soft-term derived from tracks reconstructed in the ID that are matched to the primary interaction vertex but not associated to any object, as shown in:

$$\vec{E}_T^{\text{miss}} = -\left| \sum_{i \in \text{Objects}} \vec{p}_{T,i} + \sum_{j \in \text{Soft-term}} \vec{p}_{T,j} \right|. \tag{5.5}$$

The azimuthal angle of p_T^{miss} given by:

$$\phi^{\text{miss}} = \arctan(p_y^{\text{miss}}/p_x^{\text{miss}}). \tag{5.6}$$

These objects lay the foundation for the analysis which we will discuss in Chapter 9.

Chapter 6

Monte-Carlo Simulations

"You're always one spin away form winning. Don't stop..."
— Iain Mclean, when asked about healthy habits.

MC simulations are a key part of any ATLAS physics analysis as they provide a reference to data with defined physics inputs in order to understand the detector response. Generation of the MC events start: the Matrix Element (ME) for the hard-scattering process of interest. However, with any hard-scattering event comes soft emissions that if would be too computationally intensive to calculate in the same way as the hard scattering events and so these emissions are simulated using a Parton Shower (PS). The resulting partons then under go hadronization, using a non-perturbative model. Remnants of the protons after collisions can also enter the detector volume and interact further, or decay forming part of the underlying event. As each bunch crossing results in around 30-40 other interactions, pile-up also needs to be modelled. Lastly the detectors effects must be taken into account and the final-state particles must be passed through a simulation of the ATLAS detector geometry. An illustration of these stages can be seen in Figure 6.1.

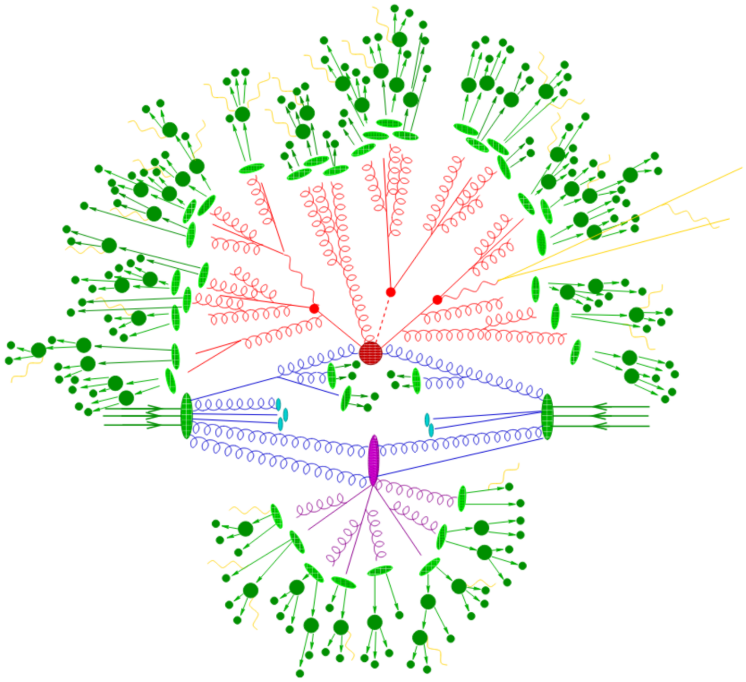


Figure 6.1 Schematic illustration of a semi-leptonic $t\bar{t}H$ event generated with a Monte Carlo event generator. The incoming proton partons are indicated by dark green blobs with parallel arrows pointing towards the interaction point. The hard-scattering process, produced via gluon–gluon fusion, is shown as a large red blob. The parton shower is depicted in blue. The decays of the top quarks and Higgs boson are represented by smaller red blobs, corresponding to the $t\bar{t}H \to bq\bar{q}\,\bar{b}\ell\nu\,b\bar{b}$ final state. Additional hard QCD radiation is shown in red, while secondary interactions from the underlying event are shown in purple. Particles prior to hadronisation are shown in light green, with hadrons after hadronisation shown in dark green. Photon radiation is depicted in yellow. [101].

6.1 Matrix Element Generation

The cross section of all QCD interactions requires knowledge of the strong coupling constant, α_s which is inversely proportional to the logarithm of the to the energy of a given interaction. Considering a *pp* collision, these collisions involve a mix of interactions at different energy scales. At low energies ($\alpha_s > 1$) higher-order perturbative contributions to the collision cross-section do not converge and thus the calculation cannot be solved using Equation 6.1.

$$\hat{\sigma}_{ab\to X} = \hat{\sigma}_0 + \alpha_S\hat{\sigma}_1 + \alpha_S^2\hat{\sigma}_2 \ + \ ..., \tag{6.1}$$

where $\hat{\sigma}_0$ describes the Leading Order (LO) (or tree level) contribution, $\alpha_S\hat{\sigma}_1$ are the Next-to-Leading Order (NLO) corrections and so on. At the next stage, in the PS, the real emissions from the higher order α_S term are described either as Initial State Radiation (ISR) or Final State Radiation (FSR) we discuss this more in Section 6.2.

6.1.1 Factorisation Theorem

In order to combat the non-perturbative nature of high order α_S terms, the factorisation theorem is deployed [102]. This states that by defining a factorisation energy scale (μ_f) as a crossing point between perturbative and non-perturbative behaviour the cross-section calculation can be factorised into high-energy and low-energy and so can be solved independently. A production cross-section of some process X from a proton-proton collision can be given by:

$$\sigma_{pp\to X} = \sum_{i,j\in\{q,\bar{q},g\}} \int dx_1 \int dx_2 f_i(x_1,\mu_f^2) f_j(x_2,\mu_f^2)\hat{\sigma}_{ij}, \tag{6.2}$$

where i and j denote the parton species of the incident particle, $f_{i(j)}(x_{1(2)},\mu_f^2)$ is the Parton Distribution Function (PDFs), and $\hat{\sigma}_{ij}$ is the partonic cross-section for parton i,j to produce X.

6.1.2 Parton Distribution Functions

In a proton we have three valance quarks (uud) and a sea of other miscellaneous quarks and gluons which aren't fully bound at high energies. Each one of these partons carry

some fraction of the protons total momentum in order to describe the "who has what" problem this description of a proton faces, we introduce the PDF. The PDF, $P_i(x)$, gives the probability for finding a parton of type i carrying a fraction x of the proton's momentum. These distributions are relevant at interaction scales typically ranging from 100 to 1000 GeV, where high-energy collisions probe the internal structure of the proton. Since PDFs arise from non-perturbative QCD dynamics, they cannot be calculated directly from first principles and are instead parametrised using experimental data. However, their dependence on the energy scale is predictable and is governed by the DGLAP evolution equations, which describe how PDFs change as the probing scale increases [103, 104, 105].

An important part of the factorisation scheme is how b-quarks are treated in QCD processes. There are two main approaches. In the Four Flavour Scheme (4FS), b-quarks are considered massive and too heavy to appear in the initial state. As a result, they don't have a parton distribution function (PDF) and are excluded from the running of α_S. This approach is better for low-energy processes where the b-quark mass matters. In contrast, the Five Flavour Scheme (5FS) is used at higher energies, where the b-quark mass becomes negligible. Here, b-quarks are treated as massless, included in the proton's structure via a PDF, and contribute to α_S running like the lighter quarks.

6.2 Parton Shower

Once we have generated our hard scattering process, we then need to apply parton shower effects, where the high-energy quarks and gluons produced begin to radiate further particles. Each parton has a finite probability of splitting, such as a gluon decaying into a quark-antiquark pair ($g \to q\bar{q}$), a quark emitting a gluon ($q \to gq$), or a gluon splitting into two gluons ($g \to gg$). Parton shower generators calculate these probabilities at different scales, simulating the step-by-step radiation as partons lose energy. This process continues until the evolution reaches scales of order 1 GeV, where perturbative QCD is no longer valid and the simulation hands over to hadronisation models to form stable particles. Radiation is typically modelled as two separate but conceptually similar cascades FSR and ISR.

FSR occurs after the hard scattering, where outgoing high-virtuality partons branch as they propagate towards lower energies. The probability that a parton does not emit

further radiation between two evolution scales Q and Q_0 is given by the Sudakov form factor,

$$\Delta(Q^2, Q_0^2) = \exp\left[-\int_{Q_0^2}^{Q^2} \frac{\mathrm{d}k^2}{k^2} \int \mathrm{d}z\, P(z)\right], \tag{6.3}$$

where $P(z)$ is the relevant Altarelli-Parisi splitting function [106]. Where the emission probability grows as the scale decreases, and the Sudakov factor acts as a survival probability against emission.

ISR describes radiation from the incoming partons before the hard scattering. ISR is typically implemented via backward evolution, starting from the hard-scattering scale and evolving to lower scales, reconstructing possible emissions in reverse. The emission probability is again governed by a Sudakov factor, but here the splitting probabilities are weighted by the ratio of parton distribution functions (PDFs). This accounts for the fact that, when evolving backward, certain branchings are suppressed if the required parent parton is rare in the proton at the relevant momentum fraction.

In both cases, the Sudakov form factor ensures the probabilistic consistency of the shower: it enforces that exactly one branching occurs at each step, with the intervals between emissions determined by the no-emission probability. The shower evolution thus resums large logarithms associated with soft and collinear radiation, providing a bridge between the hard process and the hadronisation stage.

6.3 Hadronisation

Once we have reached the showering threshold, the parton energy is low enough that quarks begin to form bound states, resulting in the formation of hadrons, a process known as hadronisation. There are two main models used to simulate hadronisation: the cluster fragmentation model and the Lund string model [107]. The cluster fragmentation model, used by SHERPA, works by first forcing gluons to split into quark–antiquark pairs [108]. These quarks then group into colour-neutral clusters, which subsequently decay into hadrons. In contrast, the Lund string model, used by PYTHIA, treats the colour field between partons as a string that stretches as they move apart. When the string's potential energy becomes large enough, it breaks, creating new quark–antiquark pairs. This process repeats until all partons are converted into colourless hadrons. Both

models are based on the principle of Local Parton–Hadron Duality, which assumes that the momentum and quantum number flow at the hadron level follows that of the parton level [109].

6.4 Underlying Event and Multi-Parton Interactions

The Underlying Event (UE) refers to all the activity in a proton–proton collision that is not part of the main hard scattering, including remnants of the incoming protons, soft radiation, and additional parton–parton scatterings known as Multi-Parton Interactions (MPI). These MPIs occur when multiple partons from each proton interact simultaneously, often producing low-momentum particles. Like the products of the hard scatter, the partons from the UE and MPI also carry colour charge and therefore must form colour-bound states. They undergo hadronisation through the same mechanisms, such as the Lund string or cluster model—forming hadrons that contribute to the soft particle activity observed in the final state of a collision event.

6.5 Pile Up

Pile-up is modelled by overlaying additional soft minimum-bias events (representing random pp collisions) on top of the hard-scatter event, These are generated using `PYTHIA8` [110]. The number of contributions can sometimes not match data and so need to be re-weighted to do so, this is called pile-up re-weighting.

6.6 Detector Simulation

The final step in the Monte Carlo event generation process is to pass the simulated events through a detector simulation, which models how the final-state particles interact with the detector hardware. For the ATLAS experiment, this is done using `GEANT4`, a detailed toolkit that simulates the passage of particles through the detector systems [111, 112]. The tools includes simulating: interactions with materials, energy loss, scattering, and detector responses. Accurately modelling these effects requires significant computational effort. As a result, the full `GEANT4`-based simulation is often the most time-consuming part of the entire MC workflow, representing the dominant share of CPU resources used

in sample production for ATLAS analyses.

An intermediate option is `ATLFAST-II`, which combines full `GEANT4` simulation for the inner detector and muon spectrometer with the parameterised calorimeter simulation `FastCaloSim`, typically interfaced via `FATRAS`. This hybrid approach delivers an order of magnitude reduction in CPU time compared to the full simulation, while maintaining significantly higher accuracy than the fully parameterised `ATLFAST` model [113, 114, 115].

More recently, `ATLFAST3` has been developed as the next-generation fast simulation platform. By combining traditional FastCaloSim techniques with machine-learning based calorimeter shower models, it achieves accuracy closer to full `GEANT4` simulation while preserving the high-speed performance needed for large event samples [116].

Chapter 7

Uncertainty Extrapolation for High-p_T PFlow Jets

"I'm all energy, no technique."
— Dr. Dwayne Spiteri, referring to his squash ability.

High-p_T flavour tagging uncertainties are essential for analyses involving boosted regions. At high transverse momentum, the kinematic properties and structure of jets evolve significantly: jets become more collimated, substructure variables are altered, and the performance of flavour tagging algorithms begins to deviate from their behaviour in the lower-p_T regime. Since PFlow jets combine calorimeter and tracking information, their response and tagging behaviour at high-p_T are sensitive to detector effects and modelling limitations that may not be fully captured by simulations. Without properly quantified uncertainties, the scale factors (SFs) applied to match data and simulation could be miss-estimated, directly impacting signal efficiencies and background rates in searches for boosted objects, like the Higgs bosons, or new resonances.

The PCBT approach, which allows finer granularity in the b-tagging decision by treating the tagger output score in efficiency bins, is particularly powerful but introduces complexity in calibration. As it treats b-tagging as a continuous observable rather than a binary one, the uncertainties must account not just for the b-tag efficiency in p_T, but also for the shape and migration across score bins. This shape sensitivity becomes even more crucial at high p_T, where the tagger response may shift due to changes in the underlying jet composition or detector resolution. As a result, PCBT requires a detailed

and reliable propagation of uncertainties, especially for PFlow jets, which are more sensitive to tracking performance and pileup effects, both of which are more variable at high energies.

Moreover, many physics analyses using PFlow jets and PCBT are designed to explore boosted topologies, where the sensitivity relies on correctly identifying high-p_T b-jets. In these regions, flavour tagging uncertainties often dominate the total systematic uncertainty budget. If the uncertainties in this regime are not well modelled or are overly simplified (e.g. due to pruning or insufficient calibration coverage), it can lead to biased estimates of signal strengths or artificial sensitivity in limits. Accurate high-p_T flavour tagging uncertainties ensure robust uncertainty propagation and enable consistent treatment across different analyses, working points, and jet definitions. This is particularly important for combinations or reinterpretations that span a wide p_T range and depend on the integrity of uncertainty correlations.

Here, we define the high-p_T uncertainty extrapolation method used in the $t\bar{t}H(b\bar{b})$ Legacy analysis as this analysis used PFlow jets and PCBT. This piece of work was done solely to improve the uncertainty handling in the high-p_T resolved and boosted regions of the $t\bar{t}H(b\bar{b})$ Legacy analysis but was later implemented in a calibration data interface (CDI) for software release 21. This means, that any analysis, using PFlow jets and PCBT in high-p_T regions, across the ATLAS collaboration would benefit from this work.

7.1 Extrapolation Uncertainty

Flavour tagging requires the use of data-to-simulation scale factors because the performance of the tagging algorithms in Monte Carlo simulations rarely matches that observed in real data. These differences arise due to imperfect modelling of detector responses, the complexity of hadronization processes, and various systematic effects inherent in the simulation of physics processes. By deriving scale factors that represent the ratio of the tagging efficiency in data to that in simulation, physicists can correct for these discrepancies. This calibration ensures that analyses relying on flavour tagging have accurate efficiencies and background estimates, ultimately leading to more reliable physical measurements and limits. Hence the first ingredient in our extrapolation recipe are these scale factors.

In order to establish the data-to-simulation scale factor for a given flavour of jet we need an abundance of events within a given p_T range. For example, to generate the efficiency scale factors used in b-tagging, both an enriched MC sample of $t\bar{t}$ events was used and well understood reference objects to provide in-situ calibration and uncertainties. However this approach sets a limit in which the efficiency scale factors can be reliably generated as the limited data statistics means there are p_T regions beyond which the data-to-simulation scale factor cannot be reliably determined. For b-tagged Pflow jets the upper limit is $p_T = 400$ GeV, this sets an in-situ calibration threshold on b-jets. The calibration threshold for each flavour of jet can be seen in Figure 7.1. The uncertainties need to then be extrapolated to 3 TeV.

In order to effectively derive accurate systematic uncertainties that assume the scale factors are constant, beyond the calibration threshold, an extrapolation method must be used. The employed flavour tagging uncertainty extrapolation method for PCBT has a heavy reliance on data-to-simulation scale factors. Let us briefly describe the extrapolation method for Pflow b-jets [117]. This method can be extended to all flavours of jets and jet types and for this study is extended to c and *light* jets. This method has already been used in the high-p_T uncertainty extrapolation of VR-jets [118].

The b-tagging efficiency is defined as:

$$\epsilon_b = \frac{\text{\#of jets that are } b\text{-tagged}}{\text{Total \# of truth } b \text{ jets}}.$$

We use this to then define our data-to-simulation scale factors (SF) e.g for b-jets as:

$$\text{SF}_{b,w} = \frac{\epsilon_{b,w}^{\text{data}}}{\epsilon_{b,w}^{\text{MC}}},$$

where $\epsilon_{b,w}^{\text{data}}$ and $\epsilon_{b,w}^{\text{MC}}$ represent the efficiency obtained by data and simulation, respectively. the sub-script of b, w represent the flavour and working point (60%, 70%, 77%, 85% and 100%), respectively.

In principle if we sum the efficiencies for a particular flavour in a single p_T bin across all working points, we should reach a closure point of 1. For this example we will assume a

b-jet using one working point to simplify the notation:

$$\sum_i \epsilon^i_{p_T} = 1.$$

However, approaching and beyond the calibration threshold of a particular flavour, this closure is not achieved. We introduce the data-to-simulation scale factor in-order to correct the efficiency, but closure is still not met:

$$\sum_i \epsilon^i_{p_T} \mathrm{SF}^i_{p_T} = 1.$$

Attempting to bring us closer to closure, we proceed to expand the scale factor to include the relative scale factor corrected efficiency, $\Delta_{rel}\mathrm{SF}_{p_T}$:

$$\sum_i \epsilon^i_{p_T} \mathrm{SF}^i_{p_T} (1 - \Delta_{rel}\mathrm{SF}_{p_T}) = 1.$$

Expanding $\Delta_{rel}\mathrm{SF}_{p_T}$ we find:

$$\Delta_{rel}(\mathrm{SF}_{pt}) = \Delta_{rel}(\mathrm{SF}_{pt,ref}) + \Delta_{ref}(\epsilon_{MC,pt}) - \Delta_{ref}(\epsilon_{MC,pt,ref}), \tag{7.1}$$

where, $\Delta_{rel}(\mathrm{SF}_{pt,ref})$ is the in-situ calibration SF relative uncertainty at the calibration threshold p_T for our desired flavour and $\Delta_{ref}(\epsilon_{MC,pt}) - \Delta_{ref}(\epsilon_{MC,pt,ref})$ is our extrapolation uncertainty which is purely based on simulation, and contains the simulation efficiency evaluated at our desired p_T and the simulation efficiency at our calibration threshold p_T, respectively. The relative uncertainty, $\Delta_{ref}(\epsilon_{MC,pt})$, for tracking, jet calibration, modelling and global is shown in Figures 7.1, 7.2, 7.3 and 7.4.

Flavour	Calibration threshold (GeV)
b	400
c	250
$light$	300

Table 7.1 The in-situ calibration thresholds for b, c and light Pflow jets.

For our desired flavour of Pflow jet we need to use the extrapolation uncertainty contained in Equation 7.1 and extrapolate all systematic uncertainties to evaluate them past the calibration threshold. The following plots are for b-jets at a 60% WP and generated using

Z' samples, which is a MC sample that is enriched with $t\bar{t}$ events in order to effectively evaluate the SFs at their desired p_T beyond any calibration threshold.

7.1.1 Tracking Uncertainties

As we see in Figures 7.1, the relative tracking uncertainties, the contributions to the overall extrapolation tracking uncertainty are minimal, peaking at ≈ 0.03 at 3 TeV. The tracking-related uncertainties are as follows: `TRK_FAKE_RATE_LOOSE_TIDE` and `TRK_FAKE_RATE_LOOSE` account for uncertainties in the fake track rate in data and simulation, with TIDE referring to a dedicated tuning for dense environments. The remaining uncertainties correspond to different sources of uncertainty in the track reconstruction efficiency for tracks passing loose quality criteria. `TRK_EFF_LOOSE_TIDE` and `TRK_EFF_LOOSE_PP0` stem from data-driven corrections and pileup modelling in the efficiency determination. `TRK_EFF_LOOSE_PHYSMODEL` addresses uncertainties from the underlying physics modelling used in simulation. `TRK_EFF_LOOSE_GLOBAL` represents a global uncertainty in track reconstruction efficiency applied across the detector. Finally, `TRK_EFF_LOOSE_IBL` captures potential miss-modelling of the IBL, which is particularly relevant for high-p_T tracks near the beamline. These components are combined in quadrature to form the total tracking uncertainty.

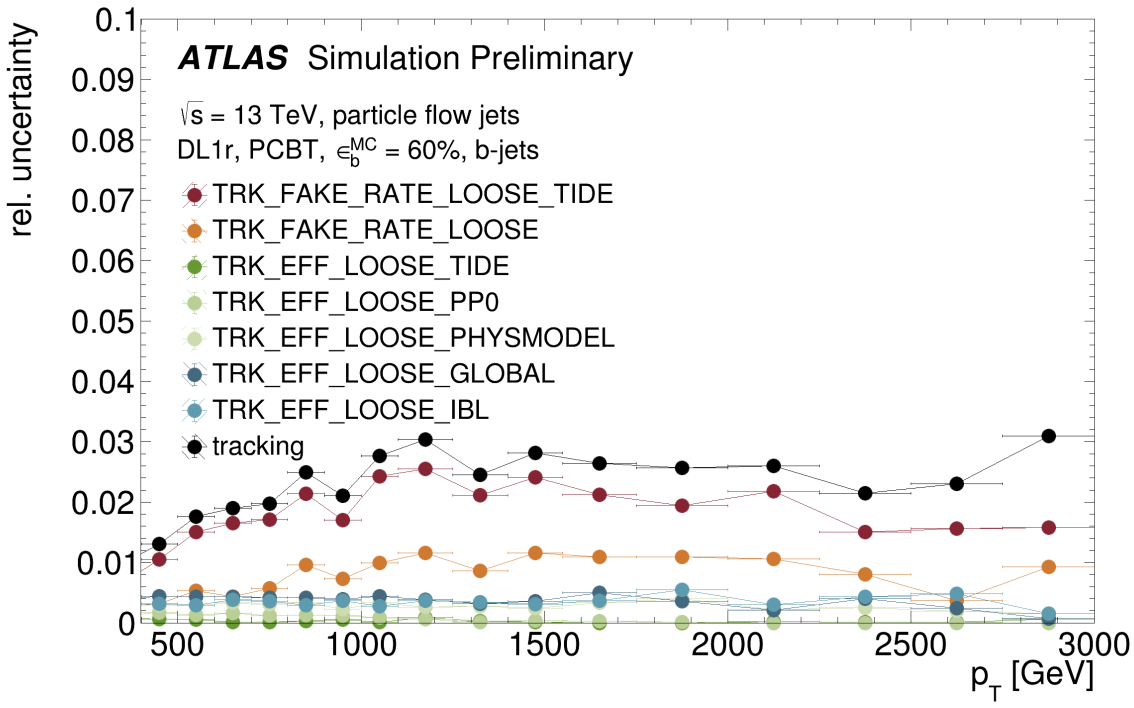


Figure 7.1 Decomposed tracking uncertainty: fake rate in dense (red) and non-dense (orange) environments; tracking efficiencies from various sources including dense environments (dark green), primary particles (light green), physical model (light blue), global (dark blue), IBL (light blue); total (black) for PFlow b-jets with DL1r at 60% efficiency.

7.1.2 Jet Calibration Uncertainties

Much like the tracking uncertainties, the jet calibration and jet calibration extrapolation uncertainties show in Figures 7.2 are similarly quite minimal. The peak of the distribution seen to be ≈ 0.05 at 3 TeV. The jet calibration uncertainties shown in Figure 7.2 is comprised of: `JET_SingleParticle_HighPt` arises from studies comparing data and simulation for isolated high-p_T tracks and assessing the accuracy of momentum reconstruction. `JET_EffectiveNP_*` represents compressed nuisance parameters derived from a combination of sources such as pileup, flavour composition, and detector modelling. These are obtained via diagonalization techniques to reduce the dimensionality of systematic uncertainties. `JET_JER_*` accounts for mismatches in the jet energy resolution between data and Monte Carlo, impacting smearing and unfolding procedures. `JET_EtaIntercalibration_*` captures uncertainties in the relative calibration between jets in different pseudorapidity regions, typically derived from di-jet balance techniques. Together, these uncertainties are combined in quadrature to form the total Jet-calibration uncertainty, which grows significantly in the high-p_T regime due to tracking-based effects [89].

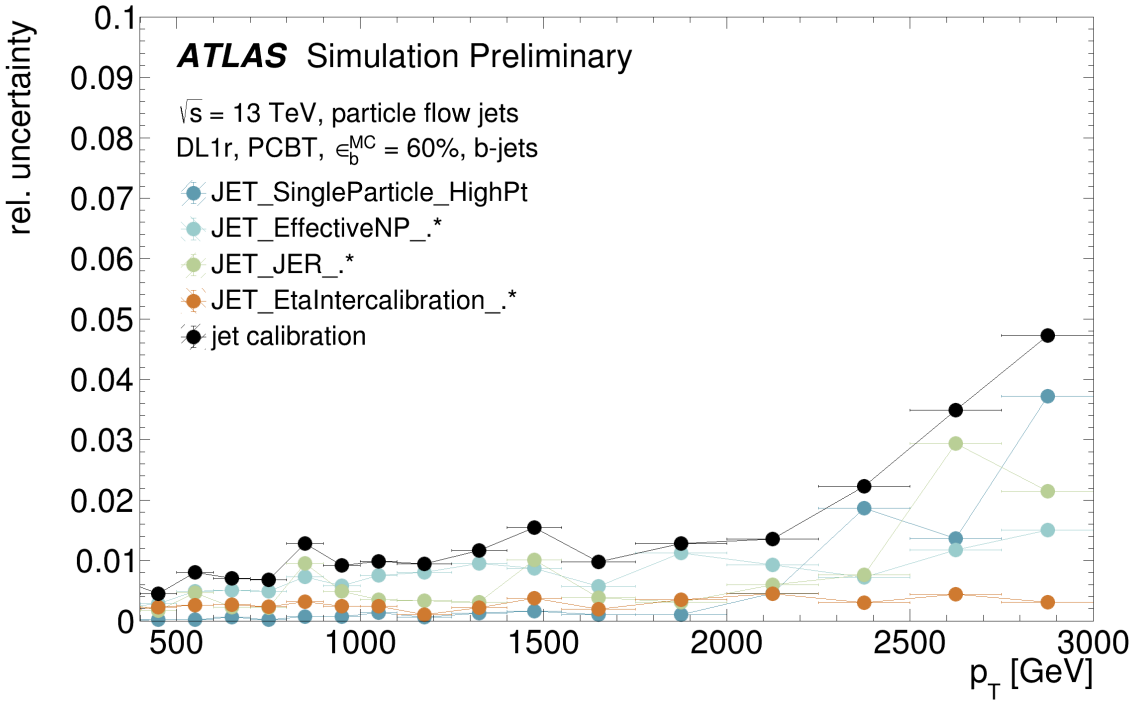


Figure 7.2 Decomposed jet calibration uncertainty: single particle at high p_T (dark blue), effective nuisance parameters (light blue), jet energy resolution (JER, green), η-intercalibration (orange), and total (black) for PFlow b-jets with DL1r at 60% efficiency.

7.1.3 Modelling Uncertainties

Compared to the systematics described in Sections 7.1.1 and 7.1.2 the modelling uncertainties are considerably larger as seen in Figure 7.3. With the main contributions coming from the parton shower and B-hadron interactions. When these are propagated in to the extrapolation uncertainty these significantly inflate the overall value at any p_T bin, peaking at 0.36 at 3 TeV. The modelling uncertainties are: `FSR` corresponds to the uncertainty from final-state radiation modelling and is typically assessed by varying the QCD radiation parameters or using alternative parton shower configurations. `parton shower` represents uncertainties associated with different choices of parton shower generators (e.g., PYTHIA vs HERWIG), which affect the development of the jet and the kinematics of the b-hadron. `B-hadron interaction` accounts for mismodelling of b-hadron decays and their interactions with the detector material, which can influence the reconstruction of secondary vertices and impact parameter significance. The `Total` uncertainty is the quadrature sum of all modelling components.

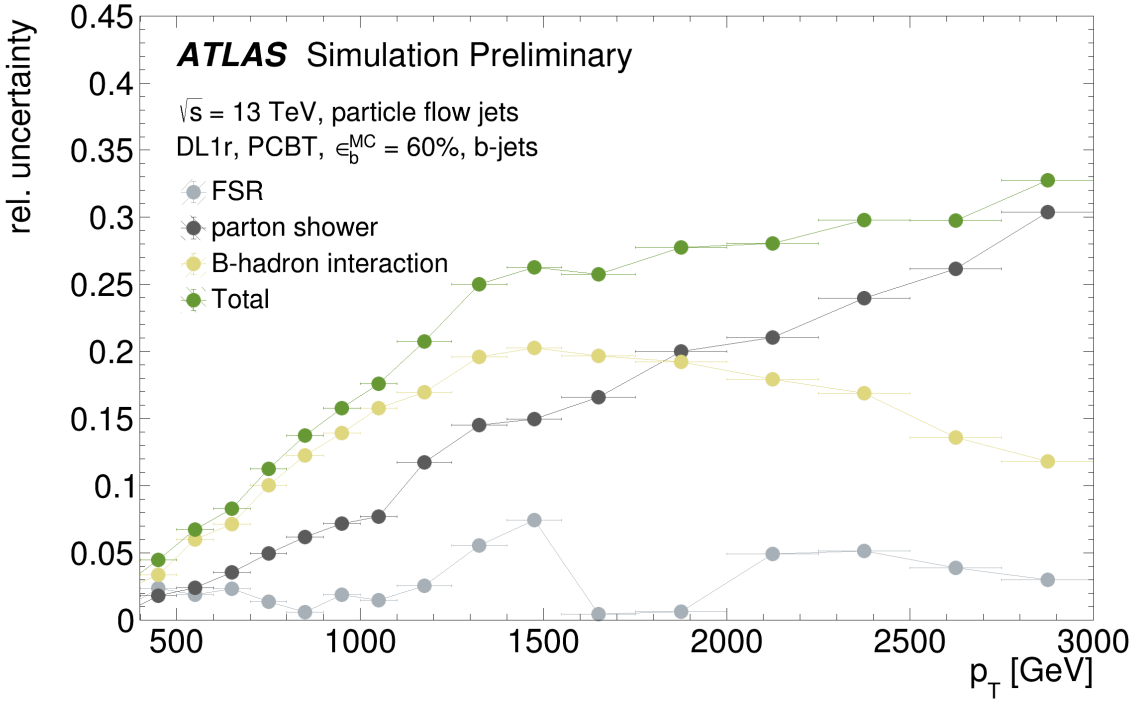


Figure 7.3 Decomposed modelling uncertainty: FSR (light grey), B-hadron interaction (yellow), parton shower (dark grey), and total (green) for PFlow b-jets with DL1r at 60% efficiency.

7.1.4 Global Uncertainty

As discussed in Section 7.1.3 the modelling uncertainties are the dominant contributors for the extrapolation uncertainty, this is confirmed in Figure 7.4.

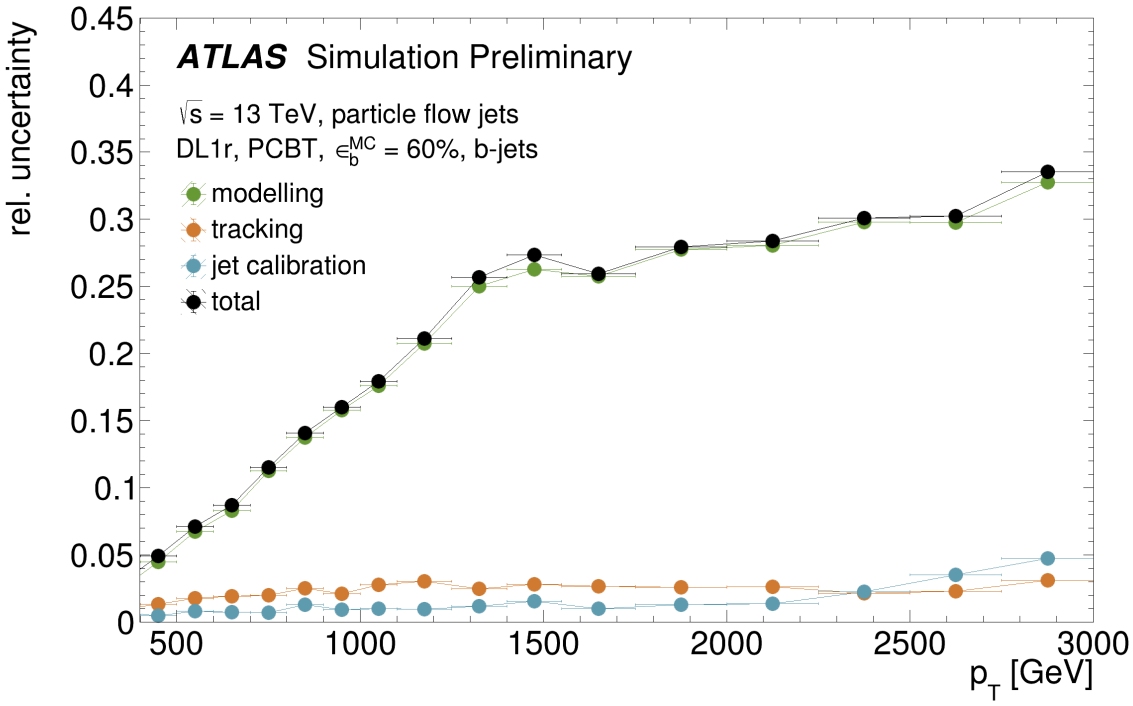


Figure 7.4 Nominal global uncertainty: modelling (green), tracking (orange), jet calibration (blue), and total (black) for PFlow b-jets with DL1r at 60% efficiency.

7.2 Eigenvector Decomposition

Folding the full set of extrapolated uncertainties into any analysis would result in a very bloated fit model, as we would have several nuisance parameters, per flavour of jet. To get around this issue, we can construct a matrix containing the contribution of each uncertainty per p_T bin for one flavour of jet. Conducting an eigenvector (EV) decomposition on this matrix, resulting in a set of EVs were the leading EVs are encoded with the majority of the information from the initial matrix, selecting these leading EVs while pruning the rest folding these into our analysis would greatly reduce the bloat in our fit model while still capturing effects of the extrapolated uncertainties.
Let us denote by N the number of p_T bins under consideration, and by M the number

of systematic uncertainty sources (e.g., tracking, pile-up, modelling, etc.) for one flavour of jet. The uncertainty contributions can be represented in a matrix

$$\mathbf{U} \in \mathbb{R}^{N \times M},$$

where the element u_{ik} represents the contribution of the k-th systematic source to the uncertainty in the i-th p_T bin. Each column $\mathbf{u}_k$ of $\mathbf{U}$ encodes the shape of the k-th uncertainty across all bins.

Covariance Matrix Construction

To study the correlations between the total uncertainties across p_T bins, we construct the covariance matrix

$$\mathbf{V} = \mathbf{U}\mathbf{U}^\top \in \mathbb{R}^{N \times N},$$

with components

$$V_{ij} = \sum_{k=1}^{M} u_{ik} u_{jk}.$$

Eigenvector Decomposition

We perform an eigenvector decomposition (EVD) on the covariance matrix $\mathbf{V}$:

$$\mathbf{V} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top,$$

where:

- $\mathbf{Q} = [\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_N]$ is an orthogonal matrix containing the eigenvectors of $\mathbf{V}$,
- $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1^2, \ldots, \lambda_N^2)$ is the diagonal matrix of eigenvalues.

Each eigenvector satisfies:

$$\mathbf{V}\vec{v}_k = \lambda_k^2 \vec{v}_k,$$

and the orthonormality condition

$$\vec{v}_i^{\top}\vec{v}_j = \delta_{ij}, \quad \Rightarrow \quad \mathbf{Q}^{\top}\mathbf{Q} = \mathbf{I}.$$

We then define a new matrix of eigenvector uncertainties:

$$\tilde{\mathbf{U}} = \mathbf{Q}\mathbf{\Lambda}^{1/2},$$

with elements

$$\tilde{u}_{ik} = \lambda_k v_{ik}.$$

Variance Preservation and Truncation

The total variance in each bin i is preserved under this transformation:

$$\sum_{k=1}^{M} u_{ik}^2 = \sum_{k=1}^{N} \tilde{u}_{ik}^2 = V_{ii}.$$

Since the eigenvalues λ_k^2 are typically ordered such that $\lambda_1^2 \geq \lambda_2^2 \geq \cdots \geq \lambda_N^2$, most of the information is contained in the first few eigenvectors. Therefore, we can truncate the expansion by retaining only the leading $L < N$ components:

$$\mathbf{V} \approx \sum_{k=1}^{L} \lambda_k^2 \vec{v}_k \vec{v}_k^{\top}.$$

This yields a reduced eigenvector uncertainty matrix:

$$\tilde{\mathbf{U}}_{\text{reduced}} = \mathbf{Q}_{(L)}\mathbf{\Lambda}_{(L)}^{1/2},$$

where $\mathbf{Q}_{(L)}$ contains only the first L eigenvectors, and $\mathbf{\Lambda}_{(L)}$ the corresponding eigenvalues.

In practice, this technique allows us to include only the dominant L orthogonal modes as nuisance parameters in the analysis, dramatically reducing the dimensionality of the fit model while faithfully preserving the core features of the extrapolated uncertainty space.

7.3 Post Truncation Validation

After some discussion, with the internal flavour tagging group of ATLAS (FTAG), the number of EVs to keep after truncation were settled on per flavour and are shown in table 7.2. The question remains: if the number of EVs retained contain enough information to effectively encapsulate uncertainties evaluated thus far.

Flavour	Number of eigenvectors kept after truncation, L
b	4
c	2
$light$	4

Table 7.2 Number of eigenvectors kept after truncation per flavour.

Figures 7.5, 7.6 and 7.7 show the full uncertainties for each p_T bin, from immediately after calibration to 3 Tev, for each PCBT WP, (black) and the uncertainties evaluated only using the truncated set of EVs (red). As we see in Figure 7.5, referring to b-jets, there is a max loss of information of $\approx 6\%$. However, this is a 6% loss of an overall uncertainty value of ≈ 5, leading to an overall precision loss of < 0.3 in our final uncertainty value. Similarly, in Figure 7.6, referring to c-jets, we see a maximum loss of information of $\approx 4\%$ of an original uncertainty value of 10 leading to an overall loss of 0.4. Finally, Figure 7.7, referring to light-jets, we see a maximum information loss of $\approx 20\%$. This is much larger than the previous flavours, however following the same logic as before, this is a $\approx 20\%$ loss of 16 resulting in an overall loss of 3.2 from our final uncertainty. The high level takeaway from this discussion is that: The bins immediately after calibration suffer the most from information loss in the truncated regime across all flavours. However, the magnitude of their information loss is minimal due to the relative low value of the raw uncertainty, hence considered acceptable.

After considering the information loss with the truncated set of EVs, the loss was deemed minor and FTAG accepted the recommendation to reduce the number of EVs detailed in 7.2. With this recommendation in hand, FTAG propagated this work into the final CDI of release 21 and so any analysis using PFlow jets and the PCBT methods with high-p_T regions could benefit from this extrapolation method and will provide more reliable flavour-tagging uncertainties past the calibration threshold.

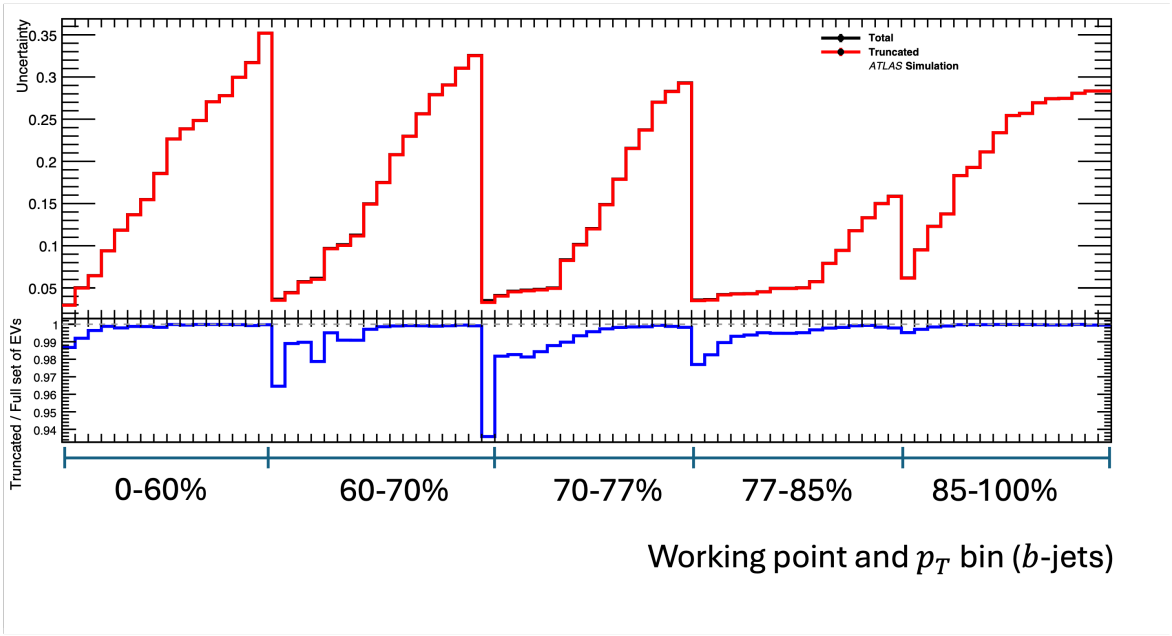


Figure 7.5 b-jets: The uncertainty for a particular p_T bin after calibration and WP up to 3 TeV and the ratio of total (black) and pruned (red) eigenvectors, we keep 4 eigenvectors after the pruning stage.

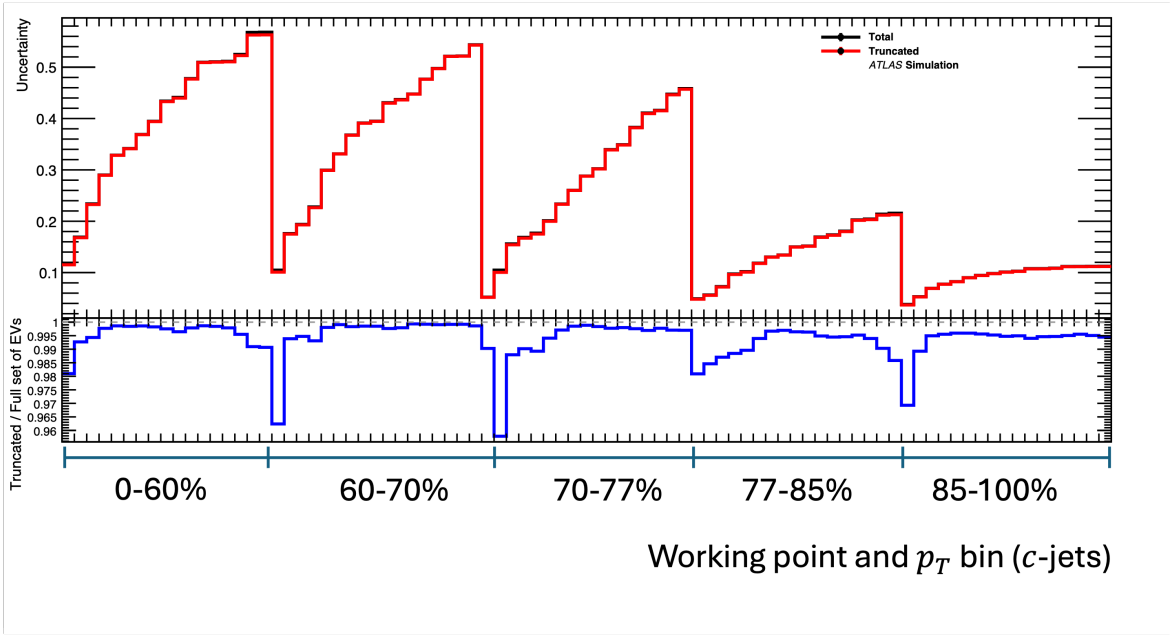


Figure 7.6 c-jets: The uncertainty for a particular p_T bin after calibration and WP up to 3 TeV and the ratio of total (black) and pruned (red) eigenvectors, we keep 2 eigenvectors after the pruning stage.

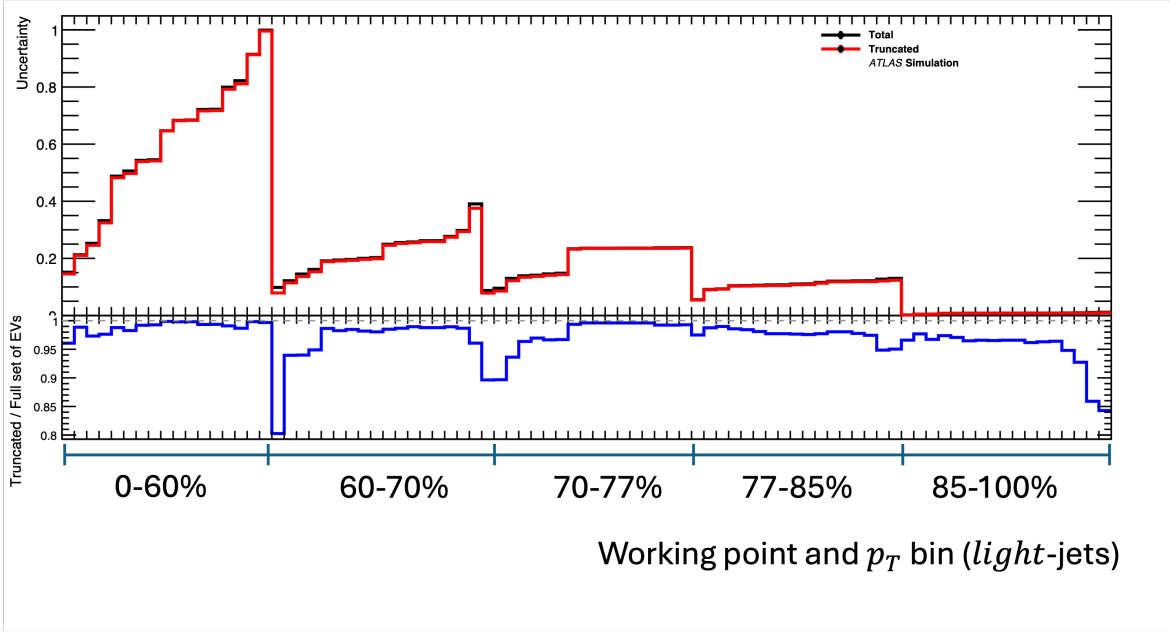


Figure 7.7 *light*-jets: The uncertainty for a particular p_T bin after calibration and WP up to 3 TeV and the ratio of total (black) and pruned (red) eigenvectors, we keep 4 eigenvectors after the pruning stage.

Chapter 8

Analysis Techniques

"Hesitation is defeat, my dear Sekiro."
— Isshin Ashina

Thus far in this thesis, we have covered the theory of particle physics, how we detect and reconstruct the particles and the MC techniques. Yet some questions still remain, such as: How do we separate events the into a signal and background event? And how do we construct a statistically robust framework, using our signal and background events? To answer the first question, usually a multivariate analysis incorporating a machine learning technique is used. For the $t\bar{t}H(b\bar{b})$ analysis a multi-class neural network based on a permutation-invariant transformer architecture with attention mechanism is used [119]. Using the output of this framework a combined profile likelihood fit that simultaneously determines the event yields for the signal and for the most important background components, while constraining the overall background model within the assigned systematic uncertainties is constructed.

This chapter will serve as a general introduction to the transformer architecture and profile likelihood fit construction before we apply these concepts in Chapter 9.

8.1 Profile Likelihoods

8.1.1 Poisson Distributions

In each pp collision at the LHC, the final state may originate from the signal process or from a variety of background processes. These cannot be separated event-by-event with perfect accuracy, but their relative contributions can be inferred statistically using the measured distributions in signal- and control-enriched regions. The data in these regions are organized into histograms of discriminant variables, such as the output of a neural network classifier or a reconstructed invariant mass or Higgs boson p_T. Each of these events has a small probability of producing the final state of interest. This matches the conditions under which the Poisson distribution applies: rare independent trials with a constant rate. The Poisson model correctly describes both low-statistics bins and high-statistics bins (where it smoothly approximates a Gaussian distribution). Therefore, the likelihood in each bin is taken as the Poisson probability of the observed count given the predicted mean from signal and background. We define the Poisson distribution as:

$$P(n|\nu) = \frac{\nu^n e^{-\nu}}{n!}, \tag{8.1}$$

where ν is the expected mean yield in the bin, determined from the signal and background model. n is the observed number of events in the bin. $e^{-\nu}$ ensures that the total probability over all $n \in \{0, 1, 2, \dots\}$ sums to one. The factorial $n!$ accounts for the number of distinct permutations of n indistinguishable events.

8.1.2 The Likelihood Function

The likelihood function, denoted $\mathcal{L}(\boldsymbol{\Theta})$, quantifies the probability of observing the data x given a set of model parameters $\boldsymbol{\Theta}$. In other words, it measures the degree to which a model, parametrised by $\boldsymbol{\Theta}$, agrees with the observed data. For a set of n statistically independent measurements $x_1, x_2, \dots, x_n$, each described by a probability density function $f(x; \boldsymbol{\Theta})$, the likelihood is given by:

$$\mathcal{L}(\boldsymbol{\Theta}) = \prod_{i=1}^{n} f(x_i \mid \boldsymbol{\Theta}). \tag{8.2}$$

The assumption of independence allows the likelihood to factorise as a product of the individual PDFs, such that:

$$P(x_1, x_2, \ldots, x_n|\Theta) = \prod_{i=1}^{n} P(x_i|\Theta). \tag{8.3}$$

In practice, it is more convenient to work with the logarithm of the likelihood. Taking the logarithm converts the product into a sum:

$$\ln \mathcal{L}(\boldsymbol{\Theta}) = \sum_{i=1}^{n} \ln f(x_i|\boldsymbol{\Theta}). \tag{8.4}$$

It is standard in high-energy physics to minimise the negative log-likelihood:

$$-\ln \mathcal{L}(\boldsymbol{\Theta}) = -\sum_{i=1}^{n} \ln f(x_i|\boldsymbol{\Theta}), \tag{8.5}$$

since optimisation algorithms are typically designed to locate minima of functions. Minimising the negative log-likelihood is mathematically equivalent to maximising the likelihood[i].

8.1.3 The Likelihood Function With Binned Data

For the $t\bar{t}H(b\bar{b})$ legacy analysis in Chapter 9 the model parameters are: signal strength μ, normalisation factors k and the set of Nuisance Parameters (NPs) θ (these are usually modelled as Gaussian priors), which we use to describe the effects of systematic uncertainties in the expectations and treated as extra degrees of freedom inside the fit. For a binned distribution, the number of expected events for bin i can be expressed using the experimentally observed number of events in bin i, given by:

$$N_i^{\text{exp}}(\mu, k, \theta) = \sum_s \mu_s \cdot N^s_{\text{Sig},i}(\theta) + \sum_b k_b \cdot N^b_{\text{Bkg},i}(\theta), \tag{8.6}$$

where $N^s_{\text{Sig},i}$ is the number of expected signal events from a signal process s in bin i. $N^b_{\text{Bkg},i}$ is the number of background events for a given background process b in bin i. μ_s and k_b are the signal strength and background normalisation factor for a given background process. These background normalisation factor(s) can either be set to one, assuming exactly SM like behaviour or left to freely float with an initial value of one, allowing

[i]Its always easier to differentiate a sum than a product.

the fit to then adapt to observed and expected number of events. The signal strength is given by:

$$\mu_s = \frac{\sigma^s}{\sigma^s_{\mathrm{SM}}}, \tag{8.7}$$

where σ^s is the observed cross-section of the measured signal process, s and σ^s_{SM} is the expected SM cross-section for the same process. When quoting signal strength for an inclusive measurement only one μ is used. For a Simplified Template Cross-Section[ii] (STXS) measurement a μ is used for each STXS region. The resulting binned likelihood function is given my:

$$\mathcal{L}(\mu, k, \theta) = \prod_i^M \frac{N_i^{\mathrm{exp}}(\mu, k, \theta)^{N_i}}{N_i!}\, e^{-N_i^{\mathrm{exp}}(\mu,k,\theta)}, \tag{8.8}$$

where M is the total number of bins and N_i is the number of events in bin i. Equation 8.7 shows that large or small values of cross-section can lead to speculation of the presence of new physics, provided no large systematic or modelling uncertainty dominates this potential deviation. A change of bias ($\mu \rightarrow$Wilson Coefficient) in Equation 8.8 can be used to re-interpret a SM measurement in the EFT framework discussed in Section 2.8.

8.1.4 Profile Likelihoods and Test Statistic

The profile likelihood fit is the statistical framework used to extract the best-fit value of a Parameter Of Interest (POI). Here its μ, while consistently incorporating the effects of systematic uncertainties, NPs and normalisation factors. In this approach, for each tested value of μ, the likelihood is maximized with respect to all nuisance parameters, yielding the profile likelihood:

$$\mathcal{L}(\mu) = \mathcal{L}(\mu, \hat{\hat{k}}, \hat{\hat{\theta}}), \tag{8.9}$$

where $\hat{\hat{\theta}}$ is the value of a given NP that maximizes the likelihood for a fixed value of μ. The value of μ that maximizes the full likelihood function is the best-fit signal strength $\hat{\mu}$, and the shape of the profile likelihood around $\hat{\mu}$ determines the statistical uncertainty.

[ii] This is an agreed framework that partitions Higgs production processes into kinematic regions to enable precise, model-independent, measurements of Higgs properties.

The ratio of these two likelihood methods is given by λ_μ:

$$\lambda_\mu = \frac{\mathcal{L}(\mu, \hat{\hat{k}}, \hat{\hat{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{k}, \hat{\theta})} \tag{8.10}$$

To quantify the compatibility of different μ values with the observed data, the test statistic is defined as:

$$q_\mu = -2 \ln \lambda_\mu \tag{8.11}$$

This test statistic is used both to construct confidence intervals for μ, and to quantify the level of compatibility between the observed data and the hypothesis parameter setting μ. Larger values of q_μ correspond to observed data is less compatible with the hypothesised value of μ.

8.1.5 Significance

The deviation from the background-only hypothesis is quantified by the signal significance S, defined as:

$$S = \sqrt{\lambda_{q_0}}, \tag{8.12}$$

where λ_0 is the likelihood ratio evaluated at $\mu = 0$ (background only). The significance expresses how strongly the observed result deviates from the background-only expectation, typically quoted in units of standard deviations (σ). In particle physics, two common thresholds are used: a significance of at least 3σ is referred to as evidence, corresponding to a p-value of approximately 0.3%, while a significance of 5σ is required for a discovery, corresponding to a p-value of about 3×10^{-7}. Here, the p-value represents the probability of observing a deviation as large as, or larger than, the measured result under the assumption that the background-only hypothesis is true.

8.1.6 Nuisance Parameter Pulls

NPs nominally have a central value of $\theta = 0$, describing the best possible value associated with a given systematic uncertainty. If this central value were to shift we call this a pull

in our fit, which is described as:

$$\text{Pull} = \frac{\hat{\theta} - \theta_0}{\Delta\theta}, \tag{8.13}$$

where $\hat{\theta}$ is the fitted value of the NP, θ_0 is the nominal (pre-fit) value and $\Delta\theta$ is the uncertainty of that NP (pre-fit or post-fit, depending on convention). These pulls can indicate potentially problematic modelling parameters or tensions in the fit. As a diagnostic tool a significant pull is usually the first step in any statistical detective work when trying to understand problems in the fit.

8.1.7 Asimov Data Sets

An Asimov fit is performed using a specially constructed dataset where each bin contains exactly the expected yield from the model without any statistical fluctuations. This dataset represents the median expectation under a given hypothesis (often the SM signal strength $\mu = 1$ or the background-only case $\mu = 0$). Fitting the Asimov dataset allows one to evaluate the expected performance of the analysis, such as the anticipated uncertainty on the signal strength, the ranking of nuisance parameters, and the pulls on NPs, in the absence of statistical noise. By comparing the results of an Asimov fit to the fit on real data, one can identify deviations in nuisance parameter pulls or impacts that may indicate tensions between data and the model or underestimated systematic uncertainties. These comparisons are another key tool in the statistical detective's repertoire.

8.1.8 Tools Used

To use the statistical framework we have described thus far a number of tools are used. Namely tools provided by the RooStat framework [120, 121], which is applied using the TRExFitter framework, and is based on the HistFactory framework [122]. The minimisation is performed using the Minuit library implemented in ROOT [123, 124].

8.2 Transformer Neural Networks

HEP experiments produce vast and complex datasets describing the properties and interactions of subatomic particles. Extracting meaningful patterns from these data has

long been a central challenge, historically addressed using carefully engineered features and domain-specific algorithms. In recent years, advances in machine learning (ML) have transformed the analysis landscape, enabling models to learn directly from raw detector outputs and simulation data. Techniques such as boosted decision trees, convolutional neural networks (CNNs), and graph neural networks (GNNs) have all demonstrated remarkable performance in classification, reconstruction, and anomaly detection tasks.

Transformers represent the next stage in this evolution. Originally developed for natural language processing, transformers replace sequential recurrence with a global attention mechanism, allowing the model to capture long-range and complex dependencies between input elements. This architecture is highly flexible: by treating the input as a set of feature vectors rather than an ordered sequence, transformers can naturally operate in a permutation-invariant fashion.

In this section, we present the core components of the transformer architecture, with emphasis on the attention mechanism and its adaptation to HEP data. We describe how input particle features are embedded, how self and cross-attention layers propagate information, and how the overall architecture can be constructed to preserve physical symmetries. The goal is to illustrate why transformers are a powerful method for the classification and reconstruction problems faced by the $t\bar{t}H(b\bar{b})$ analysis. All components of this discussion are taken from [125].

8.2.1 The Attention Mechanism

The Transformer architecture uses attention mechanisms to capture connection between the input representations[iii]. These connections are derived via dot-product self-attention which transforms the input data into *queries*, *keys* and *values*, with attention weights to computed according to the feature in question. The final architecture for the transformer is represented by a stack of shallow networks with blocks of self-attention to allow for the exchange of relevant information, giving rise to a more expressive model compare to traditional neural networks. Let us begin to build the concepts for this architecture and start with dot-product self-attention.

[iii] input representations are just input features, which is our raw data, post embedding.

Dot-Product Self-Attention

The core operation of self-attention, denoted **sa**[•], is to determine the relevance of one input object to another via a similarity measure. Given a set of N input vectors $\{x_1, x_2, \ldots, x_N\}$, where $x_m \in \mathbb{R}^{D\times 1}$,[iv] we first compute a *query* vector q_n for the object and a *key* vector k_m for each object in the set:

$$q_n = b_q + \Omega_q x_n, \quad k_m = b_k + \Omega_k x_m, \tag{8.14}$$

where $\Omega_q, \Omega_k \in \mathbb{R}^{D\times D}$ are learnable weight matrices and $b_q, b_k \in \mathbb{R}^D$ are biases. The layman's understanding of the query is simply "what this input object is looking for" in the other input objects. Similarly, a key can be thought of as "what an input object has to offer" for comparison, while the value is the actual information that will be shared if that object is deemed relevant. The similarity between the query of one object and the key of another is computed using the scaled dot product between q_n and k_m:

$$s_{nm} = \frac{q_n^\top k_m}{\sqrt{d_k}}, \tag{8.15}$$

where d_k is the dimensionality of the keys. The scaling factor $1/\sqrt{d_k}$ prevents the dot products from becoming large in magnitude, which would otherwise push the softmax into regions with small gradients[v]. These similarity scores form the basis for computing attention weights, which determine how strongly each input object incorporates information from all others.

Computing the Attention Weights

The raw similarity scores s_{nm} quantify how closely the query vector of object n matches the key vector of object m. However, these scores are unbounded real numbers and cannot be used directly to weight the values. To convert them into a set of interpretable, non-negative coefficients that sum to one, we apply the softmax function across the m

[iv] In the context of the $t\bar{t}H(b\bar{b})$ analysis the inputs will be particle objects such as p_T, mass, PCBT bin and so on.

[v] The scaling factor $1/d_k$ prevents dot products from growing with d_k, which would otherwise drive the softmax into saturation, producing vanishing gradients through the attention weights.

index for each fixed n:

$$a[x_m, x_n] = \frac{\exp(s_{nm})}{\sum_{j=1}^{N} \exp(s_{nj})}. \tag{8.16}$$

The softmax function has two important effects: It exponentiates the scores, which accentuates differences between high and low similarity values, thereby sharpening the focus of the attention mechanism. It also normalises the set of scores for each n so that they form a valid probability distribution over the N input objects. These weights satisfy:

$$a[x_m, x_n] \geq 0, \quad \sum_{m=1}^{N} a[x_m, x_n] = 1, \tag{8.17}$$

meaning they can be interpreted as the fraction of "attention" that the updated representation of object n assigns to the information coming from object m. A large value of $a[x_m, x_n]$ indicates that the features of m are highly relevant to updating n, whereas a small value indicates limited influence. In the context of the $t\bar{t}H(b\bar{b})$ analysis, these weights encode learned, event-specific relationships between different reconstructed objects, enabling the network to dynamically decide which inputs are most informative for a given event interpretation.

From Dot-Product Attention to the Self-Attention Block

In addition to the queries and keys, each input object x_m is also projected into a value vector:

$$v_m = b_v + \Omega_v x_m, \tag{8.18}$$

where $\Omega_v \in \mathbb{R}^{D\times D}$ and $b_v \in \mathbb{R}^D$ are learnable parameters. This transformation is analogous to the affine mapping in a standard feed-forward layer, but here no non-linear activation is applied. The values represent the raw information content associated with each input object that will be shared with others once the attention weights have been determined. The output representation for object n is then computed as a weighted sum over the value vectors of all N input objects:

$$\mathrm{sa}_n[x_1, \ldots, x_N] = \sum_{m=1}^{N} a[x_m, x_n]\, v_m. \tag{8.19}$$

Here, $a[x_m, x_n]$ controls how much of v_m is incorporated into the updated representation of x_n. A high attention weight means that the information from object m is strongly relevant to updating n, while a low weight means it contributes very little. In this way, the self-attention block enables each object to gather contextual information from every other object in the event, producing context-aware representations that are dynamically tailored to the specific event being processed.

8.2.2 Multi-Headed Attention and Matrix Formulation

Up to this point, the description of self-attention has been expressed in terms of individual input objects and their corresponding query, key, and value *vectors*. While this view is useful for building intuition, in practice the computation is implemented more efficiently using a *matrix* representation. Instead of processing each input x_m separately, all N input representations are arranged into a single matrix:

$$X = \begin{bmatrix} x_1^\top \\ x_2^\top \\ \vdots \\ x_N^\top \end{bmatrix} \in \mathbb{R}^{N \times D}. \tag{8.20}$$

The queries, keys, and values for all inputs are then computed in parallel by multiplying X with learned weight matrices and adding a basis matrix:

$$\mathbf{Q} = \mathbf{X}\,\mathbf{\Omega}_Q + \mathbf{1}\,\mathbf{b}_Q^\top, \quad \mathbf{K} = \mathbf{X}\,\mathbf{\Omega}_K + \mathbf{1}\,\mathbf{b}_K^\top, \quad \mathbf{V} = \mathbf{X}\,\mathbf{\Omega}_V + \mathbf{1}\,\mathbf{b}_V^\top, \tag{8.21}$$

where $W_Q, W_K, W_V \in \mathbb{R}^{D \times d_k}$ are learnable parameter matrices. Here, $\mathbf{Q}$, $\mathbf{K}$, and $\mathbf{V}$ have shapes $\mathbb{R}^{N \times d_k}$, and each row corresponds to the query, key, or value of one input object. The attention mechanism in matrix form can then be written as:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\,\mathbf{K}^\top}{\sqrt{d_k}}\right)\mathbf{V}, \tag{8.22}$$

where the matrix $\mathbf{Q}\,\mathbf{K}^\top \in \mathbb{R}^{N \times N}$ contains all pairwise similarity scores between queries and keys, the softmax is applied row-wise, and the resulting attention weight matrix multiplies $\mathbf{V}$ to produce the updated representations.

Multi-Headed Attention

A single attention head uses one set of parameters $(\mathbf{\Omega}_Q, \mathbf{\Omega}_K, \mathbf{\Omega}_V, \mathbf{b}_Q, \mathbf{b}_K, \mathbf{b}_V)$, which limits it to learning one type of relationship at a time. *Multi-headed attention* overcomes this by introducing H parallel attention heads, each with its own parameters:

$$Q^{(h)} = XW_Q^{(h)}, \quad K^{(h)} = XW_K^{(h)}, \quad V^{(h)} = XW_V^{(h)}, \quad h = 1, \ldots, H. \tag{8.23}$$

Each head computes its own attention output:

$$\text{head}_h = \text{Attention}\left(\mathbf{Q}^{(h)}, \mathbf{K}^{(h)}, \mathbf{V}^{(h)}\right). \tag{8.24}$$

These outputs are concatenated along the feature dimension and projected back to the model dimension D:

$$\text{MultiHead}(\mathbf{X}) = \text{Concat}\left(\text{head}_1, \ldots, \text{head}_H\right) \mathbf{W}_O, \tag{8.25}$$

where $\mathbf{W}_O \in \mathbb{R}^{(H \cdot d_v) \times D}$ is a learned projection matrix. Multi-headed attention allows each head to specialise, some may focus on local relationships between similar objects, while others capture global event-wide correlations. By combining these complementary perspectives, the transformer can model complex dependencies more effectively than with a single attention head.

Cross-Attention

Self-attention operates on a single set of input representations, using the same source for queries, keys, and values. In some cases, however, it is necessary to relate information from *two distinct sets* of representations. This is achieved with *cross-attention*. Let $\mathbf{X}^{(Q)} \in \mathbb{R}^{N_Q \times D}$ denote the matrix of query-side input representations (the "target" sequence or set) and $\mathbf{X}^{(K,V)} \in \mathbb{R}^{N_K \times D}$ the matrix of key/value-side representations (the "source" sequence or set). The query set might represent, for example, the intermediate hidden states of a decoder network, while the key/value set could represent encoded detector data or reconstructed event objects. The projections for cross-attention are defined as:

$$\mathbf{Q} = \mathbf{X}^{(Q)}\,\mathbf{\Omega}_Q + \mathbf{1}\,\mathbf{b}_Q^\top, \quad \mathbf{K} = \mathbf{X}^{(K,V)}\,\mathbf{\Omega}_K + \mathbf{1}\,\mathbf{b}_K^\top, \quad \mathbf{V} = \mathbf{X}^{(K,V)}\,\mathbf{\Omega}_V + \mathbf{1}\,\mathbf{b}_V^\top. \tag{8.26}$$

Note that $\mathbf{Q}$ is derived exclusively from the query-side inputs, while $\mathbf{K}$ and $\mathbf{V}$ are derived from the key/value-side inputs. The cross-attention operation then proceeds exactly as in self-attention:

$$\text{CrossAttention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\,\mathbf{K}^{\top}}{\sqrt{d_k}}\right)\mathbf{V}. \tag{8.27}$$

In this setting:

- Each query object attends to *all* objects in the key/value set.
- The attention weights represent the relevance of each key/value-side object to the query-side object.
- The resulting output is an updated version of the query-side representations, enriched with information drawn from the key/value side.

In the context of the $t\bar{t}H(b\bar{b})$ analysis, cross-attention provides a mechanism to relate two distinct representations of the event. The *query-side* inputs can represent high-level event or hypothesis-level embeddings, such as a candidate $t\bar{t}H$ system formed from a preliminary aggregation of event features. The *key/value-side* inputs are the set of reconstructed objects in the event for example, jets, leptons, and missing transverse energy each described by its own learned representation.

In this setup, the query vectors pose the question: *"Which reconstructed objects are most relevant for refining my current event-level hypothesis?"* The keys encode each object's properties, allowing the network to determine their relevance, and the values carry the information that will be incorporated into the updated event-level representation.

Through training, the attention mechanism learns to focus on objects that are most informative for distinguishing $t\bar{t}H(b\bar{b})$ signal events from background processes, such as identifying b-jets from the Higgs boson decay or leptons from top quark decays. In this way, cross-attention enables a targeted, learned transfer of information from the object level to the hypothesis level, improving the network's discriminative power.

Multi-Headed Cross-Attention

As with self-attention, cross-attention can be extended to multiple heads, with each head using its own parameter set:

$$\mathbf{Q}^{(h)} = \mathbf{X}^{(Q)}\,\mathbf{\Omega}_Q^{(h)} + \mathbf{1}\,\mathbf{b}_Q^{(h)\top}, \quad \mathbf{K}^{(h)} = \mathbf{X}^{(K,V)}\,\mathbf{\Omega}_K^{(h)} + \mathbf{1}\,\mathbf{b}_K^{(h)\top}, \quad \mathbf{V}^{(h)} = \mathbf{X}^{(K,V)}\,\mathbf{\Omega}_V^{(h)} + \mathbf{1}\,\mathbf{b}_V^{(h)\top}. \tag{8.28}$$

Each head produces:

$$\text{head}_h = \text{CrossAttention}\left(\mathbf{Q}^{(h)}, \mathbf{K}^{(h)}, \mathbf{V}^{(h)}\right), \tag{8.29}$$

and the heads are concatenated and projected back to the model dimension:

$$\text{MultiHeadCross}(\mathbf{X}^{(Q)}, \mathbf{X}^{(K,V)}) = \text{Concat}\left(\text{head}_1, \ldots, \text{head}_H\right)\,\mathbf{W}_O. \tag{8.30}$$

Multi-headed cross-attention allows the network to learn multiple complementary ways of relating the two input sets, which is especially useful when the sources encode different modalities or levels of abstraction. In the $t\bar{t}H(b\bar{b})$ analysis, each attention head can specialise in a different aspect of the relationship between the event-level hypothesis and the reconstructed object set. For example, one head may focus on identifying b-jets consistent from a Higgs boson decay, another may emphasise correlations between leptons and jets from top quark decays, while a third could capture global event features such as missing transverse momentum patterns. By combining these complementary attention patterns, multi-headed cross-attention allows the network to integrate diverse and potentially weak signals into a coherent, high-quality event-level representation, improving discrimination between $t\bar{t}H(b\bar{b})$ signal and background events.

Input Embedding

Before an event's reconstructed objects can be processed by the attention mechanism, their raw features must be transformed into a common *embedding space* of dimension D. Let $\mathbf{f}_m \in \mathbb{R}^F$ denote the raw feature vector of object m, where F is the number of input features (e.g., p_T, mass, η, ϕ, particle identification scores, categorical bins such as PCBT, etc.). These features are mapped into the model space via a learned affine

transformation:

$$\mathbf{x}_m = \mathbf{b}_{\text{emb}} + \mathbf{\Omega}_{\text{emb}}\, \mathbf{f}_m, \tag{8.31}$$

where $\mathbf{\Omega}_{\text{emb}} \in \mathbb{R}^{F\times D}$ is the embedding **basis matrix** and $\mathbf{b}_{\text{emb}} \in \mathbb{R}^{D}$ is a bias vector. The result $\mathbf{x}_m \in \mathbb{R}^{D}$ is the initial *input representation* of object m that will serve as the starting point for subsequent attention layers. This embedding step serves several purposes:

- It projects heterogeneous input features into a uniform representation space, enabling efficient batch processing.
- It allows the network to learn optimal combinations of raw features, rather than relying on fixed physics-derived variables.
- It provides the dimensionality D required for compatibility with the learned basis matrices in the query, key, and value projections.

In the $t\bar{t}H(b\bar{b})$ analysis, this means that jets, leptons, and other reconstructed objects are all expressed in the same learned representation space, allowing the transformer to treat them uniformly while still preserving the physics content of their raw features.

8.2.3 The Transformer Architecture

The transformer architecture consists of stacked layers that combine multi-headed attention with position-wise feed-forward networks, operating on a set of input embeddings, each mapped to a D-dimensional embedding vector via a learned affine transformation. These embeddings form the input matrix $\mathbf{X}$ to the first transformer layer.

Within each layer, a multi-headed self-attention block updates each object's representation by aggregating information from all others, weighted according to learned attention scores. For tasks requiring integration between different levels of representation, such as refining an event-level hypothesis using object-level data, cross-attention is employed, with queries drawn from the hypothesis representation and keys/values from the object set.

The output of the attention block is passed through a feed-forward sublayer, which applies an affine transformation and non-linear activation independently to each object representation. Residual connections and layer normalisation are applied around both

the attention and feed-forward sublayers to improve gradient flow and stabilise training. A visual representation of this work flow is shown in Figure 8.1.

By stacking multiple such layers, the transformer progressively builds rich, context-aware representations that capture both local and global relationships within the event. The final representation can then be pooled or otherwise aggregated to produce a compact event-level feature vector, which is passed to a classification head to distinguish $t\bar{t}H(b\bar{b})$ signal events from background processes.

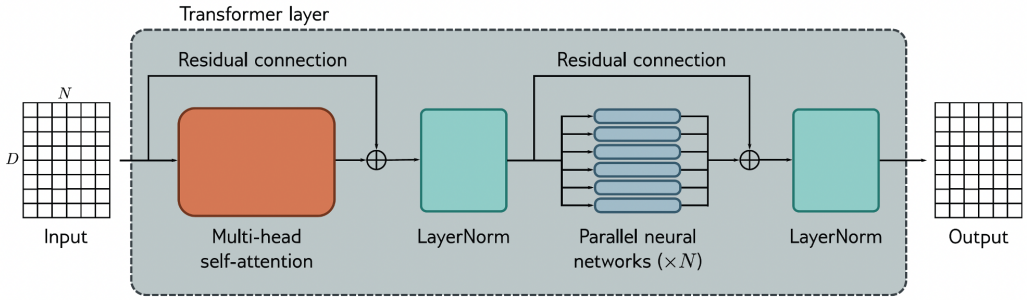


Figure 8.1 The transformer takes as input a $D \times N$ matrix, where each of the N columns is the D-dimensional embedding of an input object. The output has the same shape. Each transformer layer applies a sequence of operations. First, a multi-headed attention block allows the embeddings to exchange information, producing updated representations for each object. This is done within a residual block, so the original inputs are added back to the attention output. Second, layer normalisation is applied. Third, a second residual block applies the same fully connected feed-forward network independently to each of the N object representations. Finally, layer normalisation is applied again [125].

Chapter 9

$t\bar{t}H(b\bar{b})$ Legacy Analysis

"12 and 5"

— Dr. Elliot Watton, on whisky and burgers.

We now arrive at the first of two main results and analyses presented in this thesis; the measurement of the Higgs boson production in association with a pair of top quarks ($t\bar{t}H$), focusing on the decay channel $H \to b\bar{b}$, using data recorded by the ATLAS detector. The analysis is based on the full Run 2 dataset, collected between 2015 and 2018, corresponding to an integrated luminosity of 140 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. This same dataset was previously used to measure the $t\bar{t}H(b\bar{b})$ production cross-section, which yielded a signal strength of $\mu_{t\bar{t}H} = 0.35^{+0.36}_{-0.34}$ [16]. That result indicated a shortfall in the $H \to b\bar{b}$ decay channel compared with expectations from other Higgs decays, creating some tension that motivated a more refined re-analysis. We will refer to the current analysis as the legacy analysis and the previous analysis as the 2022 result.

9.1 Overview

Before we can delve into the legacy analysis lets recap the 2022 result, which will allow us to discuss some of its shortfalls and how they are improved for the legacy analysis. This enables the key differences in the legacy analysis to be highlighted.

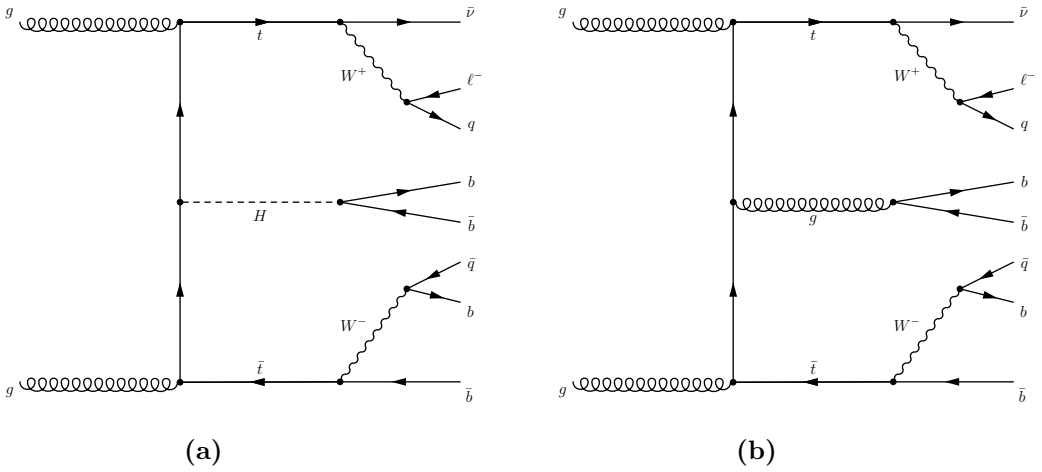


Figure 9.1 Single-lepton topology for $t\bar{t}H$ with $H \to b\bar{b}$ signal (a) and $t\bar{t}b\bar{b}$ background (b).

9.1.1 Previous Result

The 2022 result was given as an Inclusive measurement and a Simplified Template Cross-Section (STXS) result, where the signal event are split by Higgs p_T.

Inclusive Measurement

The $t\bar{t}H(b\bar{b})$ process has three main final states dictated by the decay modes of the $W^{\pm}$ bosons from the $t\bar{t}$ decay. The decay mode with the largest branching ratio (about 45.7%) is the fully hadronic channel, where both $W^{\pm}$ bosons decay into quarks. Although this mode is the most common, it is also the hardest to analyse because of the overwhelming background from QCD-driven multi-jet production. For this reason, this final state has been excluded from both the 2022 result and the legacy analysis.

The next mode is the semi-leptonic (or single-lepton) channel, in which one $W^{\pm}$ boson decays to quarks while the other decays to a lepton and a neutrino. This topology benefits from a high branching fraction (43.8%) and the presence of a charged lepton, which helps separate signal from background. Since only one neutrino appears in the final state, its momentum can be inferred from the missing transverse energy (E_T^{miss}), making the reconstruction comparatively straightforward. The sizable branching ratio also ensures good statistics and a favourable signal-to-background ratio, even at high p_T. To optimize the analysis, this channel is divided into resolved and boosted categories, where reclustered jets are employed to capture both b-quarks from the Higgs decay within

a single jet.

Finally, the dileptonic (or di-lepton) channel involves both $W^{\pm}$ bosons decaying leptonically. This channel has the smallest branching fraction (around 10.5%) and includes two neutrinos, which makes reconstruction more difficult. Nevertheless, it is still considered because the two final-state leptons strongly suppress multi-jet backgrounds, yielding the cleanest experimental signature of all the channels.

The inclusive measurement of the 2022 result used single lepton resolved and boosted channels along with a di-lepton resolved channel. The final inclusive measurement was seen to be $\mu_{t\bar{t}H} = 0.35^{+0.36}_{-0.34}$ with an observed (expected) significance 1.0 (2.7) standard deviations. The per channel breakdown can be see in Figure 9.2.

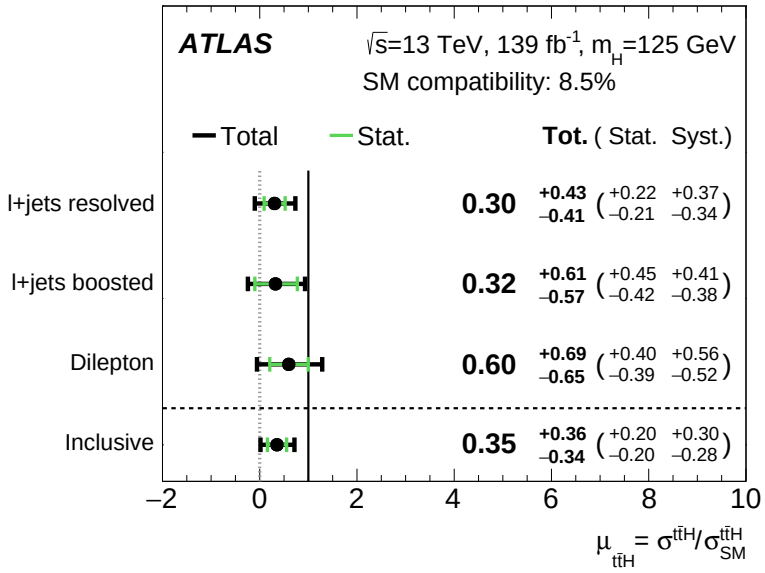


Figure 9.2 The inclusive result for 2022 result [16].

The overall uncertainty of the result is largely driven by systematic effects, with the dominant contribution originating from the modelling of the $t\bar{t}+ \geq 1b$ background. A detailed decomposition of the uncertainties and their impact on $\mu_{t\bar{t}H}$ is presented in Table 9.1.

Uncertainty source	$\Delta\mu$	
Process modelling		
$t\bar{t}H$ modelling	+0.13	−0.05
$t\bar{t}+\geq 1b$ modelling		
$t\bar{t}+\geq 1b$ NLO matching	+0.21	−0.20
$t\bar{t}+\geq 1b$ fractions	+0.12	−0.12
$t\bar{t}+\geq 1b$ FSR	+0.10	−0.11
$t\bar{t}+\geq 1b$ PS & hadronisation	+0.09	−0.08
$t\bar{t}+\geq 1b$ p_T^{bb} shape	+0.04	−0.04
$t\bar{t}+\geq 1b$ ISR	+0.04	−0.04
$t\bar{t}+c$ modelling	+0.03	−0.04
$t\bar{t}+light$ modelling	+0.03	−0.03
tW modelling	+0.08	−0.07
Background-model statistical uncertainty	+0.04	−0.05
b-tagging efficiency and mis-tag rates		
b-tagging efficiency	+0.03	−0.02
c-mis-tag rates	+0.03	−0.03
l-mis-tag rates	+0.02	−0.02
Jet energy scale and resolution		
b-jet energy scale	+0.00	−0.01
Jet energy scale (flavour)	+0.01	−0.01
Jet energy scale (pile-up)	+0.00	−0.01
Jet energy scale (remaining)	+0.01	−0.01
Jet energy resolution	+0.02	−0.02
Luminosity	+0.01	0.00
Other sources	+0.03	−0.03
Total systematic uncertainty	+0.30	−0.28
$t\bar{t}+\geq 1b$ normalisation	+0.04	−0.07
Total statistical uncertainty	+0.20	−0.20
Total uncertainty	+0.36	−0.34

Table 9.1 A breakdown of the contributions to the uncertainty in μ is shown from the 2022 result [16]. The impact of each source of uncertainty is evaluated after performing the fit. Showing the overall uncertainties are dominated by the $t\bar{t}+\geq 1b$ modelling.

Simplified Template Cross-Section

The STXS framework [126, 127] was developed in close collaboration between theorists and the LHC experiments. It enables the measurement of signal strengths in exclusive

kinematic regions of phase space, referred to as STXS bins. These bins are designed to minimise dependence on theoretical modelling assumptions while facilitating the combination of results across different experiments and analysis channels. In Higgs boson measurements, the framework is particularly useful because many decay channels can be fully reconstructed, allowing the transverse momentum of the Higgs boson (p_T^H) to be measured. Accordingly, STXS bins are typically defined as a function of p_T^H. To ensure compatibility across measurements, the classification is based on the truth-level p_T^H ($\hat{p}_T^H$) from Monte Carlo predictions rather than the reconstructed p_T^H. For the measurement of $t\bar{t}H$, the LHC experiments have agreed upon the common set of STXS bins listed in Table 9.2.

STXS Bin	$\boldsymbol{p_T^H}$ **range [GeV]**
STXS1	$0 < p_T^H \leq 60$
STXS2	$60 < p_T^H \leq 120$
STXS3	$120 < p_T^H \leq 200$
STXS4	$200 < p_T^H \leq 300$
STXS5	$300 < p_T^H \leq 450$
STXS6	$p_T^H > 450$

Table 9.2 Commonly agreed STXS bins for $t\bar{t}H$ measurements, as defined by the LHC working group.

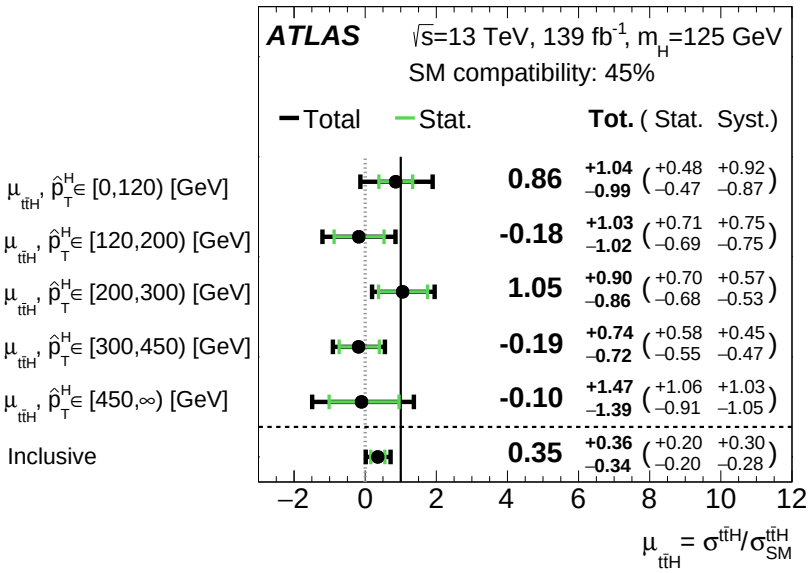


Figure 9.3 The differential result 2022 result for μ as a function of p_T^H [16].

9.1.2 Improvements on the Previous analysis

A number of significant updates have been introduced in this re-analysis compared to the most recent public ATLAS result [16]. These changes span detector reconstruction, event selection, signal discrimination, and background modelling, each contributing to improved sensitivity to the $t\bar{t}H(b\bar{b})$ process.

First, the analysis is based on `Athena` release 21, which incorporates a new ATLAS detector reconstruction framework. This includes refined particle-flow (PFlow) jet reconstruction and the DL1r b-tagging algorithm, together with updated calibrations [97]. As a consequence, the full jet energy resolution model is employed, accompanied by an improved treatment of b-tagging uncertainties, including the recommended high-p_T extrapolations.

The event pre-selections have also been relaxed to increase statistics. In the single-lepton channel, events are accepted with at least five jets ($p_T > 25$ GeV, $|\eta| < 2.5$), of which at least three must be b-tagged at the 70% working point (WP) of the DL1r algorithm. For the dilepton channel, the requirement is reduced to at least three jets, with at least three tagged at the 85% WP and two at the 70% WP. Additionally, a dedicated boosted selection is introduced in the single-lepton channel, requiring at least four jets, with three b-tagged at the 85% WP, and one reclustered jet with $p_T > 300$ GeV.

The reconstruction of the $H \to b\bar{b}$ candidate has been substantially revised. Events are categorised into six regions (signal and five control regions: $t\bar{t}$+light, $t\bar{t}c$, $t\bar{t}1b$, $t\bar{t}2b$, and $t\bar{t}B$) using a newly developed attention-based transformer model. This network exploits only the particle four-vectors (jets, leptons, and missing transverse energy) as inputs, with b-tagging WP information included. The b-jet kinematics are further corrected using soft-muon information within jets. The model is trained on both nominal and systematic variations of the signal and background samples, ensuring robustness to theoretical and detector uncertainties. A dedicated boosted signal region is also defined in the ℓ+jets channel using the transformer output, and the boosted DNN has been retrained without relying on variables sensitive to differences between `AtlFast` and full simulation, ensuring stability without degrading performance.

On the background side, major improvements have been made to the $t\bar{t} + b\bar{b}$ modelling. A consistent description based on four-flavour scheme matrix-element calculations is employed, with a corresponding systematic model that fully accounts for the theoretical uncertainties in both the $t\bar{t}b$ and $t\bar{t}$+light/$t\bar{t}c$ categories (the latter derived from inclusive $t\bar{t}$ five-flavour scheme predictions). Thanks to good pre-fit data–MC agreement, most modelling uncertainties are now correlated across analysis regions. The H_T reweighting of inclusive $t\bar{t}$ samples is applied, and the normalisations of the $t\bar{t} + c$ and $t\bar{t}$+light backgrounds are left free-floating in the fit, facilitated by the newly defined dedicated control regions.

Additional updates include the introduction of fake-lepton estimates to account for the looser pre-selections: in the ℓ+jets channel these are derived with the fake-factor method, while in the dilepton channel they are taken from simulation. The single-top background and its uncertainties are evaluated using the latest dynamic-scale samples. Furthermore, the output of the transformer classifier is used directly as the fit observable in both signal and control regions, providing maximal separation power.

Finally, the signal region definition incorporates refinements motivated by the Higgs simplified template cross-section (STXS) framework: the lowest-p_T bin is now split into the $[0, 60]$ GeV and $[60, 120]$ GeV intervals, as in STXS version 1.2, ensuring consistency with other Higgs decay channels such as $H \to \gamma\gamma$.

Taken together, these methodological developments substantially improve both the

statistical reach and the systematic control of the analysis, enabling a more precise determination of the $t\bar{t}H(b\bar{b})$ production rate than in previous studies. These improvements are detailed below.

9.2 Event Selection

There are four main stages to the event selection that define the phase space for these analyses, these stages are designed to increase the selection for possible $t\bar{t}H$ events while also reducing the background. The four main stages are: the raw ATLAS Run-2 dataset, the trigger selection, objects definitions and finally channel and region definitions. This section provides an overview on these stages.

9.2.1 ATLAS Run-2 Dataset

This analysis uses the full ATLAS Run-2 dataset. The dataset was recorded during the 2015-2018 data taking stage and consists of proton-proton collision events at a centre of mass energy $\sqrt{s} = 13$ TeV with a total integrated luminosity of 140fb^{-1}. The events considered for analysis are required to pass the ATLAS data quality criteria, meaning that at the time of recording all detector subsystems were required to be operational and performing at a high level. Runs that pass this criteria are consider for the Good Run List (GRL) which concatenates blocks of data together to be used in physics analyses.

9.2.2 Trigger Selection

The events selected for this analysis, for both the single and di-lepton channels, are required to pass a set of un-prescaled single-lepton triggers that have different p_T thresholds. Only single-lepton triggers are used in order to remain orthogonal with multi-lepton events. Un-prescaled refers to the absence of a trigger rate scale factor, so all events that pass the criteria are used. The trigger system employs a logical OR combination, allowing an event to be recorded if it satisfies at least one of the available triggers. This strategy maintains high efficiency across the relevant thresholds, ensuring, for instance, that high-p_T leptons which do not meet isolation requirements are still retained. To be selected, an event must either pass a low-p_T trigger with an isolation condition on the lepton, or a higher-p_T trigger with looser identification and no isolation

requirement. For electrons, the lowest threshold corresponds to $p_T = 24$ GeV, paired with a medium identification working point and no isolation. The highest electron threshold is $p_T = 120$ GeV, where a loose identification and no isolation are required. For muons, the lowest threshold is $p_T = 20$ GeV, with medium identification and a loose isolation requirement, while the highest threshold is $p_T = 50$ GeV, with medium identification and no isolation. Table 9.3 shows a summary of the single-lepton triggers used the Run-2 data analyses.

lepton	p_T (GeV)		Identification		Isolation	
	2015	2016–2018	2015	2016–2018	2015	2016–2018
electrons	24	26	medium	tight	–	loose
	60	60	medium	medium	–	–
	120	140	loose	loose	–	–
muons	20	26	medium	medium	loose	medium
	50	50	medium	medium	–	–

Table 9.3 Summary of the single-lepton triggers used in the analysis Run 2 data. The lower p_T thresholds and looser identification and isolation requirements present for the 2015 data are due to the decreased pile-up conditions with respect to 2016–18 data.

9.2.3 Object Definitions

Events are required to contain at least one reconstructed vertex with at least two associated tracks of $p_T > 500$ MeV. The primary vertex is chosen as the one with the largest summed squared transverse momentum of its associated tracks, provided it is consistent with the average beam-spot position.

Electrons must have $p_T > 10$ GeV and be reconstructed within $|\eta| < 2.47$, excluding the transition region $1.37 < |\eta| < 1.52$. Muons are required to have $p_T > 10$ GeV and $|\eta| < 2.5$. Electron (muon) candidates must satisfy the MEDIUM (LOOSE) identification criteria. Impact parameter requirements are imposed: $|z_0| < 0.5$ mm for both, and $|d_0|/\sigma(d_0) < 5$ (3) for electrons (muons), respectively.

Jets are reconstructed using the anti-k_T algorithm with radius $R = 0.4$ and must satisfy $p_T > 25$ GeV and $|\eta| < 2.5$. Events containing jets that fail quality requirements,

designed to reject non-collision backgrounds or detector noise, are discarded. For jets with $p_T < 60$ GeV, the tight working point of the jet vertex tagging algorithm [96] is applied to suppress pile-up contributions.

b-jets are identified with the DL1r algorithm, a neural-network method that exploits track impact parameters, displaced vertices, and jet kinematics. Each jet is assigned a pseudo-continuous score, enabling classification at multiple efficiency working points (85%, 77%, 70%, and 60%). Scale factors are applied to simulation to correct for differences between data and Monte Carlo in tagging efficiencies and mis-tag rates.

In boosted single-lepton topologies, reclustered jets are built from the small-R jets using the anti-k_T algorithm with $R = 1.0$. These jets are required to have $p_T > 200$ GeV and $|\eta| < 2.0$, an invariant mass $M > 50$ GeV, at least two constituent small-R sub-jets, and an angular separation $\Delta R > 1.0$ from all selected electrons.

9.2.4 Channel Definitions

The final stage of the event selection is categorising the channels based on the decay mode of the W-bosons produced in the top-quark decays. The relaxation of the b-tagging requirements on this iteration of the analysis, results in more signal and background events and to allow better control of their subsequent modelling, and improved statistics for the training of the machine-learning algorithms used to classify the events. The channel selection criteria are summarised in Table 9.4.

Di-lepton Channel

In the di-lepton channel, events must contain exactly two electrons or muons satisfying the baseline lepton criteria. At least one lepton is required to have $p_T > 27$ GeV and to match a trigger lepton within $\Delta R < 0.15$. The di-lepton invariant mass must exceed 15 GeV to suppress low-mass Drell-Yan backgrounds, and events with $83 < m_{\ell\ell} < 99$ GeV are excluded to reduce Z-boson contributions. In the ee channel, the sub-leading electron must have $p_T > 15$ GeV, while for $\mu\mu$ and μe channels the threshold is $p_T > 10$ GeV. Events containing hadronic τ candidates are vetoed to preserve orthogonality with other $t\bar{t}H$ analyses. Selected events must include at least three jets, with at least three b-tagged at the 85% efficiency working point and at least two at the 70% working point.

Single-lepton Channel

The single-lepton channel, events are categorized into resolved and boosted topologies. The resolved selection requires at least five jets, with a minimum of three b-tagged at the 70% efficiency working point, and exactly one electron or muon with $p_T > 27$ GeV. At most one hadronic τ candidate is allowed. The boosted selection requires at least four small-R jets, one large-R jet with $p_T > 300$ GeV, and at least three b-tagged jets at the 85% efficiency working point. Events not satisfying the boosted criteria are assigned to the resolved category. These pre-selections are designed to enhance sensitivity to $t\bar{t}H$ events with leptonic $t\bar{t}$ decays and $H \to b\bar{b}$ final states. The selection efficiency for $t\bar{t}H(b\bar{b})$ is 6.3%, corresponding to a signal acceptance more than three times larger than in the previous iteration of the analysis, while simultaneously suppressing the dominant Standard Model backgrounds.

Channel	# Jets	# b-tags (DL1r)		# e/μ	# τ_{had}	# RC Jets
		70% WP	85% WP			
dilepton	≥ 3	≥ 2	≥ 3	2	0	–
ℓ+jets resolved	≥ 5	≥ 3	≥ 3	1	≤ 1	–
ℓ+jets boosted	≥ 4	–	≥ 3	1	≤ 1	≥ 1

Table 9.4 Object selection requirements for both the di-lepton and single-lepton channels, with the latter divided into resolved and boosted categories. Shown are the minimum number of jets, the number of b-tagged jets at 70% WP and 85% WP, the number of selected electrons or muons (e/μ), the number of hadronic tau candidates (τ_{had}), and the number of RC jets. RC jets must additionally satisfy $p_T > 300$ GeV.

9.3 Event Modelling

In order to model the signal and background events, MC generators are used. The generators used to create the particular MC samples are shown in Table 9.5. This section highlights the differences of each MC sample and so provides an insight to their generation.

9.3.1 Signal Modelling

The $t\bar{t}H$ signal is modelled using the POWHEG-BOX V2 generator [128] at NLO accuracy with the NNPDF3.0NLO set of parton distribution functions [129] in the five-flavour scheme. The h_{damp} parameter is set to $\frac{3}{4}(m_t + \bar{m}_t + m_H) = 352.5$ GeV, while the renormalisation and factorisation scales are chosen as $\mu_R = \mu_F = \sqrt[3]{m_T(t)\, m_T(\bar{t})\, m_T(H)}$, where m_T denotes the transverse mass. Parton showering and hadronisation are simulated with PYTHIA 8.230 [110].

To evaluate systematic uncertainties associated with the parton shower and matching scheme, two alternative samples are produced. In one case, POWHEG-BOX V2 is interfaced with HERWIG 7.04 [86] instead of PYTHIA. In the other, the nominal generator settings are retained but interfaced with PYTHIA 8.230 using `pthard`=1, modifying the definition of the hardness scale in the matching procedure.

All $t\bar{t}H$ decay modes are generated, although only a small contribution from non-$H \to b\bar{b}$ channels survives the analysis selection. The simulated samples are normalised to the predicted inclusive $t\bar{t}H$ cross section, $\sigma_{t\bar{t}H} = 507^{+35}_{-50}$ fb, which includes both NLO QCD and electroweak corrections.

9.3.2 Background Modelling

$t\bar{t} + b\bar{b}$

To reduce the modelling uncertainties previously observed in the $t\bar{t} + b\bar{b}$ processes, dedicated samples are generated at NLO in the four-flavour scheme using the POWHEG-BOX RES generator [130] interfaced with OPENLOOPS [131]. The `NNPDF3.1nlo` $n_f = 4$ PDF set is employed, and parton showering and hadronisation are provided by PYTHIA 8.244. Both the matrix-element and shower generators are configured with specialised settings to further constrain the associated uncertainties. The renormalisation and factorisation scales are defined as:

$$\mu_R = \tfrac{1}{2} \sum_{i=t,\bar{t},b,\bar{b},j} m_{T,i} \tag{9.1}$$

$$\mu_F = \tfrac{1}{2}\sqrt[4]{m_T(t)\, m_T(\bar{t})\, m_T(b)\, m_T(\bar{b})}, \tag{9.2}$$

while the damping parameter is set to:

$$h_{\text{damp}} = \tfrac{1}{2} \sum_{i=t,\bar{t},b,\bar{b}} m_{T,i}. \tag{9.3}$$

In addition to the nominal sample, three dedicated variations are produced to assess systematic effects. These probe the choice of parton-shower and hadronisation model (using Herwig 7.04 instead of Pythia), the matrix-element–shower matching scheme (varying the `pthard` setting in Pythia 8.244 to 1), and the recoil model (switching from global to dipole recoil in Pythia 8.244).

A b-quark mass of $m_b = 4.95$ GeV is used, consistent with the b quarks produced in top decays. In addition to the nominal sample, three variations are generated to probe systematic uncertainties. The first tests the parton-shower and hadronisation model by replacing Pythia 8.244 with Herwig 7.04. The second explores the matrix-element–parton-shower matching by setting `pthard`=1 in Pythia 8.244. The final variation alters the recoil scheme, switching from the global recoil configuration used in the nominal sample to the dipole recoil option in Pythia 8.244.

$t\bar{t}$ Inclusive modelling

In addition to the dedicated $t\bar{t} + b\bar{b}$ samples, the remaining $t\bar{t}$+jets background is taken from inclusive five-flavour scheme $t\bar{t}$ samples, excluding events classified as $t\bar{t} \geq 1b$. The inclusive $t\bar{t}$ process is generated with Powheg-Box V2 at NLO QCD accuracy and interfaced with Pythia 8.230 for parton showering and hadronisation, using settings similar to those of the $t\bar{t}H$ signal samples. The `hdamp` parameter is set to $1.5\, m_t$, and the renormalisation and factorisation scales are defined as:

$$\mu_R = \mu_F = \sqrt{m_t^2 + p_T^2}.$$

Three additional variations are produced to assess modelling uncertainties. The first two test the parton-shower and hadronisation model (by replacing Pythia with Herwig) and the matching scheme between matrix element and shower (via the `pthard` setting), following the same strategy as for the signal process. The third variation changes the damping parameter to $h_{\text{damp}} = 3\, m_t$ to probe its impact on the event kinematics.

Process	ME Generator	ME PDF	PS	Normalisation
Signal and other Higgs samples				
$t\bar{t}H$	Powheg Box V2	NNPDF3.0nlo	Pythia 8.230	NLO+NLO (EW)
	Powheg Box V2	NNPDF3.0nlo	Herwig 7.04	NLO+NLO (EW)
$tHjb$	MadGraph5_aMC@NLO v2.6.2	NNPDF3.0nlo nf4	Pythia 8.230	–
tWH	MadGraph5_aMC@NLO v2.8.2	NNPDF3.0nlo	Pythia 8.245	–
	MadGraph5_aMC@NLO v2.8.2 [DR]	NNPDF3.0nlo	Pythia 8.245	–
$t\bar{t}$+ jets background				
$t\bar{t}$	Powheg Box V2	NNPDF3.0nlo	Pythia 8.230	NNLO+NNLL
	Powheg Box V2	NNPDF3.0nlo	Herwig 7.04	NNLO+NNLL
	Sherpa 2.2.10	NNPDF3.0nlo	Sherpa	NNLO+NNLL
$t\bar{t}+b\bar{b}$	Powheg Box Res	NNPDF3.1nlo nf4	Pythia 8.230	–
	Powheg Box V2	NNPDF3.1nlo nf4	Herwig 7.1.6	NNLO+NNLL
	Sherpa 2.2.10	NNPDF3.0nlo nf4	Sherpa	–
Other backgrounds				
tW	Powheg Box V2 [DR]	NNPDF3.0nlo	Pythia 8.309	NNLO+NNLL
	Powheg Box V2 [DS]	NNPDF3.0nlo	Pythia 8.309	NNLO+NNLL
	Powheg Box V2 [DS]	NNPDF3.0nlo	Herwig 7.2.3	NNLO+NNLL
t-channel	Powheg Box V2	NNPDF3.0nlo nf4	Pythia 8.230	NLO
s-channel	Powheg Box V2	NNPDF3.0nlo	Pythia 8.230	NLO
$t\bar{t}W$	Sherpa 2.2.10 (NLO[1j], LO[2j])	NNPDF3.0nlo	Sherpa	NLO+NLO (EW)
	MadGraph5_aMC@NLO v2.3.3	NNPDF3.0nlo	Pythia 8.210	NLO+NLO (EW)
$t\bar{t}Z(q\bar{q},\nu\nu)$	MadGraph5_aMC@NLO v2.8.1	NNPDF3.0nlo	Pythia 8.244p	NLO+NLO (EW)
	MadGraph5_aMC@NLO v2.8.1	NNPDF3.0nlo	Herwig 7.2.1	NLO+NLO (EW)
W+jets	Sherpa 2.2.11 (NLO [2j], LO[5j])	NNPDF3.0nlo	Sherpa	NNLO
Z+jets	Sherpa 2.2.11 (NLO [2j], LO[5j])	NNPDF3.0nlo	Sherpa	NNLO
$Z(\tau\tau)$+jets	Sherpa 2.2.14 (NLO [2j], LO[5j])	NNPDF3.0nlo	Sherpa	NNLO
VV (had.)	Sherpa 2.2	NNPDF3.0nlo	Sherpa	–
VV (lep.)	Sherpa 2.2	NNPDF3.0nlo	Sherpa	–
$VV+jj$ (EW)	Sherpa 2.2 (LO[EWI])	NNPDF3.0nlo	Sherpa	–
$t\bar{t}t\bar{t}$	MadGraph5_aMC@NLO v2.3.3	NNPDF3.1nlo	Pythia 8.230	NLO+NLO (EW)
tWZ	MadGraph5_aMC@NLO v2.6.2 [DR]	NNPDF3.0nlo	Pythia 8.235	–
tZq	MadGraph5_aMC@NLO v2.6.0	NNPDF3.0nlo	Pythia 8.230	–

Table 9.5 Summary of the MC simulated samples used in the analysis, including the ME generator, PDF set, parton shower, and normalisation scheme. The nominal samples are listed first, followed by alternative samples used for systematic variations in signal and background modelling. The label "nf4" denotes the $N_f = 4$ scheme of the NNPDF3.0 set [129]. Higher-order cross-sections applied for normalisation are shown in the final column. Empty entries indicate that no additional K-factor was applied.

9.4 Multivariate Analysis Techniques

In order to classify events into signal or background, advanced Multivariate Analysis (MVA) techniques are used. Similar MVA techniques are used to reconstruct Higgs boson candidates from the b-jets in an event through an explicit matching procedure. This allows us to then categories the Higgs events into STXS bins based on the reconstructed Higgs p_T. For the classification and reconstruction tasks required of this analysis, the transformer architecture outlined in Chapter 8.

9.4.1 Event Classification and Higgs Reconstruction

Network Overview Inputs

The multi-class classification transformer is trained to assign probabilities p_i for an event to originate from either the $t\bar{t}H$ signal or one of five $t\bar{t}$ background categories, yielding six output classes: $t\bar{t}H$, $t\bar{t}+\geq 2b$, $t\bar{t}+1b$, $t\bar{t}+1B$, $t\bar{t}+\geq 1c$, and $t\bar{t}$+light. The flavour categorise are defined according to the number of additional b-, c-, or light jets in the event.

The network input consists of reconstructed physics objects such as: jets, leptons, and the missing transverse momentum E_T^{miss}. Each object is represented by a vector of low-level kinematic features, primarily its four-momentum, supplemented by additional information. The features include the Cartesian momentum components (p_x, p_y, p_z), the energy, transverse momentum p_T, mass, pseudorapidity η, and azimuthal angle ϕ. To improve convergence, redundant features such as $\sin\phi$ and $\cos\phi$ are provided explicitly, encoding the periodicity of the azimuthal angle. Jet-specific inputs include the DL1r pseudo-continuous b-tagging score, while leptons are augmented with type, charge, and object-type flags. Where applicable, feature distributions are standardised to zero mean and unit variance, as summarised in Table 9.6.

The classifier is trained on nominal $t\bar{t}H$ signal and $t\bar{t}$+jets background samples, supplemented with alternative simulated datasets to enhance training statistics and stabilise convergence. In parallel, a reconstruction network is trained exclusively on $t\bar{t}H$ events, requiring two jets to be truth-matched to the Higgs decay products.

Feature	Description	Standardisation
Kinematics		
p_x	Object momentum in x-direction.	$\mu = 0,\ \sigma = 1.$
p_y	Object momentum in y-direction.	$\mu = 0,\ \sigma = 1.$
p_z	Object momentum in z-direction.	$\mu = 0,\ \sigma = 1.$
energy	Object energy.	$\mu = 0,\ \sigma = 1.$
p_T	Object transverse momentum.	$\mu = 0,\ \sigma = 1.$
mass	Object mass.	$\mu = 0,\ \sigma = 1.$
η	Object pseudo-rapidity.	$\mu = 0,\ \sigma = 1.$
ϕ	Object azimuthal angle.	$\mu = 0,\ \sigma = 1.$
$\cos\phi$	Cosine of object azimuthal angle.	$\mu = 0,\ \sigma = 1.$
$\sin\phi$	Sine of object azimuthal angle.	$\mu = 0,\ \sigma = 1.$
Identification		
PCBT score	DL1r pseudo-continuous b-tagging score assigned to jets. $\text{Score} = \begin{cases} 1, & \text{if un-tagged,} \\ 2, & \text{if tagged at [85\%, 77\%],} \\ 3, & \text{if tagged at [77\%, 70\%],} \\ 4, & \text{if tagged at [70\%, 60\%],} \\ 5, & \text{if tagged at 60\%.} \end{cases}$	None.
lepton type	Lepton type of input objects. ($e = 1$, $\mu = 2$).	None.
lepton charge	Charge of lepton objects.	$\mu = 0,\ \sigma = 1.$
E_T^{miss} flag	Whether input object is E_T^{miss} (1) or not (0).	None.

Table 9.6 Kinematic and identification features used as inputs to the classification network, together with the corresponding preprocessing standardisation.

9.4.2 Network architecture

The architectures of the transformer networks is shown in Figure 9.4. Both models adopt the encoder structure of the transformer, differing only in their treatment of the final layers. Input features are first embedded into a latent space through n_{embed} fully connected layers, producing d-dimensional latent vectors. These are processed by n_{attblock} transformer blocks, each consisting of a multi-headed self-attention mechanism followed by a feed-forward neural network. The latent dimension d is held constant to allow the use of residual connections, whereby the block input is added to the transformed output. This technique improves gradient flow and stabilises training.

In the classification network, the output of the final attention block is aggregated by a cross-attention pooling layer, which outperforms other pooling strategies such as max- and mean-pooling. The pooled features are mapped through a softmax layer to produce class probabilities, satisfying $\sum_i p_i = 1$. The reconstruction network instead

applies a jet-pairing layer to identify jets associated with the Higgs boson decay. A pairing score S_{ij} is computed for each jet pair (i, j) using the dot product of their latent feature vectors. The pair with the highest score, restricted to $j > i$, is taken as the Higgs candidate. The corresponding four-vector is then reconstructed, from which quantities such as p_T^H are extracted.

To improve class separation, the classification probabilities p_i are transformed into discriminant observables d_i defined as:

$$d_i = \frac{p_i}{\sum_{j \neq i} p_j \, \hat{N}_{ij}}, \tag{9.4}$$

where,

$$\hat{N}_{ij} = \frac{N_j}{\sum_{k \neq i} N_k}, \tag{9.5}$$

and N_j denotes the predicted yield of class j. This weighting prevents high-yield backgrounds from dominating and yields more balanced discriminants, enabling effective separation of processes with very different cross-sections. These observables are used in the definition of the signal and control regions.

An alternative set of observables, the so-called fractions, were also studied, defined as

$$f_i = \frac{p_i \, N_i}{\sum_j p_j \, N_j}. \tag{9.6}$$

Although this construction provides a natural yield-weighted probability, it was found to penalise classes with smaller yields. In contrast, the discriminants d_i not only offered slightly better classification performance, but also ensured more balanced control regions. For this reason, they were chosen as the final template variables in the profile likelihood fit used to extract the signal strength and cross-section.

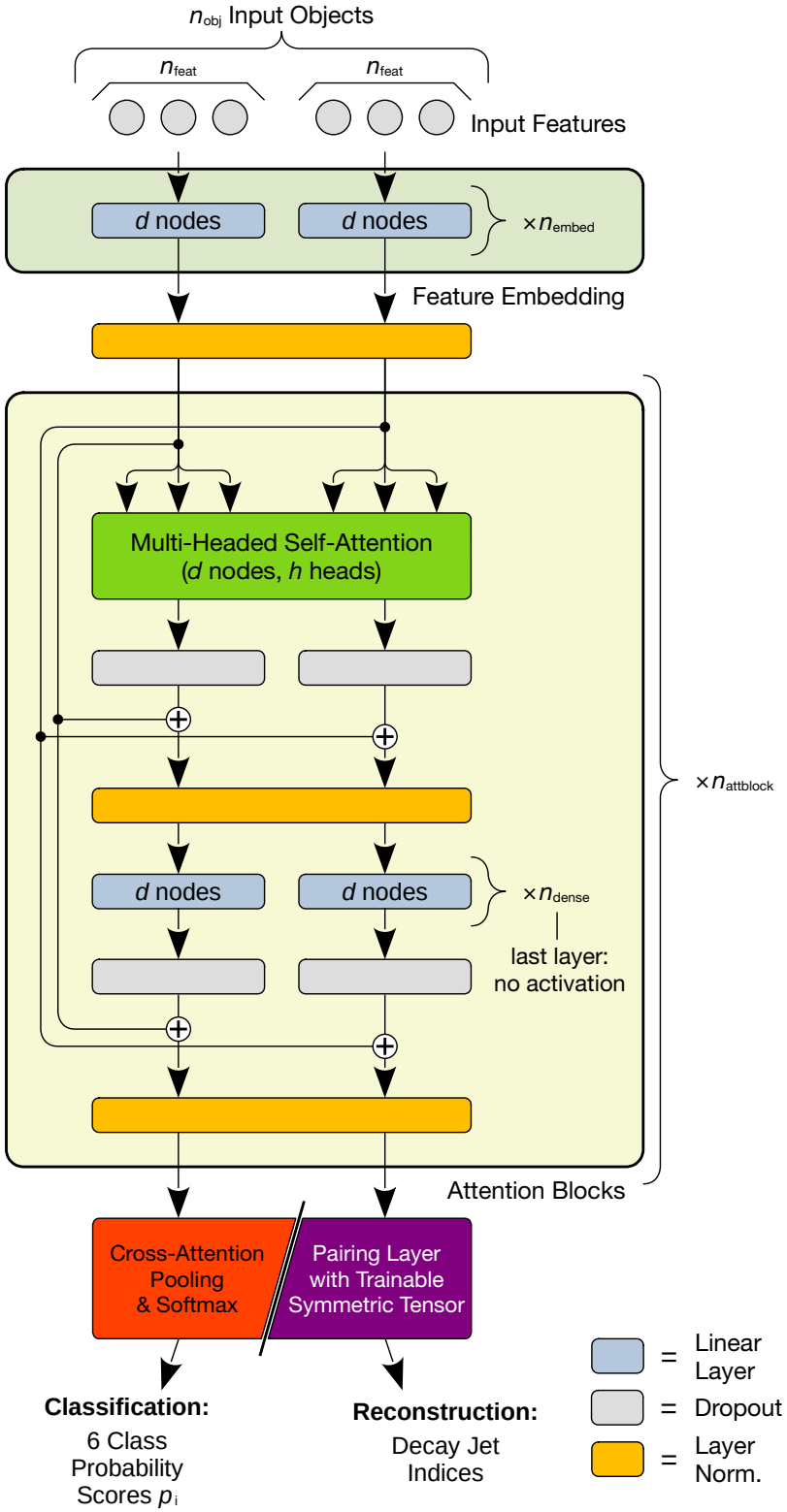


Figure 9.4 Schematic of the classification and reconstruction networks based on a permutation-invariant transformer. Input features are embedded into latent representations and processed through multi-headed attention blocks. For classification, outputs are aggregated via a cross-attention layer, while for p_T^H reconstruction, a pairing layer identifies the Higgs-decay jets [132].

9.4.3 Boosted Higgs Candidates Reconstruction

The boosted Higgs reconstruction uses a separate set of NNs from our transformer network. The boosted classification DNN is designed to distinguish RC-jets into three categories: Higgs, top, and QCD, and serves as the primary algorithm for Higgs candidate identification. Its architecture comprises three fully connected hidden layers, each with 100 neurons, employing the ReLU activation function. A final softmax layer produces class probabilities $P(H)$, $P(t)$, and $P(\text{QCD})$, constrained such that: $P(H)+P(t)+P(\text{QCD}) \approx 1$. The model is trained using the categorical cross-entropy loss function and optimised with stochastic gradient descent (SGD). Seventeen input features are provided, with jets ordered in p_T within each event. These features are constructed using physics-motivated RC-jet substructure observables chosen for their known discriminating power between Higgs, top, and QCD jets. Overtraining is controlled through odd–even event splitting, early stopping based on training–validation divergence. Validated by stable retraining performance and good Data/MC agreement of both inputs and network outputs [133]. The network is trained with 17 input features, constructed from p_T-ordered subjets within each RC-jet:

1. invariant mass of the RC-jet,
2. number of subjets contained in the RC-jet,
3. b-tagging working point (WP) of the leading subjet,
4. b-tagging WP of the sub-leading subjet,
5. maximum b-tagging WP among all subjets in the RC-jet,
6. minimum b-tagging WP among all subjets in the RC-jet,
7. sum of b-tagging WPs of all subjets, excluding the leading and sub-leading subjets,
8. transverse momentum p_T of the leading subjet,
9. transverse momentum p_T of the sub-leading subjet,
10. substructure variable d_{12},
11. substructure variable d_{23},
12. substructure variable Q_w,
13. invariant mass of all b-tagged subjets at the loosest 85% WP,

14. invariant mass of all non-b-tagged subjets,
15. angular separation ΔR between the leading and sub-leading subjets,
16. maximum ΔR separation among subjets in the RC-jet,
17. minimum ΔR separation among subjets in the RC-jet.

9.5 Analysis Regions and Pre-Fit modelling

Once the events pass the kinematic preselection, events are classified into their respective channels and regions. In the single-lepton (dilepton) channel, the selection is further divided into eight (five) control regions and six (six) resolved signal regions, defined according to the STXS binning scheme. The single-lepton channel also benefits from a dedicated boosted selection, whose events are subsequently included in the STXS 5 and 6 regions. In contrast, no boosted signal region is defined for the di-lepton channel, owing to the limited statistical power of this decay mode. The signal and background region septation for the single lepton channel is shown in Figure 9.5.

The assignment of events to analysis regions follows a hierarchical procedure. The first step is to check whether an event passes the $t\bar{t}H$ classifier requirement, defined by $d_{t\bar{t}H} > 4.072$ (9.031) for single-lepton (dilepton) events.

Events failing this requirement are classified into background control regions. Single-lepton events are further split into resolved or boosted categories. Resolved single-lepton and di-lepton events are assigned to background regions according to the maximum discriminant output of the flavour classifiers. This ensures that each region is enriched in a specific $t\bar{t}$+jets flavour component. For instance, events in the $t\bar{t}+1b$ regions must have the $t\bar{t}+1b$ discriminant value greater than all competing background discriminants. Boosted events, on the other hand, are collected into a dedicated control region.

For single-lepton events with a $t\bar{t}+\geq 2b$ maximum classifier, an additional split is applied in bins of the reconstructed Higgs transverse momentum, p_T^H. This improves the modelling of the $t\bar{t}+\geq 2b$ background and facilitates a more reliable extrapolation to the STXS signal regions. In the dilepton channel, the $t\bar{t}+\geq 2b$ region is kept inclusive in p_T^H due to limited statistics.

Events passing the $t\bar{t}H$ classifier threshold are directed to signal regions. Here, the reconstruction transformer network provides the estimate of p_T^H, and events are assigned to bins optimised to maximise purity with respect to the truth p_T^H. This binning strategy enhances the fraction of correctly assigned events, particularly in the highest p_T^H bins ($p_T^H > 300$ GeV), where this analysis provides unique sensitivity within the global Higgs combinations. The reconstructed p_T^H boundaries used are:

- Single-lepton channel: [0, 60, 114, 192, 282, 408, ∞] GeV,
- Dilepton channel: [0, 60, 114, 186, 270, 402, ∞] GeV.

For the single-lepton $t\bar{t}+ \geq 2b$ control regions, a coarser split is used: [0, 114, 282, ∞] GeV, matching the corresponding signal region binning.

With events split into their assumed classes by the MVA strategy, pre-fit checks can occur to inspect the in order to validate the MVA strategy and catch any major mis-modelling issues that occurred. This is accomplished with Data/MC comparisons to check region purity or simply checking the truth MC fractions would suffice, using these methods the Higgs reconstruction and event classification can be checked for all regions, signal and background, and for both channels single lepton and di-lepton.

In all distributions presented figures below, the $t\bar{t}$+jets normalisation scaling factors are applied together with the H_T re-weighting factors. The data/MC comparisons include the full uncertainty envelope from all systematic variations, with the exception of the free-floating $t\bar{t}$+jets normalisation factors, which carry no predefined pre-fit uncertainty. All discriminant outputs shown in the distributions are rescaled to the range $[0, 1]$ by applying a logistic transformation, defined as

$$f(x) = \frac{1}{1 + e^{-x}} \,. \tag{9.7}$$

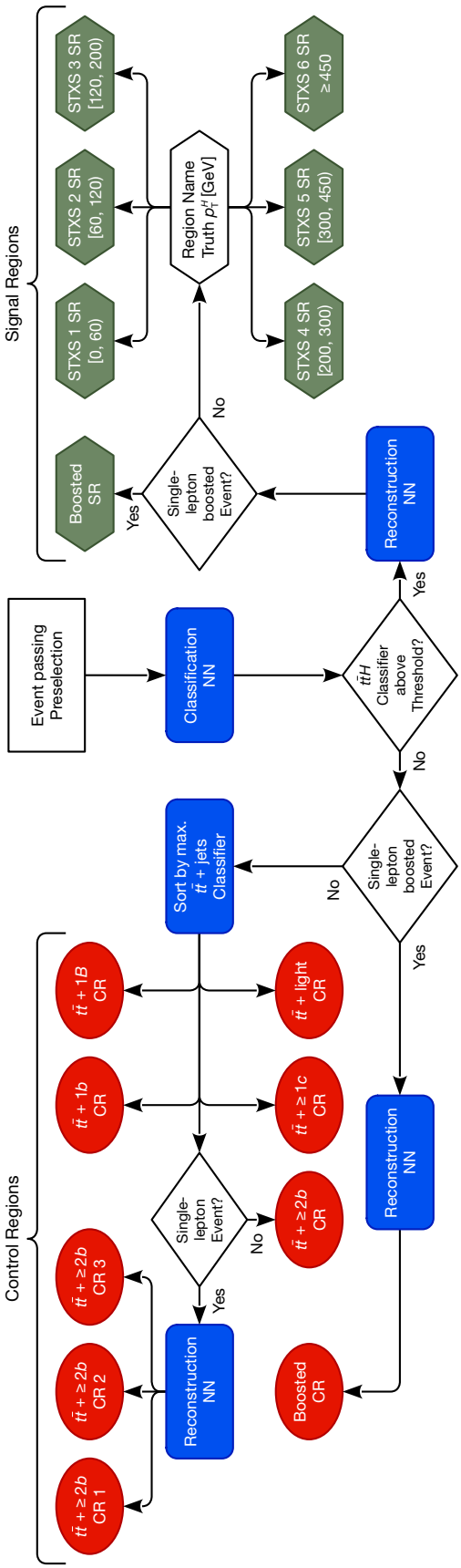
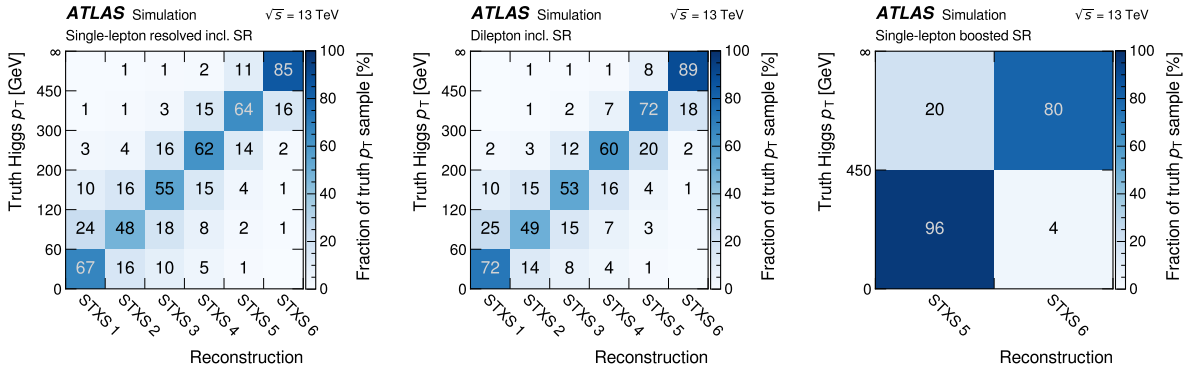


Figure 9.5 Flowchart illustrating the event classification and Higgs boson p_T reconstruction strategy employed in the analysis.

9.5.1 Higgs Reconstruction Performance

The performance of the reconstruction transformer and boosted DNN is shown in Figure 9.6. This is particularly important for differential measurements and demonstrates improved separation power as indicated by the high fraction of truth MC in their correct STXS bin.



(a) Single Lepton **(b)** Di-lepton **(c)** Single lepton boosted

Figure 9.6 Migration matrices illustrating the performance of the reconstruction transformer in assigning the Higgs boson candidate p_T^H, to the truth bins of the STXS framework. Each matrix element corresponds to the fraction of events from a given truth bin that are reconstructed in a particular reconstructed bin.

9.5.2 Pre-Fit Modelling

For this section only the single lepton pre-fit modelling distributions are shown, the di-lepton pre-fit modelling can be found in Appendix A.1. The single lepton pre-fit modelling validation can be broken up into two main sets of plots: the region composition plots, shown in Figure 9.26 and the data/MC distribution plots for both signal, shown in Figure A.5 and background regions, shown in Figure 9.28.

Each CR is enriched in the targeted $t\bar{t}$+jets flavour component. In particular, the $t\bar{t}$+light regions are very pure, with more than 50% of events originating from the $t\bar{t}$+light process. The $t\bar{t}+ \geq 2b$ and $t\bar{t}+ \geq 1c$ regions are also found to be relatively pure, with around 50% of the events coming from the respective processes. In the signal SR, the $t\bar{t}+ \geq 2b$ background dominates, while the $t\bar{t}+1b$ process provides the second largest contribution in the lower p_T^H bins. Other background categories only become significant at higher p_T^H.

In the single-lepton channel, the signal regions are presented in Figure A.5, consisting of six resolved regions and one boosted region. The pre-fit agreement with data is excellent across all regions, with the lower ratio panels showing the data-to-MC ratio within the defined uncertainties. The boosted signal region is binned in p_T^H, with bin boundaries aligned to the two highest STXS bins.

A total of eight control regions are also defined, as shown in Figure 9.28. These are constructed to target the main $t\bar{t}$+jets flavour components, and the pre-fit agreement with data is again good.

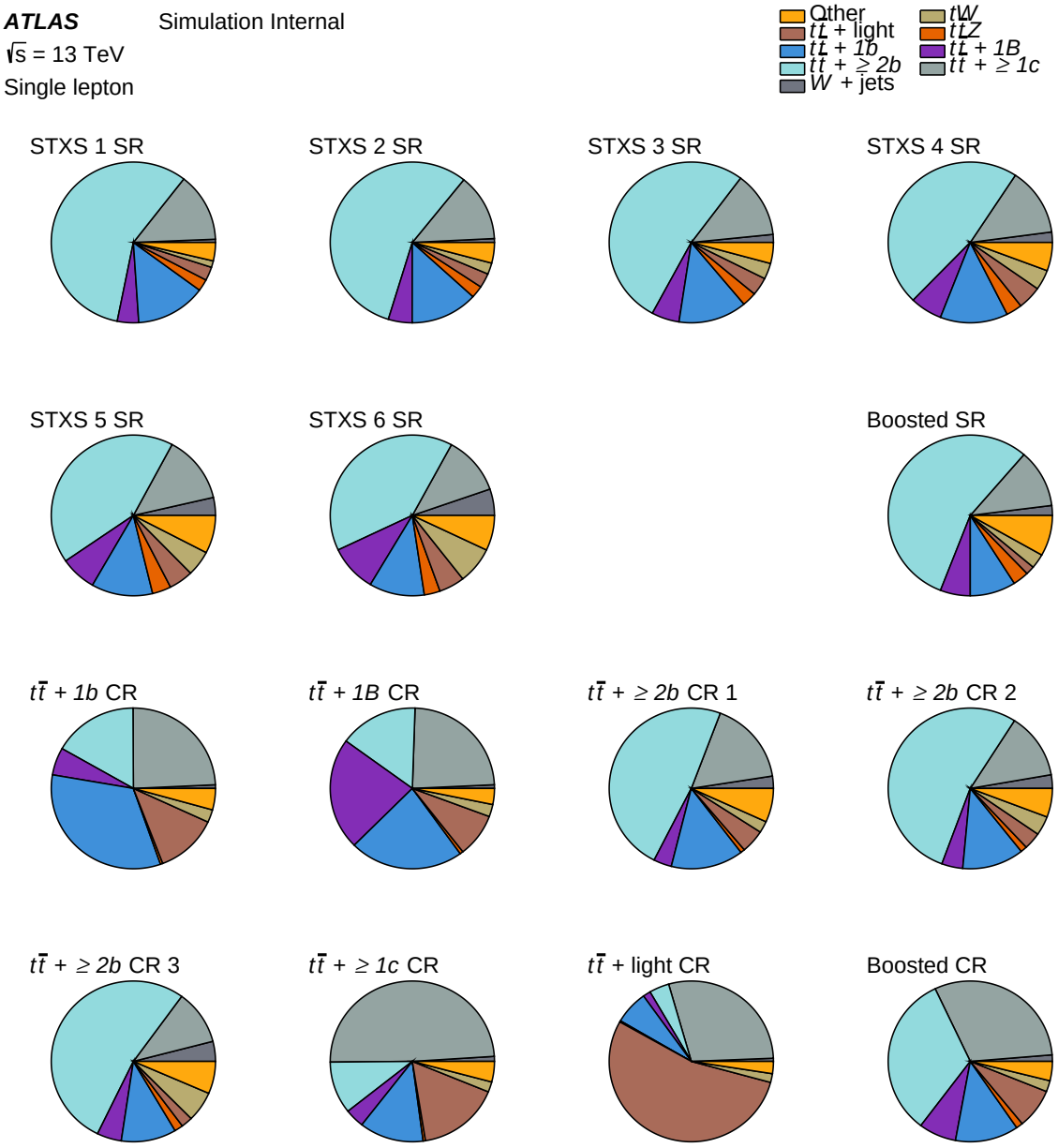
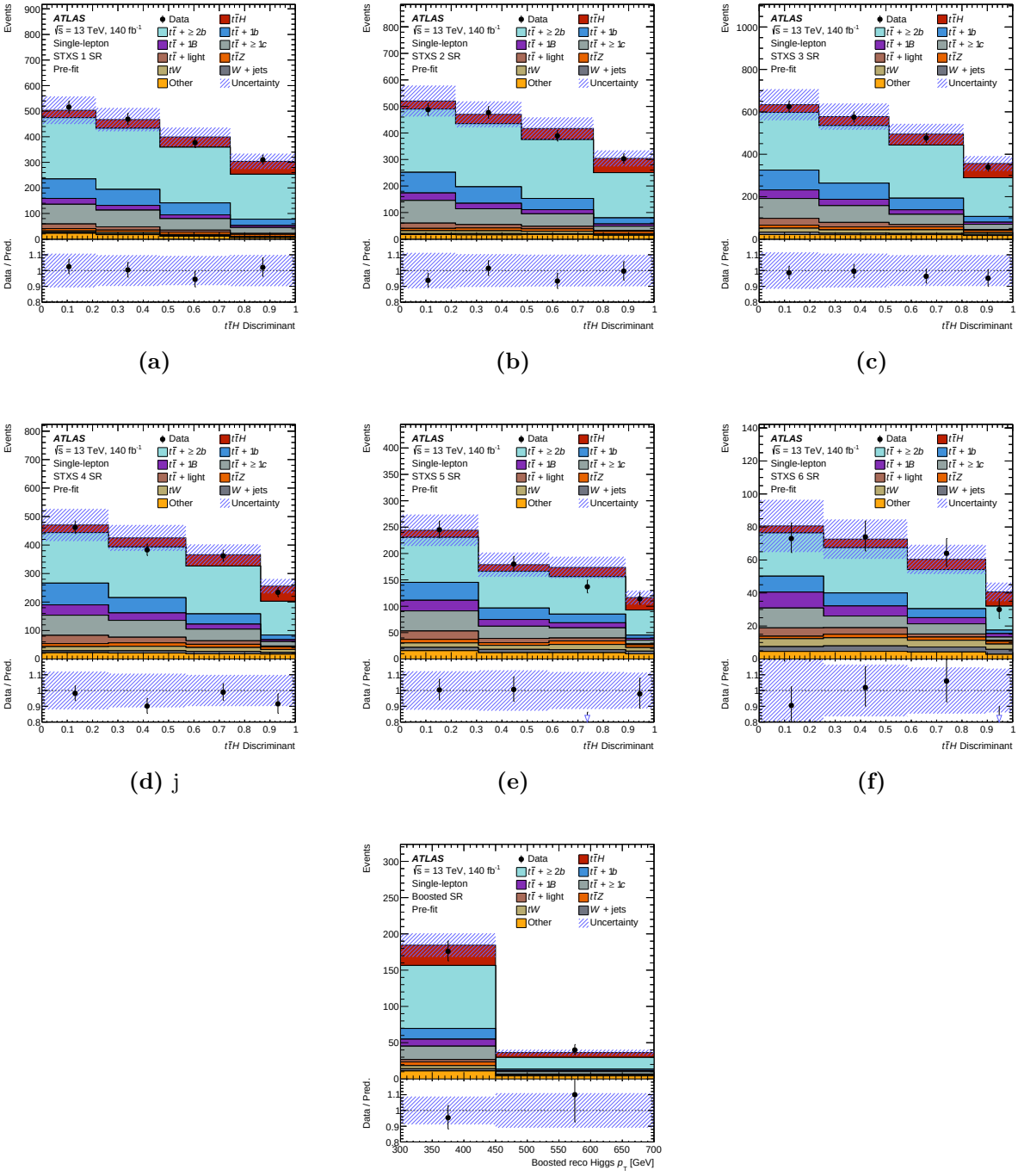


Figure 9.7 Composition of the signal and control regions of the single-lepton channel. The pie charts indicate the relative contributions of all processes, with one chart per analysis region. Each control region is enriched in the corresponding output class of the transformer network, while the signal region is subdivided into six bins of p_T^H.

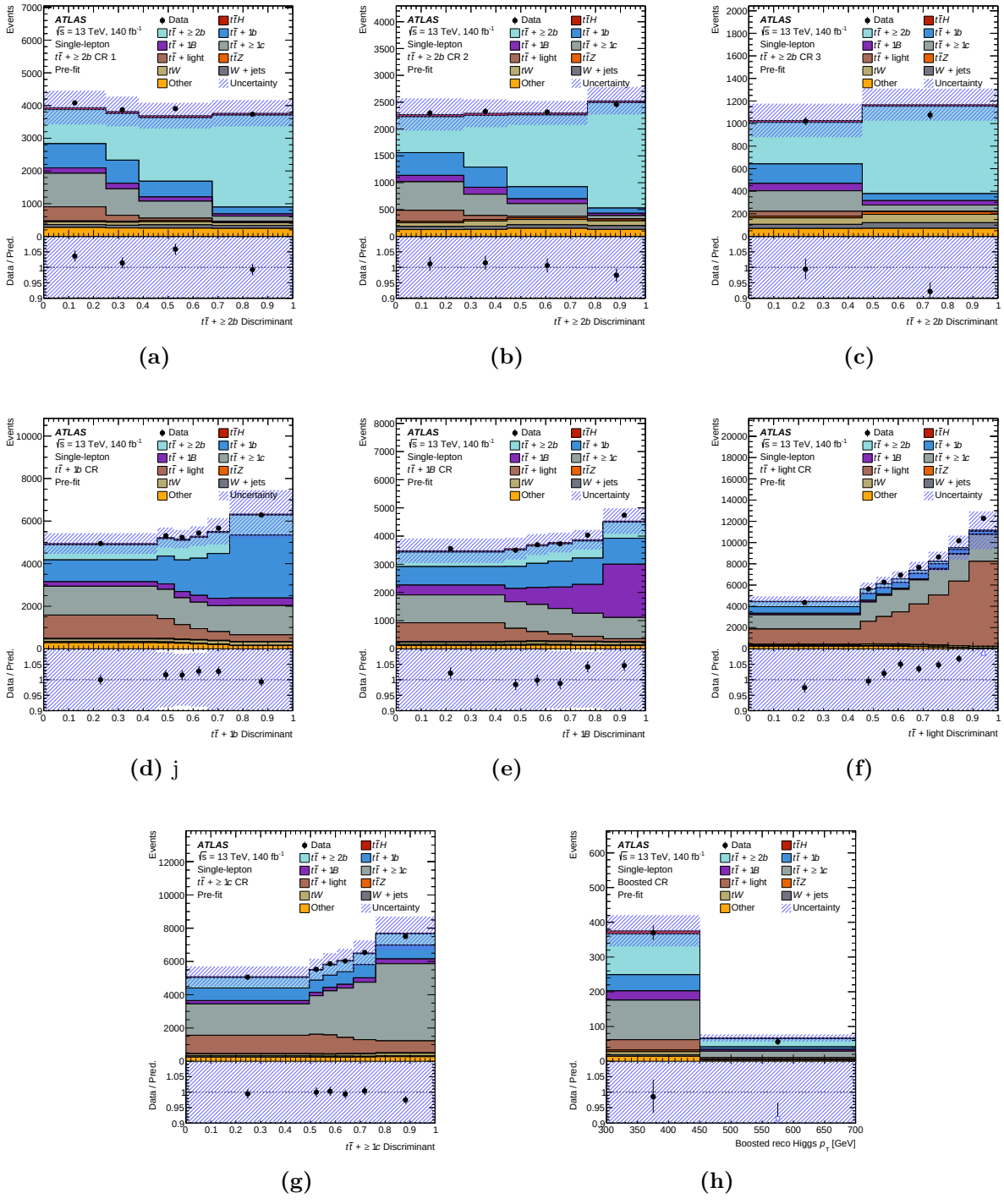


(a) **(b)** **(c)**

(d) j **(e)** **(f)**

(g)

Figure 9.8 Pre-fit distributions of the seven signal regions in the single-lepton channel: (a) STXS 1, (b) STXS 2, (c) STXS 3, (d) STXS 4, (e) STXS 5, (f) STXS 6, (g) boosted signal region. The uncertainty band consists of all statistical and systematic uncertainties, excluding the free-floating normalisations of the individual $t\bar{t}$+jets flavour components. All discriminant scores are rescaled to the interval $[0, 1]$ using a logistic function. The boosted region is binned in p_T^H, with boundaries aligned to the two highest STXS bins [127].



(a) **(b)** **(c)**

(d) **(e)** **(f)**

(g) **(h)**

Figure 9.9 Pre-fit distributions of the eight control regions in the single-lepton channel: (a) $t\bar{t}+\geq 2b$ STXS 12, (b) $t\bar{t}+\geq 2b$ STXS 34, (c) $t\bar{t}+\geq 2b$ STXS 56, (d) $t\bar{t}+1b$, (e) $t\bar{t}+1B$, (f) $t\bar{t}$+light, (g) $t\bar{t}+\geq 1c$, (h) boosted $t\bar{t}H$ control region. Each region is constructed to be enriched in the corresponding flavour component of the additional jets in the dominant $t\bar{t}$+jets background. The uncertainty band includes both statistical and systematic contributions, excluding the free-floating normalisations of the individual $t\bar{t}$+jets flavour components. All discriminant scores are rescaled to the interval $[0, 1]$ using a logistic function. The boosted signal region is binned in p_T^H.

9.6 Systematic Uncertainties

Various sources of systematic uncertainties are taken into account in the $t\bar{t}H(H \to b\bar{b})$ legacy analysis. These include experimental uncertainties associated with the reconstruction and identification of physics objects, as well as modelling uncertainties, which are estimated using alternative samples produced with different generators or generator settings. In this section, the systematic uncertainties considered in the analysis are briefly discussed, with emphasis on modelling uncertainties due to their dominant impact on the measurement.

9.6.1 Experimental Uncertainties

Luminosity and Pile-Up Modelling

The luminosity of the full Run 2 dataset, collected between 2015 and 2018, is measured to be 140.1 ± 1.2 fb^{-1}, corresponding to a relative uncertainty of 0.83% [134]. The uncertainty in pile-up modelling is estimated using a reweighting-based procedure, achieved by varying the nominal scale factors (SFs) applied to the pile-up distribution in data.

Leptons

For leptons, systematic uncertainties arise from the trigger, reconstruction, identification, and isolation procedures. Both electrons and muons exhibit efficiency differences between simulation and data, which are corrected using scale factors derived with the tag-and-probe method on $Z^0 \to \ell^+\ell^-$ events [83] The SFs are accounted for using four (ten) independent nuisance parameters (NPs) for electrons (muons).

Additional uncertainties stem from corrections to the energy and momentum scale and resolution of leptons. These corrections are obtained using $Z^0 \to \ell^+\ell^-$ and $J/\psi \to \ell^+\ell^-$ events, with further corrections for electrons derived from the E/p ratio measured in $W^\pm \to e\nu$ events [135]. In total, three (four) NPs are considered for electrons (muons).

Jet Energy Scale (JES)

JES uncertainties are derived by combining information from test-beam data, collision data, and simulation. Among the NPs there are, three flavour-specific: the b-jet energy scale, the flavour response, and the flavour composition. A more detailed description of all sources and their associated NPs is provided in [89].

Jet Energy Resolution (JER)

JER uncertainties are determined using dijet events in both data and simulation, where the balance in jet p_T and η is studied [89]. Thirteen independent uncertainty sources are considered, represented by 12 effective terms plus one term for the comparison between data and Monte Carlo.

b-Tagging

Systematic uncertainties in b-tagging arise from the efficiency calibration for b-jets and from the mis-tag rates for c- and light-flavour jets in the DL1r algorithm. Uncertainties are evaluated through a principal component analysis, yielding a total of 45 eigenvalues for b-jets, 20 for c-jets, and 20 for light-flavour jets. Additional uncertainties are introduced at high p_T due to the limited kinematic coverage of the calibrations. These extrapolation uncertainties are the same derived in Chapter 7.

Missing Transverse Momentum

All uncertainties related to the energy scale and resolution of the reconstructed physics objects described above are propagated to the E_T^{miss} algorithms. In addition to these hard contributions, extra uncertainties are included for the soft terms to account for discrepancies between data and simulation [100]. In total, three independent nuisance parameters are considered.

9.6.2 Signal Modelling Uncertainties

The impact of parton distribution function and α_s variations on the total $t\bar{t}H$ cross-section is estimated to be about $\pm 3.6\%$ [106], while the effect on the p_T^H distribution is

found to be negligible. Additional theoretical uncertainties on the $t\bar{t}H$ cross-section are considered from possible migrations between STXS bins.

To probe the uncertainty from initial-state radiation (ISR), dedicated weights are used in the matrix element and parton shower simulations. The renormalisation and factorisation scales, μ_R and μ_F, are varied by a factor of 0.5 in combination with the Var3c up tune [110], corresponding to enhanced parton radiation, and by a factor of 2 with the Var3c down tune to simulate reduced radiation. A similar approach is applied for final-state radiation (FSR), where the renormalisation scale of the QCD emission in the parton shower is varied by factors of 0.5 and 2.

The overall theoretical uncertainty on the $t\bar{t}H$ cross-section due to higher-order QCD corrections is evaluated by the LHC Higgs Cross-Section Working Group and amounts to 9.2%. For the Higgs boson branching fractions, uncertainties of $(+1.2/-1.3)\%$ are applied for $H \to b\bar{b}$, $(+1.6/-1.5)\%$ for $H \to W^{\pm}W^{\mp}$, and $\pm 5\%$ for other decay modes.

To account for uncertainties from NLO matching, the nominal sample is compared to a POWHEG-BOX V2 sample interfaced with PYTHIA 8 with the `pthard` parameter set to 1, shown in Table 9.5. A similar comparison is performed for parton shower and hadronisation modelling, where the nominal setup is tested against a POWHEG-BOX V2 sample interfaced with HERWIG 7.04.

9.6.3 Background Modelling Uncertainties

$t\bar{t}$+jets

The treatment of $t\bar{t}$+jets uncertainties is tailored to the different subprocesses. The $t\bar{t}$+light contribution is relatively well constrained in data, owing to its distinct diagrammatic structure compared to $t\bar{t} + 1b$ and $t\bar{t} + 1c$. In contrast, the $t\bar{t} + 1b$ and $t\bar{t} + 1c$ processes can receive additional contributions depending on the chosen flavour scheme of the parton distribution functions. Moreover, the mass difference between the c- and b-quarks introduces further distinctions for the heavy-flavour ($t\bar{t}$+HF) components.

Acceptance And Shape Uncertainties

Uncertainties in acceptance and shapes are evaluated by comparing the nominal predictions with alternative setups. For the $t\bar{t}+1b$ subprocess, the use of a profile likelihood fit (PLF) to determine normalisation factors in data complicates this comparison, as unmodified alternatives would strongly distort the fraction of $t\bar{t}+1b$ events within the analysis phase space. To mitigate this, the alternative predictions are reweighted to preserve the same $t\bar{t}+1b$ fraction as in the nominal sample.

Parton Shower And Radiation Modelling

ISR and FSR uncertainties are assessed using a weight-based method, varying the renormalisation and factorisation scales by factors of 0.5 and 2. To probe the impact of the parton shower and hadronisation model, the nominal sample is compared against a Powheg-Box V2 sample with identical settings interfaced to Herwig 7.16. An additional variation is produced to test recoil modelling in the initial-state shower, replacing the global recoil scheme in Pythia 8 with dipole recoil.

NLO Matching And h_{damp}

NLO matching uncertainties are evaluated by comparing the nominal setup with a Powheg-Box V2 sample interfaced with Pythia 8 where the `pthard` parameter is set to 1. In addition, variations of the h_{damp} parameter are included, as described in Chapter 5.

H_T Reweighting

Uncertainties associated with the H_T reweighting are evaluated in both channels. Two variations are considered: (i) alternative reweightings obtained using adapted post-fit normalisation factors for $t\bar{t}+1c$ and $t\bar{t}$+light, with central values shifted by their uncertainties, and (ii) switching the reweighting on and off entirely. The on/off variation is applied independently to each N_{jets} reweighting, while the adapted post-fit variation affects the overall H_T reweighting.

Fake And Non-Prompt Leptons

For the single-lepton channel, fake and non-prompt leptons are assigned an overall normalisation uncertainty of 50%. In addition, variations are derived using an alternative parametrisation of the two-dimensional fake efficiencies, binned in lepton p_T and $|\eta|$. For the di-lepton channel, a single overall normalisation uncertainty of 50% is applied.

Other Backgrounds

In addition to the $t\bar{t}$+jets processes, further background contributions are taken into account. For these processes, the associated systematic uncertainties and their sources are summarised in Table 9.7.

Process	Source of Systematic Uncertainty	Description
$t\bar{t}V$ [136]	Cross-section	±15%
	NLO matching	`MadGraph5_aMC@NLO` + `Pythia 8` vs `Sherpa`
	PS and hadronisation	`MadGraph5_aMC@NLO` + `Pythia 8` vs `Sherpa`
Single-t [137, 138, 139]	NLO cross-section	±5%
	NLO matching	`Powheg-Box V2` + `Pythia 8` vs `Powheg-Box V2` + `Pythia 8` (pthard)
	PS and hadronisation	`Powheg-Box V2` + `Pythia 8` vs `Powheg-Box V2` + `Herwig 7`
	tW and $t\bar{t}$ interference [140]	Diagram Removal (DR) vs Diagram Subtraction (DS) scheme
W+jets	Cross-section	±40%
	Cross-section	±30%
Z+jets	Normalisation	±35%
Diboson [141]	Cross-section and additional jet production	±50%
tZq [142]	Cross-section (PDFs)	±0.9%
	Factorisation and normalisation	±7.9%
tWZ [142]	Cross-section	±50%
$tHjb$	Cross-section (PDFs)	±3.7%
	Factorisation and normalisation	$+6.5\%/-14.9\%$
tWH	Cross-section (PDFs)	±3.7%
	Factorisation and normalisation	$+6.5\%/-6.7\%$
$ttt\bar{t}$	Normalisation	±50%

Table 9.7 List of all systematic uncertainties considered for the other background processes, together with their source and description.

9.7 Statistical Model and Fit Studies

The analysis is performed under a blinded strategy to avoid bias in the interpretation of the results. This means that the data in the most sensitive signal regions is not inspected until the statistical model has been fully validated. Before unblinding, a series

of dedicated fit studies is carried out using the blinded model to ensure that the any pulls in the model are understood and not a result of mis-modelling, and that the systematic uncertainties are properly implemented. After these studies we proceeded with the unblinding to reveal our signal contributions.

This section will discuss the construction of statistical model and the fit studies conducted. A large number of fit studies were carried out between myself and the analysis team to validate the statistical model. I will briefly summarise a selection of these studies to illustrate the methodology and complexity of these studies. Finally, I present the validated, unblinded model that forms the basis of the measurement.

9.7.1 Constructing the Model

As mentioned in Chapter 8 the likelihood function is created within the RooStat framework [120, 121], which is applied using TRExFitter, which is based on HistFactory [122]. The minimisation is performed using the Minuit library implemented in ROOT [123, 124]. The likelihood function used in this analysis is defined as the product of Poisson probabilities for each bin of the signal and background templates, given the observed data, together with Gaussian constraint terms that encode systematic uncertainties. Explicitly, the likelihood takes the form:

$$\mathcal{L}(\mu, \vec{k}, \vec{\theta}) = \prod_{r \in \{\text{SR, CR}\}} \left[\prod_{i \in \text{bins}} P\Big(n_{r,i} \,|\, \nu_{r,i}(\mu, \vec{k}, \vec{\theta})\Big) \right] \times \prod_{j \in \text{syst.}} G(\theta_j^0 \,|\, \theta_j), \tag{9.8}$$

where $n_{r,i}$ denotes the observed number of events in bin i of region r, and $\nu_{r,i}$ the corresponding expected yield, which depends on the signal strength μ, the background normalisation factors $\vec{k}$, and the nuisance parameters $\vec{\theta}$. All nuisance parameters θ_j are defined such that their best-fit values correspond to $\hat{\theta}_j \equiv \theta_j^0 = 0$, with associated uncertainties normalised to unity ($\sigma_{\hat{\theta}} \equiv \Delta\theta = 1$). This convention effectively calibrates the likelihood. Finally, statistical uncertainties from the limited size of Monte Carlo simulated samples are included as additional Poisson terms.

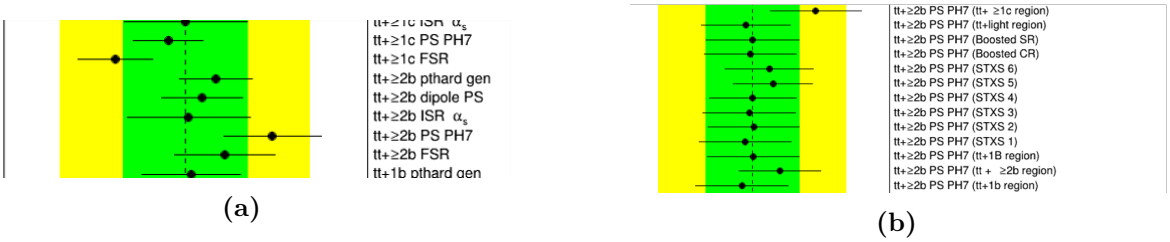
9.7.2 Fit Studies

Here we detail two example fit studies conducted during the validation of the statistical model of the analysis which required additional investigation. These studies were

performed on various versions of the fit. As such, steps were taken to relax tensions and pulls in the initial fits prior to the final fit. As, shown in Section 9.7.3, some pulls were relaxed. If pulls remain in the final fit these pull were well enough understood. Many other minor fit studies were conducted during this time most of which were simple decorrelations of NP to identify problem regions. The systematic variation plots were then inspected and either a change to the smoothing algorithm or implementing a newer version of CDI, which resulted in a relaxation of the pull. These minor studies are omitted for the sake of brevity.

$t\bar{t} \geq 2b$ PS PH7 Pull In The Single Lepton Channel

A significant pull was observed in the single-lepton channel for the $t\bar{t} \geq 2b$ parton-shower PH7 (Powheg + Herwig 7) systematic. The presence of this pull indicated that the fit was using the nuisance parameter aggressively, suggesting non-trivial mismodelling effects. To identify the source, a systematic de-correlation across analysis regions was performed. This showed that the $t\bar{t} + c$ control region was the principal driver of the tension. The plots showing the pulls are shown below in Figure 9.10.



(a) **(b)**

Figure 9.10 Pull plots showing the $t\bar{t} + 2b$ PS PH7 and $t\bar{t} + c$ FSR pull (a) the region decorrelation of the $t\bar{t} + 2b$ PS PH7 pull (b).

Inspection of the NP in this region revealed that the systematic had a pronounced shape effect rather than acting as a pure normalization shift. Comparison of pre- and post-fit distributions confirmed that data–MC discrepancies persisted even after the fit, reinforcing the hypothesis that the pull was rooted in shape mismodelling. To probe this further, the behaviour of $t\bar{t} + 2b$ events within the $t\bar{t} + c$ control region was examined shown in Figure 9.11.

The $t\bar{t} \geq 2b$ contribution exhibited a clear shape modification: the yield was enhanced at low NN($t\bar{t} + c$) scores in a way that closely followed the behaviour of the PH7 systematic.

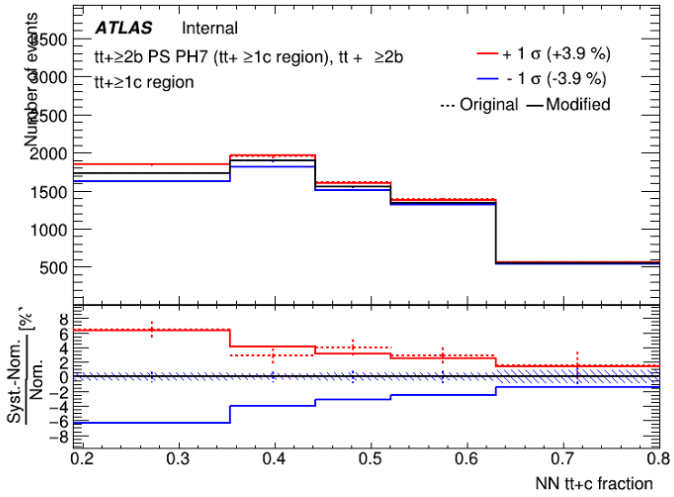


Figure 9.11 The systematic variation plot $t\bar{t} \geq 2b$ parton-shower PH7 systematic in the $t\bar{t}+2b$ sample in the $t\bar{t}+c$.

At higher NN scores, additional systematics appeared to be involved, indicating that the tension was not attributable to a single nuisance. Several candidate systematics, acting both on $t\bar{t}+c$ and $t\bar{t}+2b$ samples, were identified as capable of amplifying the apparent $t\bar{t}+2b$ contribution, particularly in the high-NN bins. Further evidence came from

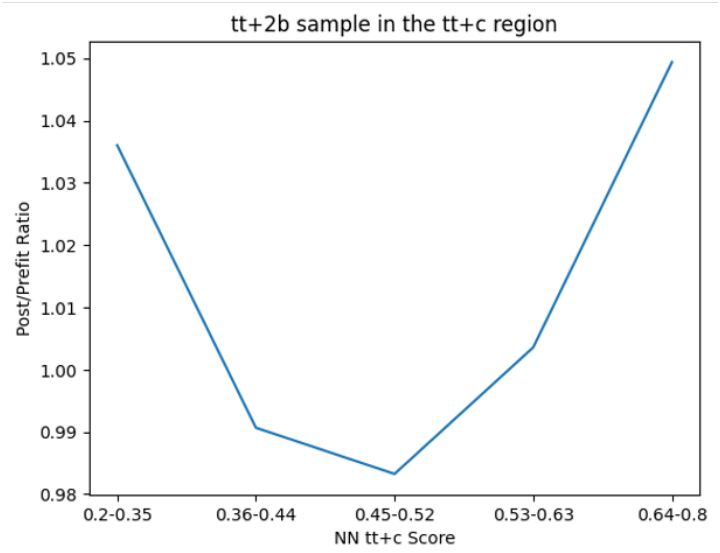
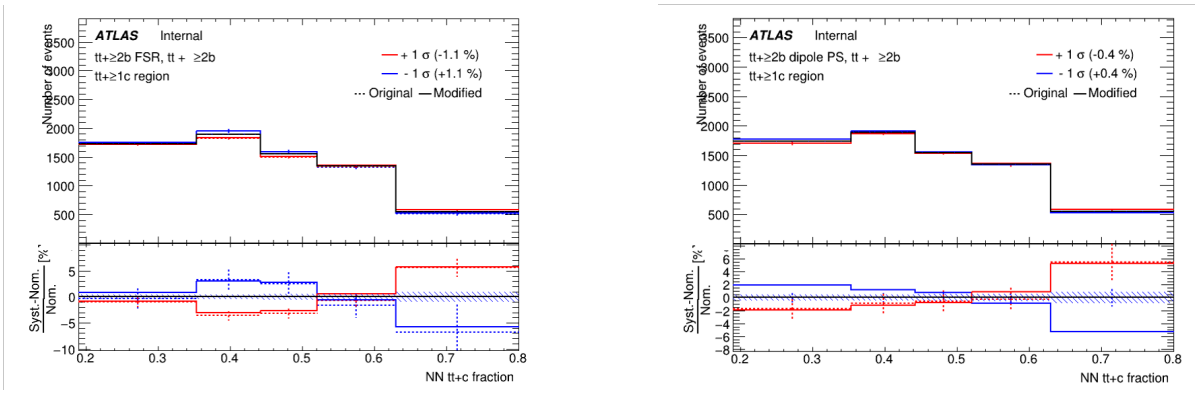


Figure 9.12 A distribution of the $t\bar{t}+2b$ sample in the $t\bar{t}+c$ region as a function of NN $t\bar{t}+c$ score.

the $t\bar{t}+c$ FSR systematic, which also showed a strong pull and introduced noticeable shape distortions shown in Figure 9.10. This suggests that multiple, entangled nuisance parameters propagate through the fit, combining to produce the observed PH7 pull. In

short, the $t\bar{t}+2b$ PH7 systematic does not reflect a single isolated modelling issue but rather the interplay of overlapping systematics in the $t\bar{t}+c$ control region. We also see that the raw systematic variation plots for $t\bar{t}+2b$ FSR and dipole systematics, shown in Figure 9.13, for the $t\bar{t}+2b$ sample in the $t\bar{t}+c$ region, show the possible increase in $t\bar{t}+2b$ events in the $t\bar{t}+c$ region at high $t\bar{t}+c$ NN score bins. This plays into the non trivial nature of the $t\bar{t}+2b$ PS PH7 systematic.



(a) **(b)**

Figure 9.13 The systematic variation plots for $t\bar{t}+2b$ FSR (a) and dipole (b) systematics for the $t\bar{t}+2b$ sample in the $t\bar{t}+c$ region.

Flavour Tagging Light EV0 Pull

In the combined fit, a moderate but non-negligible pull was observed in the light-jet EV0 systematic, shown in Figure 9.14. Importantly, this behaviour was absent in the individual single-lepton and di-lepton fits, but appeared only in the combination. The presence of such a pull indicates that the fit is using this nuisance parameter to absorb tension between channels, rather than correcting a feature of the data.

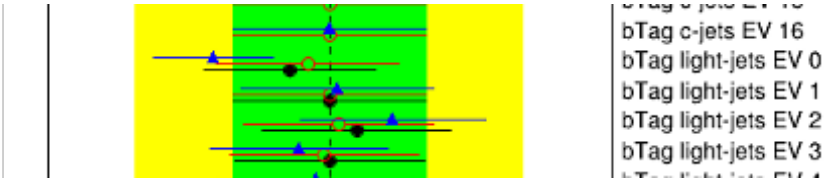
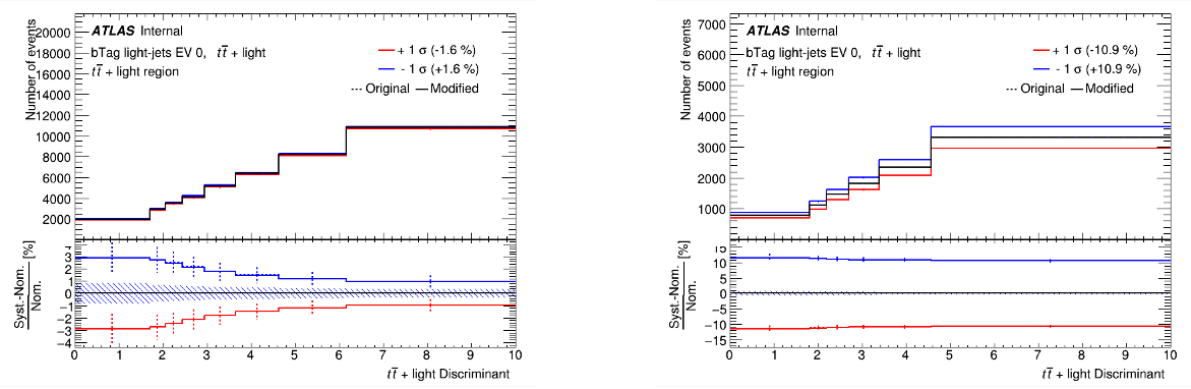


Figure 9.14 Pull plots showing the flavour tagging light jet EV0 pull in the combination fit.

The first step was to inspect the pre-fit and post-fit behaviour of the $t\bar{t}+$ light control region. The di-lepton channel showed higher yield and larger variation than the single

lepton channel, suggesting that the fit was exploiting the EV0 nuisance to balance the two. At first this was thought to be an implementation error. However, considering the event definition of each channel revealed an important difference: in the single lepton selection, the hadronic W decay can proceed via $W \to cs$, introducing a charm quark that is more likely to be mis-tagged as a b-jet. This mechanism effectively reduces the relative rate of light-jet mis-tags in the single lepton sample. By contrast, in the di-lepton channel both W bosons decay leptonically, so this source of charm contamination is absent, making the di-lepton channel more sensitive to genuine light-jet mis-tagging. The result is an intrinsic difference between the two channels in their response to the EV0 systematic. The systematic variation plots for the light EV0 systematic in the single and di-lepton channel for in the tt+light region is shown in Figure 9.15.

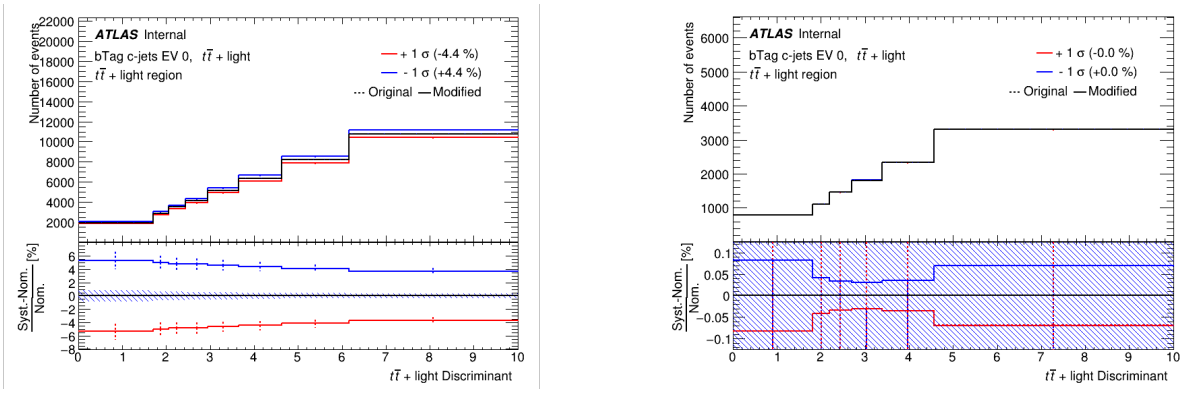


(a) **(b)**

Figure 9.15 The systematic variation plots for the light EV0 systematic in the single (a) and di-lepton (b) channel for the $t\bar{t}$+light region.

Comparing the behaviour of the c-jet EV0 nuisance across channels. Indeed, the single lepton channel showed stronger sensitivity to c-jet mis-tags, consistent with the $W \to cs$ contribution, whereas the di-lepton channel retained a stronger dependence on light-jet mistagging. This confirmed that the channel-specific physics was driving the tension.

Finally, we explored whether the pull could be relaxed by de-correlating the k-factors associated with the $t\bar{t}$ + light background across the two channels. Allowing $k^{1\ell}_{t\bar{t}+\text{light}}$ and $k^{2\ell}_{t\bar{t}+\text{light}}$ to float independently reduced the combined light-jet EV0 pull to below one standard deviation. This shows that the nuisance parameter was previously acting as a proxy to reconcile mismodelling differences between the channels, rather than indicating a fundamental problem with the flavour-tagging implementation. This study was shown to the flavour tagging group who agreed that this was a feature of the fit and not a bug



(a) **(b)**

Figure 9.16 The systematic variation plots for the c-jets EV0 systematic in the single (a) and di-lepton (b) channel in the $t\bar{t}$+light region.

in the derivation of the NPs.

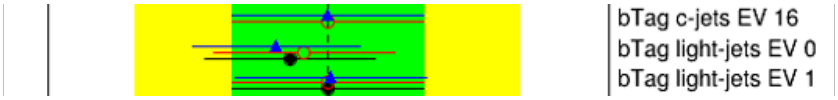


Figure 9.17 The relaxed pull of the light EV0 systematic.

This decorrelated k-factor approach, based on my recommendation, was adopted as in the final fit as shown in Figure 9.18.

9.7.3 Finalised Statistical Model

The finalised statistical model represented by the $S + B$ fit showing the background normalisation factors in Figure 9.18. The $t\bar{t}$ specific background modelling systematic pulls in Figure 9.19. The flavour tagging systematic pulls, Figure 9.20. The number of finalised extrapolation uncertainties presented in Chapter 7 and seen in the final fit are different due to some being pruned out. Finally, we have the other modelling systematic parameters shown in Figure 9.21.

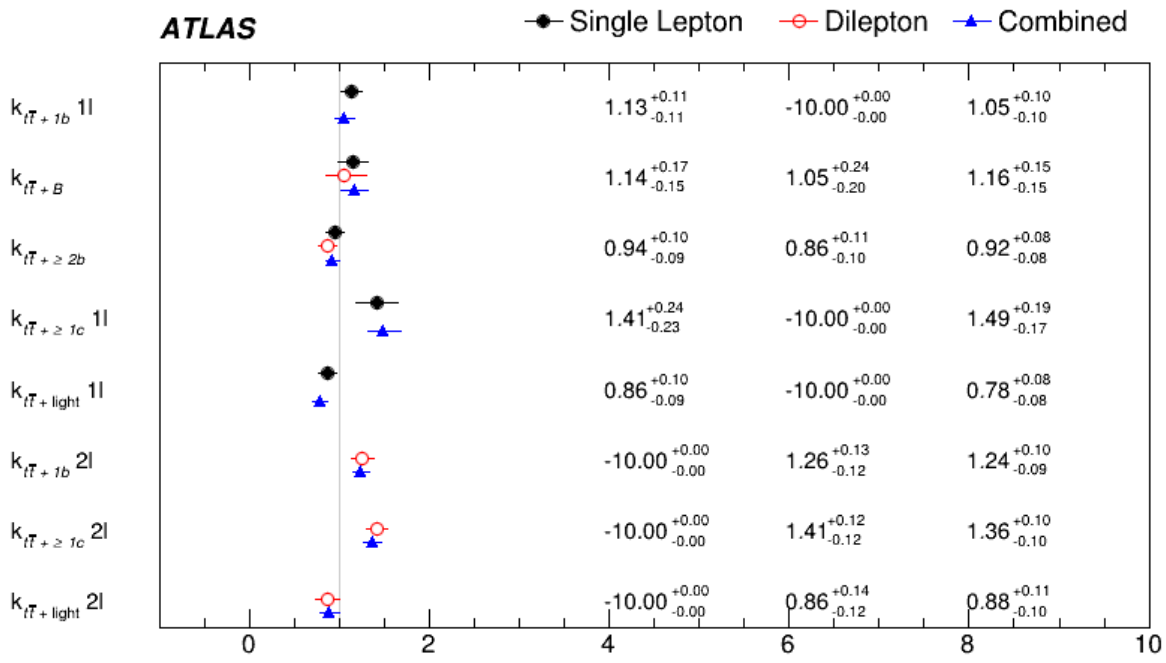


Figure 9.18 Best-fit results for all free-floating normalization factors in a fit to observed data. The background normalization factor comparison includes fits to only the single lepton and di-lepton channels, respectively, as well as the combined fit to all channels. Entries of -10 indicate that the normalization factor is not defined.

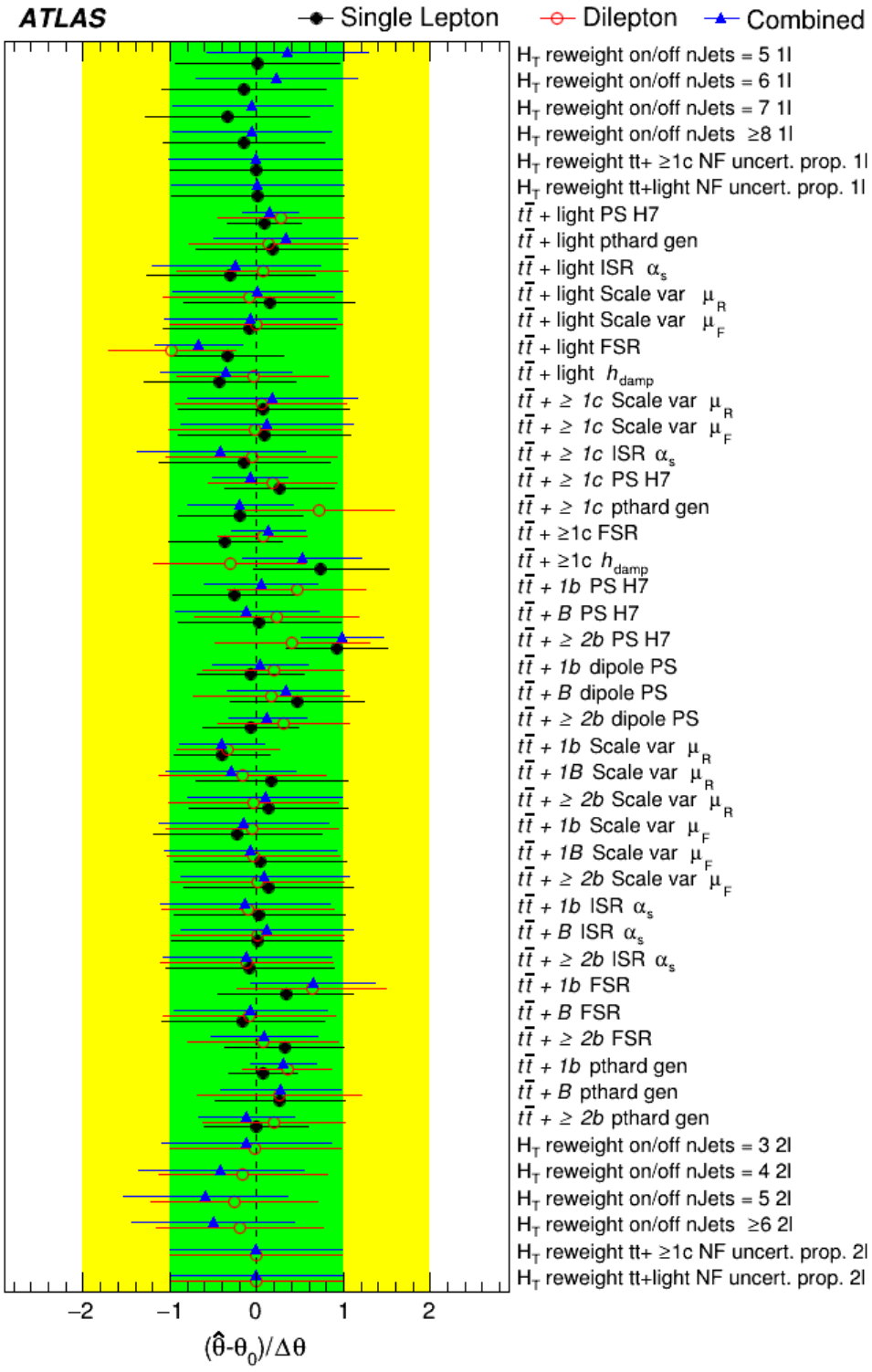


Figure 9.19 The $t\bar{t}$ specific background modelling parameters pull plots for the final fit.

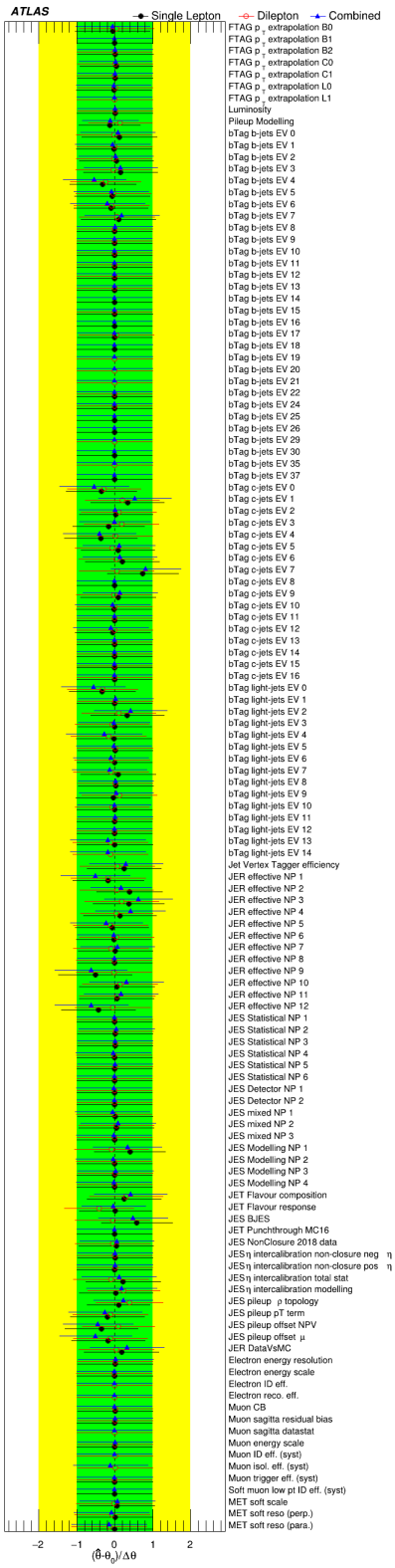
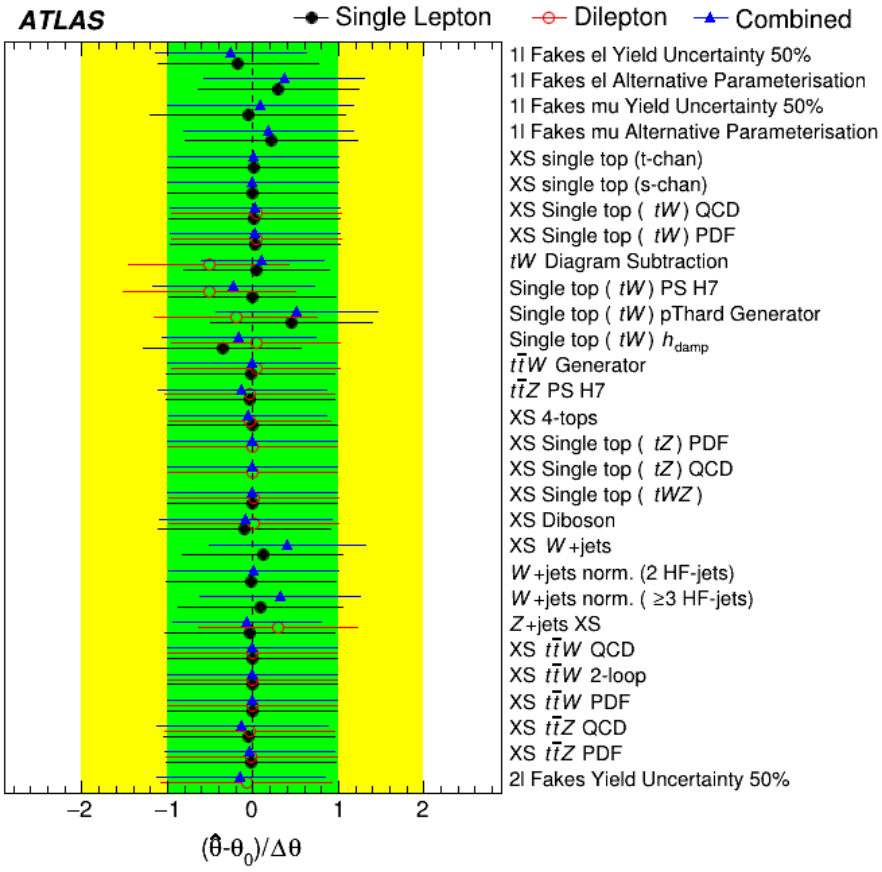
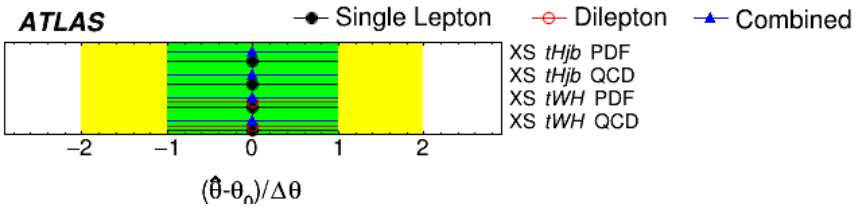


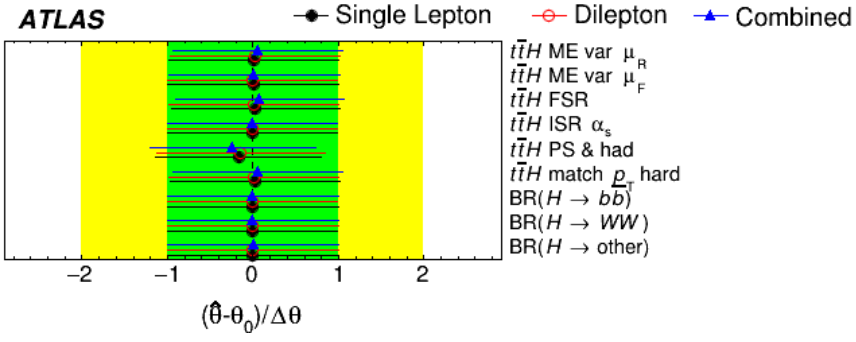
Figure 9.20 The flavour tagging systematic pulls.



(a)



(b)



(c)

Figure 9.21 Nuisance parameter pulls across modelling categories: (a) Other processes, (b) tH, (c) $t\bar{t}H$, and (d) $t\bar{t}$.

9.8 Results

The results are presented in two categories. The *inclusive measurement* refers to results where the STXS regions are combined, either across both signal channels or within each channel separately, by combining the STXS regions. This provides a single measurement of the overall $t\bar{t}H$ signal strength without resolving the individual kinematic regions or an overall per channel result. The *differential measurement* instead keeps the STXS regions separate, while combining the single-lepton and di-lepton channels. This allows the signal strength to be determined as a function of the Higgs boson transverse momentum, thereby probing the kinematic dependence of the $t\bar{t}H$ production process. The global goodness of fit is 87% for the $\sigma_{t\bar{t}H}/\sigma_{\mathrm{SM}}$ measurement in p_T bins, highlighting that good post-fit modelling is achieved.

9.8.1 Inclusive

An excess of events over the background is observed (expected) with a significance of 4.6 (5.4) standard deviations, from a combined profile-likelihood fit to data across all signal and control regions. The measured $t\bar{t}H$ signal strength for $m_H = 125.09$ is:

$$\mu_{t\bar{t}H} = 0.81^{+0.22}_{-0.19} = 0.81 \pm 0.11(\mathrm{stat.})\ ^{+0.20}_{-0.16}(\mathrm{syst.}). \tag{9.9}$$

The statistical component of the uncertainty is defined as the variation in $\mu_{t\bar{t}H}$ when all nuisance parameters associated with systematics are fixed to their best-fit values. The systematic component is then obtained as the difference in quadrature between the total and statistical uncertainties. The systematic contribution includes the theoretical uncertainties in the SM $t\bar{t}H$ cross-section described in Section 9.6. The measured signal strength in the single-lepton channel is:

$$\mu_{t\bar{t}H} = 0.72 \pm 0.12\ (\mathrm{stat.})\ ^{+0.21}_{-0.17}\ (\mathrm{syst.}), \tag{9.10}$$

and in the di-lepton channel is:

$$\mu_{t\bar{t}H} = 1.03 \pm 0.26\ (\mathrm{stat.})\ ^{+0.28}_{-0.22}\ (\mathrm{syst.}). \tag{9.11}$$

The combined $\mu_{t\bar{t}H}$ can be converted into an inclusive cross-section using the SM prediction of 507^{+35}_{-50} fb for $m_H = 125.09$, the result is:

$$\sigma_{t\bar{t}H} = 411^{+101}_{-92} \text{ fb} = 411 \pm 54 \text{ (stat.)} \, ^{+85}_{-75} \text{ (syst.) fb}, \tag{9.12}$$

with a relative uncertainty of $^{+25\%}_{-22\%}$, consistent with the SM expectation. The overall sensitivity is driven by the single-lepton channel, where systematic uncertainties dominate. In contrast, the di-lepton channel is primarily limited by the available dataset size.

The fitted normalisation factors for the $t\bar{t}+$ jets background processes, which are treated as free parameters in the likelihood fit, are listed in Table 9.8. The normalisation factors for the $t\bar{t}+1b$, $t\bar{t}+\geq 1c$, and $t\bar{t}+$light backgrounds are allowed to float independently in the single-lepton and di-lepton channels, while those for the $t\bar{t}+\geq 2b$ and $t\bar{t}+1B$ processes are constrained to be correlated across the two channels. This configuration is motivated by dedicated studies of different correlation assumptions in the combined fit. The adopted setup provides sufficient flexibility in the fit model while simultaneously minimising pulls and constraints on nuisance parameters when fitting to data. These values are consistent between the single-lepton and di-lepton channels within two standard deviations. For the $t\bar{t}+1b$ process, the normalisation factor is found to be slightly larger in the di-lepton channel than in the single-lepton channel.

Relative to the previous analysis of the same dataset, the current strategy selects 64% (29%) new events in the single-lepton (di-lepton) signal regions. The statistical correlation between the two analyses is estimated using a bootstrap method to be 19%, assuming independent systematic uncertainties. This assumption is justified by the fact that the dominant $t\bar{t}+\geq 1b$ background is treated with a different systematic model, and that experimental uncertainties have been updated. Based on this, the probability that the current measurement is compatible with the result reported in [16] is estimated to be 21%. The fitted normalisation factors for the $t\bar{t}+$jets background processes, treated as free parameters in the likelihood fit, are summarised in Table 9.8. The values are consistent between the single-lepton and di-lepton channels within two standard deviations. For the $t\bar{t}+1b$ process, the normalisation factor is found to be slightly larger in the di-lepton channel than in the single lepton channel.

Figure 9.22 shows the measured values of $\sigma_{t\bar{t}H}/\sigma_{\text{SM}}$ in bins of the Higgs boson transverse momentum, p_T^H, obtained from a combined profile-likelihood fit to data with an inde-

pendent $t\bar{t}H$ signal strength parameter in each bin. The measurement is consistent with the SM prediction, with a p-value of 89% including theoretical uncertainties in the SM $t\bar{t}H$ cross-section. The inclusive signal strength is also shown. The measurements in the [300, 450) GeV and [450, ∞) GeV p_T^H bins are limited by statistical uncertainties, while in the remaining bins the statistical and systematic contributions are comparable. The definition of boosted regions improves sensitivity in the [450, ∞) GeV bin by about 15% relative to an analysis based solely on resolved selections. The correlations between the parameters of interest are shown in Figure 9.23, where they remain below 30% and are larger at low p_T^H.

Normalisation factor	$t\bar{t}$+light	$t\bar{t}+\geq 1c$	$t\bar{t}+1b$	$t\bar{t}+1B$	$t\bar{t}+\geq 2b$
Single-lepton	$0.78^{+0.08}_{-0.08}$	$1.51^{+0.19}_{-0.18}$	$1.06^{+0.10}_{-0.10}$	$1.15^{+0.15}_{-0.14}$	$0.94^{+0.08}_{-0.08}$
Dilepton	$0.88^{+0.11}_{-0.10}$	$1.36^{+0.10}_{-0.10}$	$1.24^{+0.09}_{-0.09}$		

Table 9.8 Best-fit values of the $t\bar{t}$+jets normalisation factors obtained from the fit to data with a single inclusive $t\bar{t}H$ signal strength parameter. Independent normalisation factors for $t\bar{t}+1b$, $t\bar{t}+\geq 1c$ and $t\bar{t}$+light components are used in the dilepton and single-lepton channels. Before the fit, the $t\bar{t}$+light and $t\bar{t}+\geq 1c$ components are normalised to the $t\bar{t}$ cross-section computed at NNLO in QCD including the resummation of NNLL soft-gluon terms, while the $t\bar{t}+\geq 1b$ normalisation is taken from the $t\bar{t}b\bar{b}$ MC simulation.

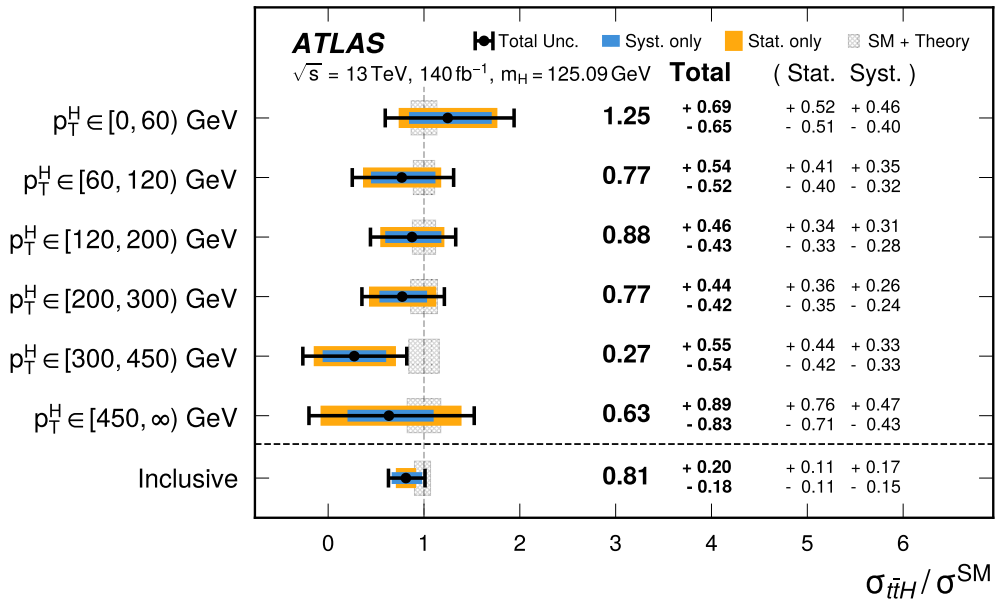


Figure 9.22 The $t\bar{t}H$ cross-sections measured in bins of the truth Higgs boson transverse momentum, p_T^H, for $|y| \leq 2.5$, and inclusively in the full phase space. All results are normalised to their SM predictions and obtained from a combined profile-likelihood fit to data in all signal and control regions. Uncertainties are shown separately for the measurement and for the theoretical prediction.

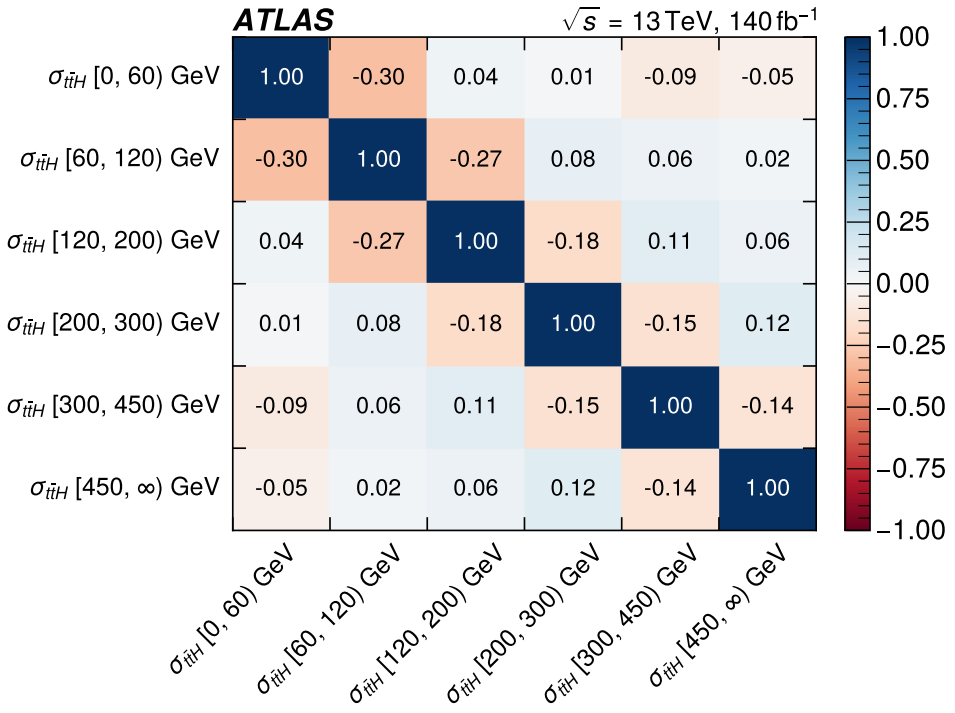
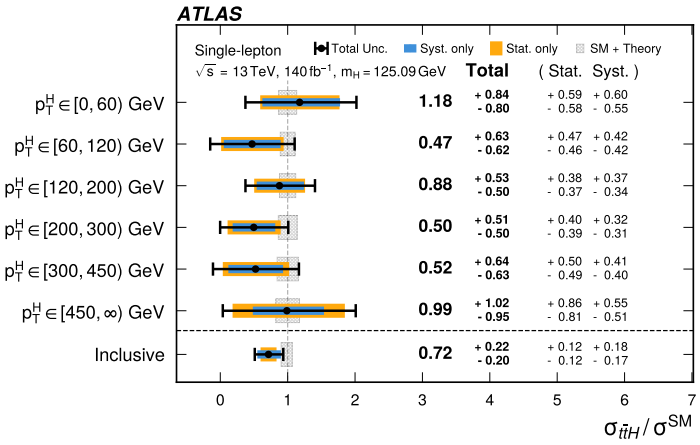


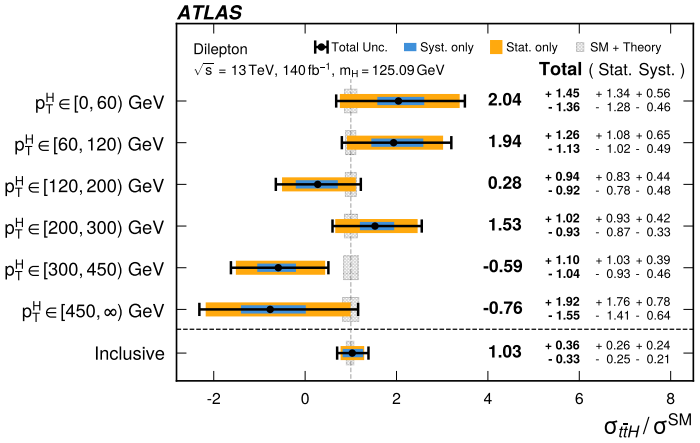
Figure 9.23 Post-fit correlations between the measured values of the $t\bar{t}H$ cross-section, $\sigma_{t\bar{t}H}$, in bins of the truth Higgs boson transverse momentum, p_T^H.

9.8.2 Differential

The individual channel fit with STXS breakdown is shown in Figure 9.24. The di-lepton channel is dominated by statistical uncertainties across all bins, demonstrating the weak statistical power of the channel.



(a)



(b)

Figure 9.24 The measured $t\bar{t}H$ cross-sections in bins of the Higgs boson transverse momentum, p_T^H, for $|y| \leq 2.5$, following the full STXS stage 1.2 binning scheme, together with the inclusive cross-section. All results are normalised to the SM prediction and obtained from a profile-likelihood fit to data in the single-lepton channel (a) and di-lepton channel (b), including all corresponding signal and control regions. The total uncertainty on each measured cross-section is indicated by the black error bar, with statistical and systematic components shown in yellow and blue, respectively. Theoretical uncertainties on the SM prediction are represented by the hashed grey band.

9.8.3 Summary Distribution

Post-fit yields of the $t\bar{t}H$ signal (S) and total background (B) are shown in Figure 9.25. They are shown as a function of $\log_{10}(S/B)$. Bins from all resolved and boosted regions in the single-lepton and di-lepton channels are merged according to their signal-to-background ratio. The signal is shown normalised to the best-fit result (red) and to the SM prediction (yellow), stacked on top of the total background. The lower panel displays the ratio of data to background compared with the best-fit prediction (solid red) and the SM expectation (dashed orange). The shaded band represents the total background uncertainty. Underflow and overflow bins are included.

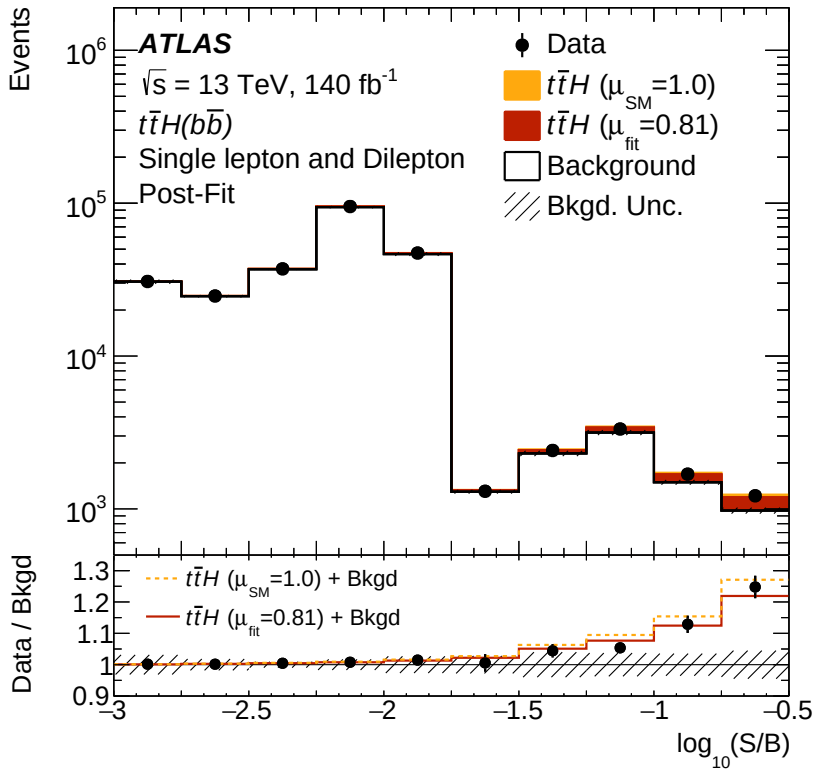


Figure 9.25 Observed and expected event yields as a function of $\log_{10}(S/B)$, where S and B denote the post-fit signal and total background yields, respectively. All bins from the fitted regions are ordered and merged according to their value of $\log_{10}(S/B)$. The signal is shown for both the best-fit signal strength, $\mu = 0.81$, and the Standard Model prediction, $\mu = 1.0$. The lower panel presents the ratio of the data to the post-fit background expectation, compared with the signal-plus-background prediction for the best-fit and SM signal strengths. The shaded band represents the total uncertainty in the background estimate. The first and last bins include the underflow and overflow contributions, respectively.

9.8.4 Post-Fit Modelling

The post-fit summary plots for the control regions and signal regions of the analysis are shown in the following distributions. Data is shown as black points, with the $t\bar{t}H$ signal contributions in red, scaled by $\sigma_{\text{obs}}/\sigma_{\text{SM}}$. The shaded uncertainty band represents the total statistical and systematic uncertainties, including their correlations. The lower panels display the ratio of data to the total post-fit yield. Excellent post-fit agreement is observed within the quoted uncertainties.

Single Lepton

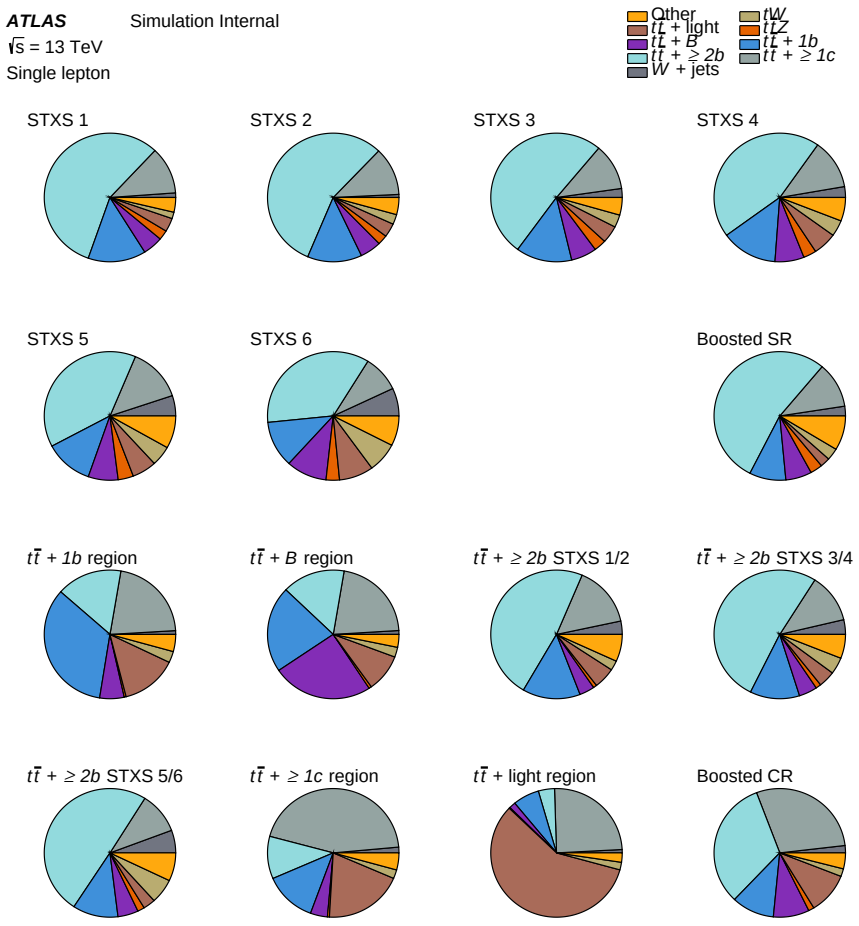


Figure 9.26 Composition of the signal and control regions of the single-lepton channel. The pie charts indicate the relative contributions of all processes, with one chart per analysis region. Each control region is enriched in the corresponding output class of the transformer network, while the signal region is subdivided into six bins of p_T^H.

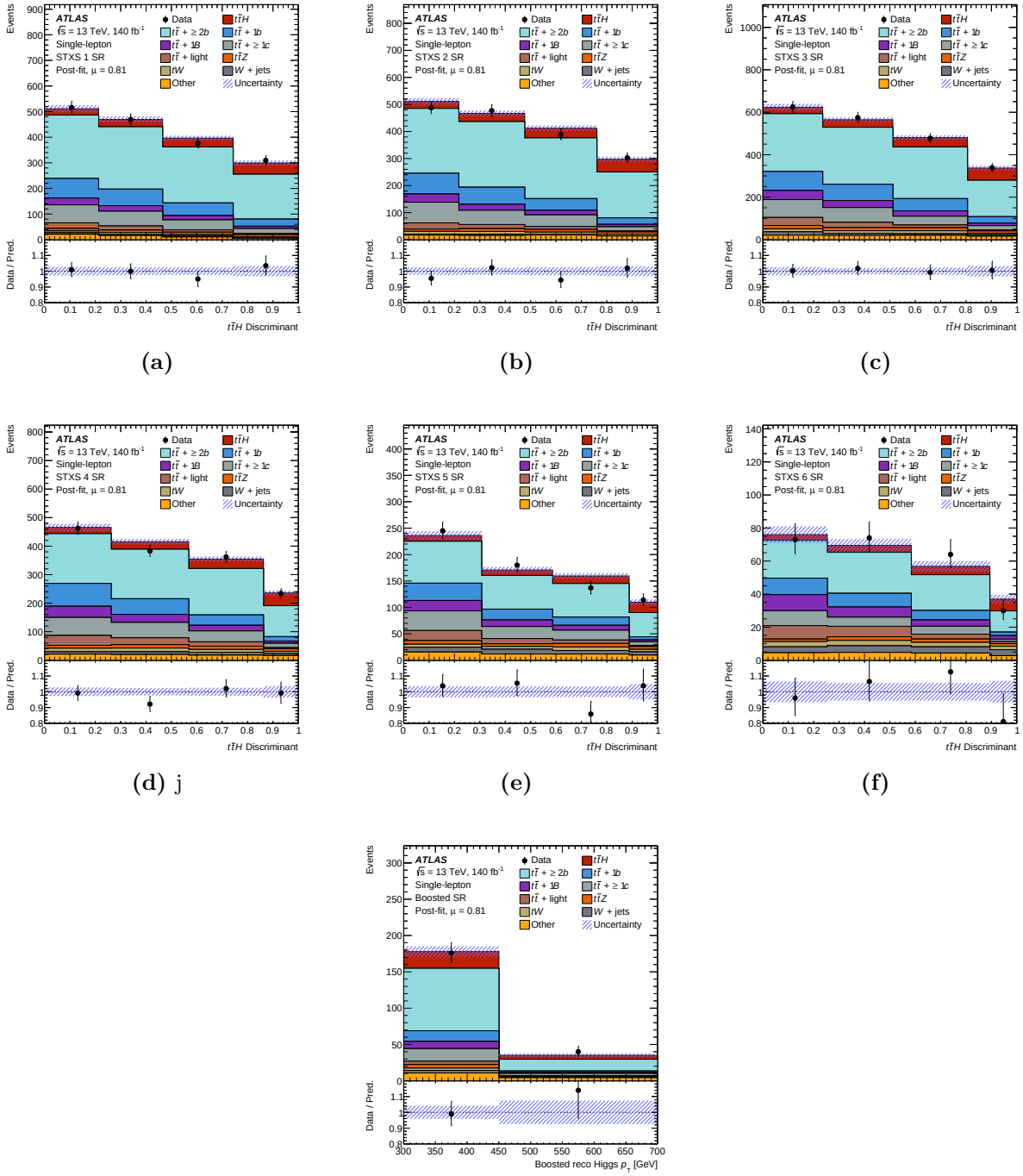


Figure 9.27 Post-Fit distributions of the seven signal regions in the single-lepton channel: (a) STXS 1, (b) STXS 2, (c) STXS 3, (d) STXS 4, (e) STXS 5, (f) STXS 6, (g) boosted signal region.

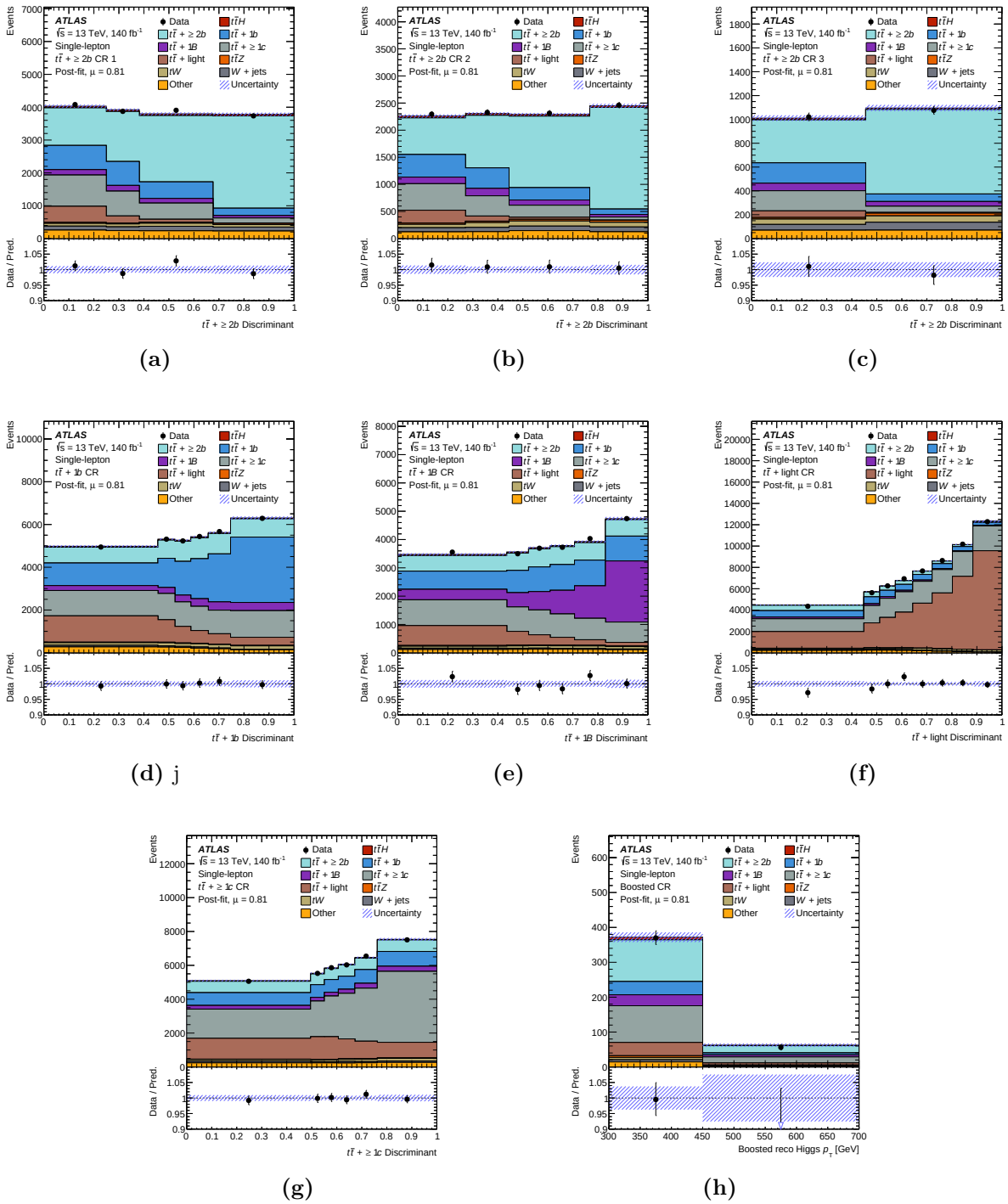


Figure 9.28 Post-fit distributions of the eight control regions in the single-lepton channel: (a) $t\bar{t}+\geq 2b$ STXS 12, (b) $t\bar{t}+\geq 2b$ STXS 34, (c) $t\bar{t}+\geq 2b$ STXS 56, (d) $t\bar{t}+1b$, (e) $t\bar{t}+1B$, (f) $t\bar{t}$+light, (g) $t\bar{t}+\geq 1c$, (h) boosted $t\bar{t}H$ control region.

9.8.5 Uncertainty Breakdown

The absolute and relative systematic uncertainties in the measurement of $\sigma_{t\bar{t}H}$, grouped by category, are summarised in Table 9.9, while the impact of individual nuisance parameters on the inclusive $\sigma_{t\bar{t}H}/\sigma_{\mathrm{SM}}$ is shown in Figure 9.29. The dominant contributions arise from signal modelling, followed by the modelling of the $t\bar{t}+\geq 2b$ background. In both cases, the largest effects are associated with FSR and the choice of parton-shower model. Among the experimental uncertainties, the leading contributions come from b-tagging and the jet energy scale.

Among the eight most influential nuisance parameters, the strongest constraints are observed for the $t\bar{t}+\geq 2b$ background uncertainties linked to the dipole shower model and the choice of matching scheme. These constraints originate primarily from the single-lepton channel, particularly in the signal regions and the $t\bar{t}+\geq 2b$ control region where this background is most significant. The constraints are further enhanced when the $t\bar{t}+\geq 2b$ control region is split into three subregions in the single-lepton channel, allowing the modelling to be probed in bins of the reconstructed Higgs boson transverse momentum.

Uncertainty source	$\Delta\sigma_{t\bar{t}H}$ (fb)		$\Delta\sigma_{t\bar{t}H}/\sigma_{t\bar{t}H}$ (%)	
Process modelling				
$t\bar{t}H$ modelling				
$t\bar{t}H$ radiation	+35	−21	+9	−5
$t\bar{t}H$ parton shower	+32	−19	+8	−5
$t\bar{t}H$ matching	<0.1	−0.3	<0.1	−0.1
$t\bar{t}H$ theory	+25	−17	+6	−4
$t\bar{t}+\geq 1b$ modelling				
$t\bar{t}+\geq 1b$ radiation	±31		±8	
$t\bar{t}+\geq 1b$ parton shower	±29		±7	
$t\bar{t}+\geq 1b$ matching	±19		±5	
$t\bar{t}+\geq 1c$ modelling	±18		±4	
$t\bar{t}+$ light modelling	±5		±1	
tW modelling	±16		±4	
Minor background modelling	±19		±5	
Flavour tagging	±36		±9	
Jet modelling	±22		±5	
Monte-Carlo statistics	±17		±4	
Other instrumental	±10		±2	
Total systematic uncertainty	+85	−75	+21	−18
Normalisation factors	±21		±5	
Total statistical uncertainty	±54		±13	
Total uncertainty	+101	−92	+25	−22

Table 9.9 Absolute and relative uncertainties in the measured $\sigma_{t\bar{t}H}$. Each entry corresponds to a distinct group of uncertainties, evaluated by re-running the fit with groups of nuisance parameters fixed to their best-fit values. Theoretical uncertainties include ISR/FSR, renormalisation and factorisation scale variations, STXS migrations, and $H \to b\bar{b}$ branching fraction. The total statistical uncertainty includes contributions from the free-floating $t\bar{t}$+jets normalisation factors.

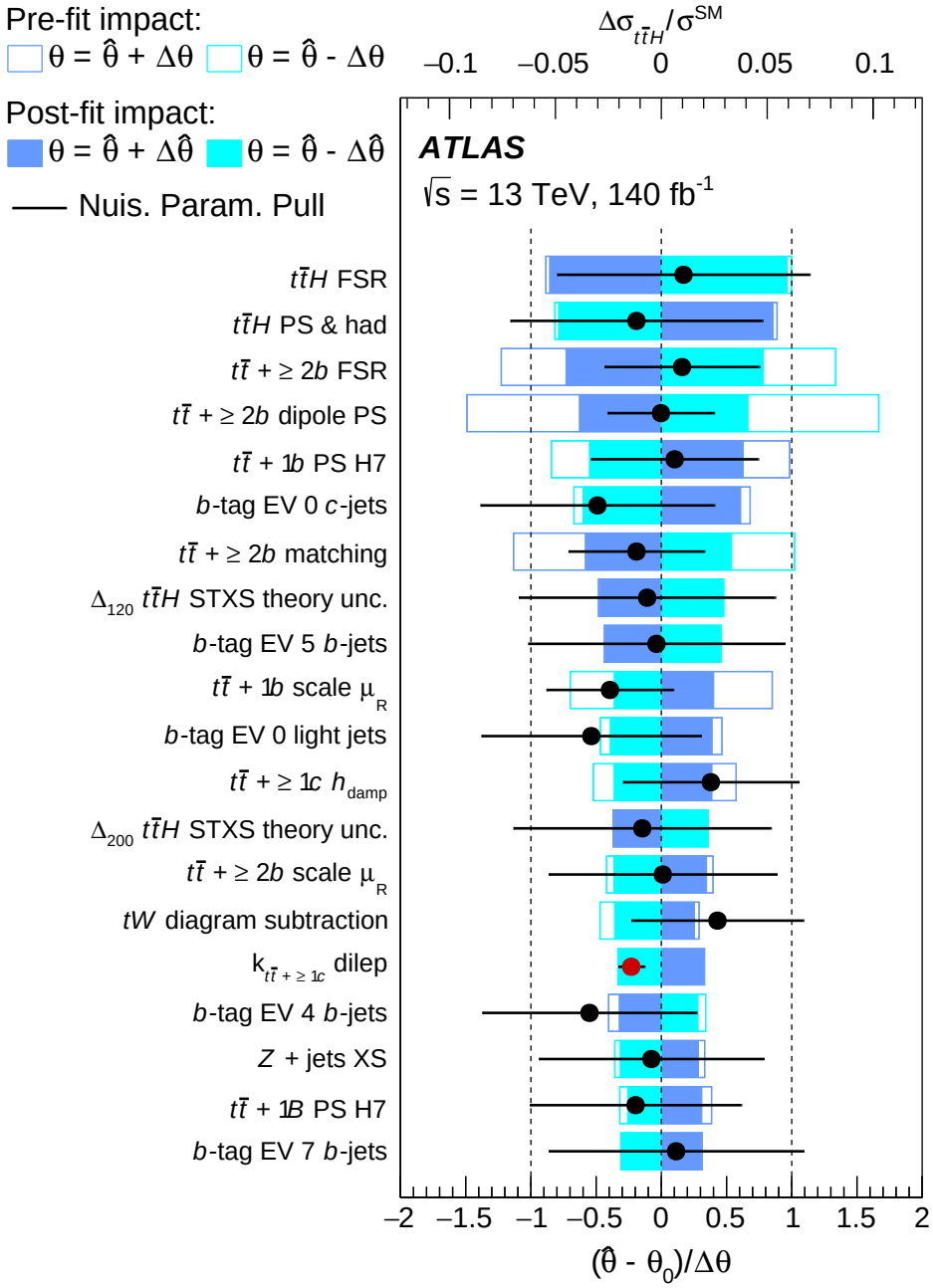


Figure 9.29 Ranking of the 20 modelling and experimental systematic uncertainties with the largest post-fit impact on the inclusive cross-section. The empty (filled) blue rectangles indicate the pre-fit (post-fit) impact on $\sigma_{t\bar{t}H}/\sigma_{\mathrm{SM}}$ and correspond to the upper axis. The impact of each nuisance parameter, $\Delta(\sigma_{t\bar{t}H}/\sigma_{\mathrm{SM}})$, is evaluated by comparing the nominal best-fit value with the result of the fit after fixing the nuisance parameter to its best-fit value $\hat{\theta}$ shifted by its pre-fit (post-fit) uncertainty $\pm\Delta\theta$ ($\pm\Delta\hat{\theta}$). Black markers show the pulls of the nuisance parameters relative to their nominal values $\theta_0 = 0$, while the red marker corresponds to the pull of the $t\bar{t}$+jets normalisation factors relative to their nominal value $\theta_0 = 1$. The pulls and their relative post-fit uncertainties, $\Delta\hat{\theta}/\Delta\theta$, are shown on the lower axis.

Chapter 10

$t\bar{t}H(b\bar{b})$ EFT Interpretation

"The S&P 500 is my favourite meme coin."
— Hannah Mehravari, after the Trump tariffs, April 2025.

With the $t\bar{t}H(b\bar{b})$ legacy analysis producing the most precise measurement of its kind to date, a SMEFT interpretation was used to test for the potential presence of BSM effects. Using this analysis we can update the precision on the Wilson Coefficients (WC), which determine the strength of the SMEFT operators.

10.1 Early Run-2 Results

Before we proceed with the analysis, we briefly review the $t\bar{t}H$ EFT precision measurement landscape. The main point of comparison for our results will be a early run-2 result from 2016 [143]. The analysis in question used a combination of results from both the CMS and ATLAS collaboration. In order to break the degeneracy between the operators the authors used H, $H + j$ and $t\bar{t}H$ analyses: the $pp \to H$ measurements used are the following: in the diphoton channel we use [144, 145] at 8 TeV and [146] at 13 TeV; in the $WW/ZZ/\tau\tau$ channels we use [147, 148, 149, 150, 151, 152]; measurements of $pp \to t\bar{t}H$ at 8 TeV in the diphoton channel are given in [153, 154], while in the multilepton and $b\bar{b}$ channels we use [153, 155, 156]. Finally, 13 TeV measurements of $t\bar{t}H$ in the multilepton and $b\bar{b}$ channels are also included [157, 158]. Both processes depend on the three operators discussed in Chapter 2. The result of this combination and re-interpretation, in terms of the most relevant operators for $t\bar{t}H$, is shown in Table 10.1. From these results we

observe constraints of $c_{t\phi}$, $c_{\phi G}$ and c_{tG} from singular WC fits, known as individual fits, and a combination of all three WCs known as combined fits. Comparing individual and combined fits we see that there is loss in precision when trying to fit more than one WC at a time, This can be attributed to the inherent degeneracy between the WCs. This degeneracy occurs when there is a set of observables and a set of operators that are correlated. This correlation leads to a loss in precision as we don't have the statistical power to separate the effects of each operator. With the addition of H and $H + j$ data, this study managed to reduce the loss in precision[i].

	Individual	c_{tG} Fixed	Combined
$c_{t\phi}$	[-3.9,4.0]	[-12,20]	[-14,31]
$c_{\phi G}$	[-0.0072,-0.0063]	[-0.022,0.031]	[-0.021,0.054]
c_{tG}	[-0.68,0.62]	-	[-1.8,1.6]

Table 10.1 The final precision obtained by the early Run-2 EFT analysis [143]. Levels of precision are presented for fits of individual parameters (left), a two parameter fit where c_{tG} is fixed (middle) and all three parameters left free in the fit (right).

[i] A total loss in precision would be a statement such as: I have 15 coins in my hand ±500. A physicists guttural reaction would of course be: negative coins are unphysical and not even a gorilla could hold 500 coins in their hand.

10.2 Analysis Methodology

For any EFT fit we first need a Monte Carlo sample with an EFT prediction in order to fit our data on to it. This is usually performed with dedicated samples created using a generator of choice. However, as we are using the finalised published analysis we can modify the provided SM samples to include an EFT prediction. We plan to re-interpret our SM result in the SMEFT framework, which can be summarised as a change of basis from μ (signal strength) to our desired WC and so our WC becomes our POI. Using cross-section weights, which provide different values per STXS region, provided by the ATLAS global EFT analysis we can read these into our signal samples and create an EFT prediction for our desired EFT WC per STXS bin [159]. With these predictions in hand we can then fit our data and extract the values of a given WC or multiple WCs in a combined fit. A flow diagram of this workflow can be seen in Figure 10.1.

When performing the fit we use `TRExFitter` just as in $t\bar{t}H(b\bar{b})$ legacy analysis. The same `TRExFitter` configuration file is also used with minimal changes made to it in order to incorporate the EFT weights and change of POI from μ to WC.

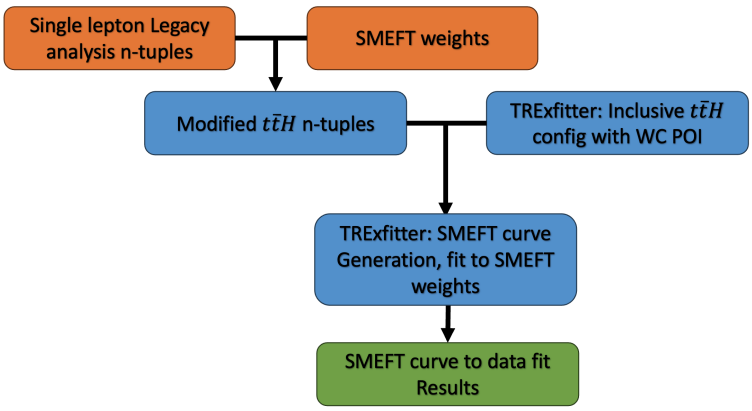


Figure 10.1 The stages of processing (orange), fitting (blue) and results (green) for the EFT fits.

10.2.1 EFT Weights

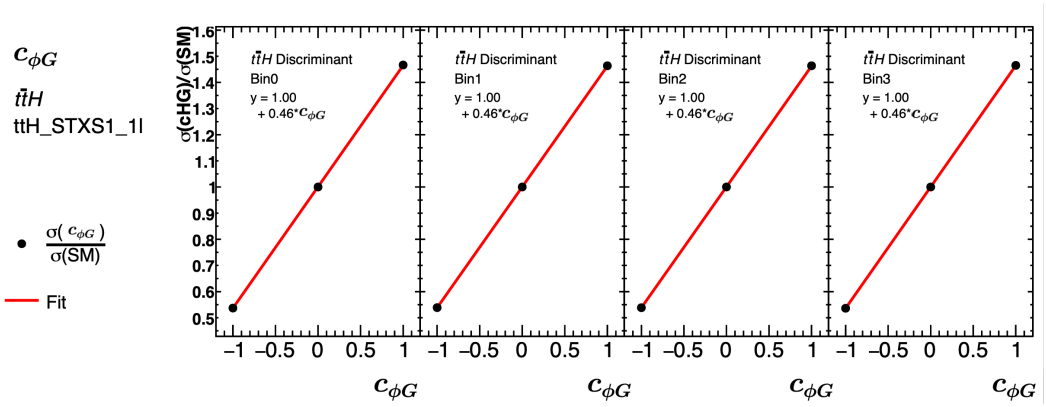
The EFT weights are used to modify the cross section as described in Equation 2.23. We read these weights into the SM sample are split by STXS bin and if they are linear or

quadratic in nature. The linear weights only include the SM-EFT mixing contributions where as the quadratic weights include contributions from both the mixing term and pure BSM term. The weights themselves are extracted from existing MC samples containing EFT predictions in them and are provided by the ATLAS global EFT analysis. The samples are generated using `MadGraph5_aMC@NLO` with two different Universal FeynRules Output (UFO) models: `SMEFTsim` for tree-level diagrams and most production/decay processes and `SMEFTatNLO` for ggF, computed at NLO in QCD. `Pythia8` is used for parton showering, hadronization, and Higgs decay modelling. A Rivet framework (with modifications) is used for the classification of events into STXS regions. The effects of each operator per STXS bin can be isolated and extracted as a weight on the SM STXS bin (either increasing or suppressing the cross-section). Throughout the fits presented, many validation and compatibility studies were conducted. No significant deviation in background normalisation factors were observed, nor were there noticeable pulls or deviations in nuisance parameters that weren't present in the baseline SM fit.

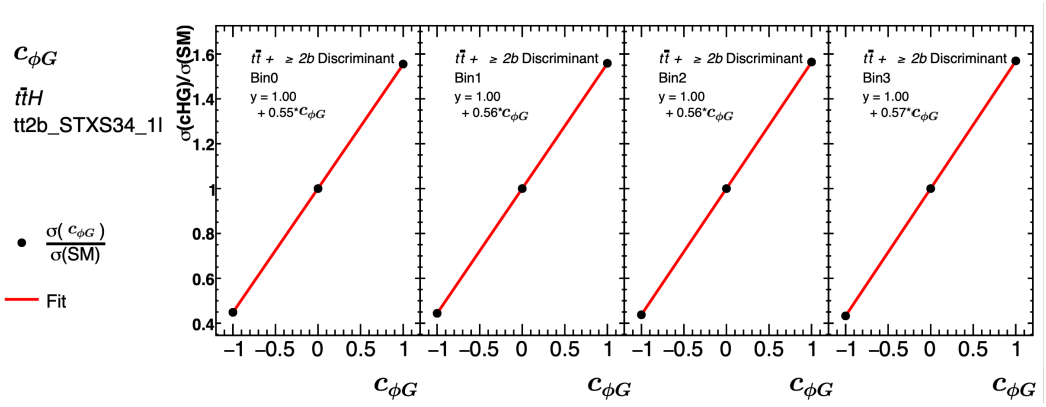
10.3 Results

10.3.1 Linear

The following results presented are the linear only fits for individual WCs and finally a combination of all three WCs. Each bin in each region requires a fit to the prediction and then a fit to the data to provide the final result. The EFT prediction is fit to a linear function in the form of $y = mx + c$, where y is the ratio of EFT and SM cross-sections ($\frac{\sigma(\text{WC})}{\sigma(\text{SM})}$), x is the WC, m is the modification to the WC and c is the y-axis intercept. For all linear fits c will always be 1, as $\frac{\sigma(\text{WC=0})}{\sigma(\text{SM})} = \frac{\sigma(\text{SM})}{\sigma(\text{SM})} = 1$. An example plot of linear fit for a signal (background) region can be seen in Figure 10.2a (Figure 10.2b). The results for these individual fits will be summarised in tables for the remainder of the linear results section. Background results are omitted from this discussion as they aren't relevant to the discussion at hand but served as a cross-check that the fits were behaving as expected. The linear fits will serve as a test as to whether or not we can resolve the contributions of the three main WCs in combination. As such the results provide a window on the effect of the improved $t\bar{t}H$ analysis on the relevant WCs to be compared to the early Run-2 results shown in Table 10.1.



(a) Example signal region linear fit for $c_{\phi G}$, STXS1.



(b) Example background region linear fit for $c_{\phi G}$, $t\bar{t} \geq 2b$ STXS 34.

Figure 10.2 Linear fits of $c_{\phi G}$ to the EFT prediction in representative signal (top) and background (bottom) regions fitted to the discriminant. The black points are the ratio $\sigma(c_{\phi G})/\sigma(\text{SM})$, and the red line shows the linear fit.

$c_{\phi G}$

The fitted value of $c_{\phi G}$ in the linear regime is $-0.50^{+0.40}_{-0.35}$. As discussed earlier $c_{\phi G}$ isn't solely constrained by the $t\bar{t}H$ production process and can be more constrained further in global fits [159]. As such we don't come close to the early Run-2 result, in terms of precision, due to the inherent low statistics of our single analysis fit.

c_{tG}

The fitted value of c_{tG} in the linear regime is $0.35^{+0.40}_{-0.28}$. As with $c_{\phi G}$, c_{tG} is better served in global fits as other production modes with will contribute to its precision.

$c_{t\phi}$

For $c_{t\phi}$ the fit yields a central value of $2.35^{+1.78}_{-1.95}$ is observed. While also consistent with the SM, the precision here is significantly improved that with early Run-2 results shown in Table 10.1. This is fit is still however still in the linear regime and not fit to the full theory.

Combined Results

Having examined the impact of each Wilson coefficient individually, we now perform a simultaneous fit to $c_{\phi G}$, c_{tG} and $c_{t\phi}$ within the linear approximation. This allows us to assess the interplay between the operators and quantify the correlations that arise when they are floated together. The combined fit provides a more realistic picture of the sensitivity achievable in $t\bar{t}H$ alone. However, we observe a loss of precision for all WCs compared to their individual fits as seen in Table 10.2. This is due to degeneracy observed in Figure 10.4. All combinations of two WCs were tested but with a similar loss in precision due to this degeneracy.

WC	Central Value	Uncertainty
$c_{\phi G}$	1.88	[-7.57,7.64]
c_{tG}	3.05	[-0.50,0.50]
$c_{t\phi}$	-9.80	[-31.70,31.18]

Table 10.2 The results for a linear three WC combination fit.

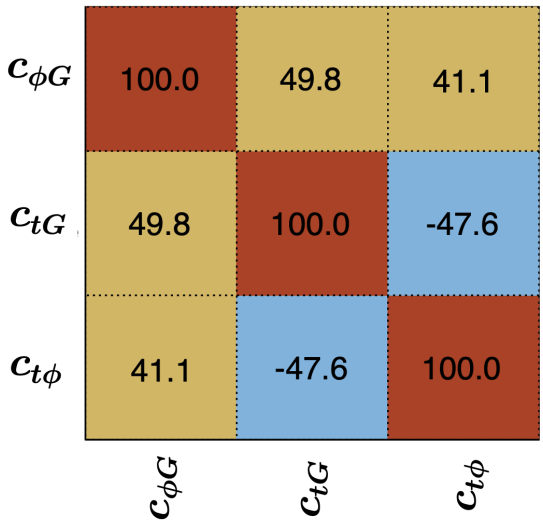


Figure 10.3 The correlations of the three WCs in the combination fit.

10.3.2 Degeneracy in the Combined Fit

Figure 10.4 shows the predicted EFT modifications to the Higgs transverse-momentum spectrum, $d\sigma/dp_T^H$. Although c_{tG} and $c_{t\phi}$ enter the theory in different ways, their impact on $d\sigma/dp_T^{Higgs}$ is highly similar across the range probed here. In particular, both act largely as rescaling of the spectrum with only modest shape variations at moderate p_T^H. Within the current statistical uncertainties these distortions overlap, so the likelihood can be equally well described by a positive shift in c_{tG} compensated by a negative shift in $c_{t\phi}$ (and vice versa). This trade-off produces the strong anti-correlation seen in Figure 10.3 and leads to the loss of precision for both coefficients in the simultaneous fit. While $c_{t\phi}$ is most affected, showing a substantial broadening of its allowed interval, c_{tG} also suffers reduced precision compared to its individual fit, underscoring the difficulty of disentangling their contributions via $t\bar{t}H$ alone. In contrast, $c_{\phi G}$ generates a harder spectrum at high p_T^H, providing a distinct handle that reduces its degeneracy with the other operators.

In contrast, the early Run-2 EFT interpretations combined information from several analyses, including inclusive $t\bar{t}$, single-top, and Higgs production channels. By exploiting the complementary sensitivities of these processes, the fits were able to break the degeneracy between c_{tG} and $c_{t\phi}$ that is evident in $t\bar{t}H$ alone. This highlights the role of global analyses in disentangling correlated operators and achieving tighter overall constraints.

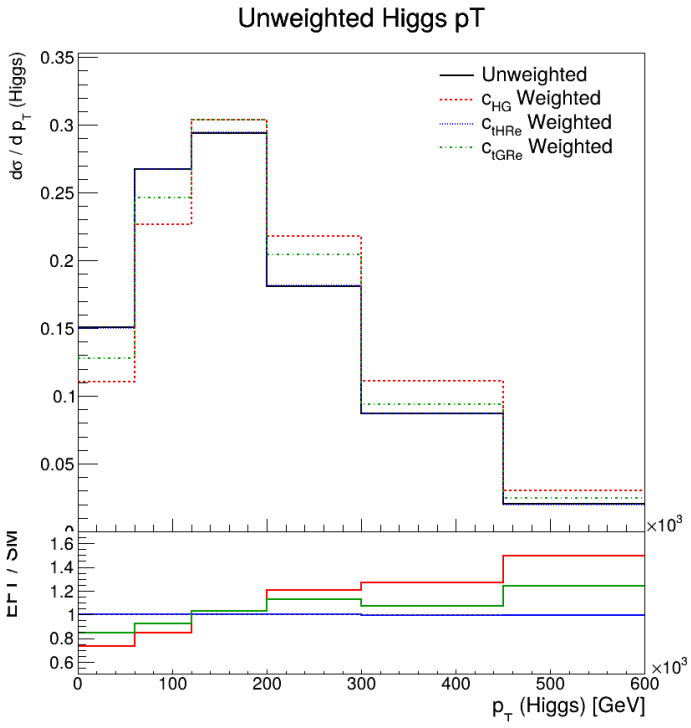


Figure 10.4 Comparing the effects of different WCs to the Higgs p_T: unweighted Higgs p_T (SM, Black), $c_{\phi G}$ (red), $c_{t\phi}$ (blue) and c_{tG} (green).

10.3.3 Quadratic

For the quadratic fits the full EFT dependence is taken into account, and these results should be regarded as the final values for comparison. As already seen in the linear case, attempting to extract all three coefficients simultaneously leads to a sharp loss of precision because of the strong degeneracies between them. The early Run-2 interpretations alleviated this by combining several processes, which provided additional handles to separate the operator effects. In contrast, the present study relies only on the single-lepton $t\bar{t}H$ dataset and therefore lacks the statistical power needed to break these correlations. For this reason no combined quadratic fit is presented; the discussion instead centres on the individual WC fits and on comparing their achieved precision with the corresponding early Run-2 results. The following fits were evaluated using the saturated model goodness of fit method [160].

$c_{\phi G}$

$$c_{\phi G} = -0.16^{+0.23}_{-0.14} = -0.16^{+0.22}_{-0.12}(\text{stat.})\ ^{+0.08}_{-0.07}(\text{syst.}). \tag{10.1}$$

The precision obtained in this analysis is lower than that achieved in the early Run-2 EFT studies. The main reason is the limited by the statistics of the single dataset considered here, restricted to the single-lepton $t\bar{t}H$ channel. The goodness of fit was found to be 0.80.

c_{tG}

$$c_{tG} = 0.20^{+0.16}_{-0.21} = 0.20^{+0.13}_{0.20}(\text{stat.})\ ^{+0.07}_{-0.07}(\text{syst.}). \tag{10.2}$$

In contrast to the behaviour of $c_{\phi G}$, the individual quadratic fit for c_{tG} shows a noticeable improvement over the early Run-2 results. The gain in precision comes from two main factors: the larger dataset now available in the single-lepton $t\bar{t}H$ analysis, and a more refined treatment of the high-p_T regions that are especially sensitive to the chromomagnetic operator. Together these enhancements shrink the uncertainty interval on c_{tG} relative to the Run-2 study and highlight the improved discriminating power of this channel on its own. Although global fits are still required to cleanly separate the effects of different operators, these results emphasise the growing impact of the $t\bar{t}H$ channel in directly constraining c_{tG}. When splitting the final uncertainty into statistical and systematic contribution, as shown in Table 10.5, we see that the statistical contributions are the dominant factor. The goodness of fit was found to be 0.81

$c_{t\phi}$

$$c_{t\phi} = 2.55^{+2.22}_{-2.15} = 2.55^{+1.46}_{-1.37}(\text{stat.})\ ^{+1.61}_{-1.60}(\text{syst.}). \tag{10.3}$$

The quadratic fit to $c_{t\phi}$ shows a clear gain in precision compared with the early Run-2 results. This is a notable outcome, since $t\bar{t}H$ is the only channel that provides direct sensitivity to $c_{t\phi}$, making the improvement both unique and valuable. In other words, the stronger constraint achieved here highlights the growing role of $t\bar{t}H$ in testing the top-Higgs Yukawa interaction. Looking at the breakdown of the uncertainty, the statistical

and systematic components are now of similar size. That balance shows the measurement is no longer limited purely by the amount of data, but is equally shaped by systematic effects. This is an encouraging sign of maturity for the $t\bar{t}H$ analysis, and a pointer to where future improvements will need to focus. The goodness of fit was found to be 0.83.

10.3.4 Omitting the High p_T Regions

As the weights provided are split by STXS region, the created predictions are also split into STXS regions. Doing so reveals how each region evolves and their regime changes from being dominated by the linear coefficient to the quadratic coefficient becoming more present as shown in Table 10.3. Thus, is worth investigating how different STXS regions contribute to the sensitivity of the fitted WC. Intuition would suggest that the higher STXS region EFT predictions would contain a more dominant quadratic regime and hence these regions would be a more sensitive, as discussed earlier. A quick test for this would be to remove these regions (STXS 5,6 and Boosted) from our fit and see if the affect on the central value and precision of our fitted WC. We expect c_{tG} and $c_{\phi G}$ to show a shift in central value and a significant loss in precision as they become more sensitive to the quadratic regime at higher p_T. Notably, $c_{t\phi}$ is unaffected in this regard, as we do not observe a shift in the linear or quadratic modifications as we increase in p_T.

Region	**Bin**	$c_{\phi G}$		$c_{t\phi}$		c_{tG}	
		Lin	**Quad**	**Lin**	**Quad**	**Lin**	**Quad**
STXS 1	1	-0.845	0.589	-0.123	0.004	0.465	0.296
	2	-0.843	0.579	-0.123	0.004	0.462	0.287
	3	-0.843	0.579	-0.123	0.004	0.463	0.287
	4	-0.845	0.587	-0.123	0.004	0.464	0.291
STXS 2	1	-0.881	0.688	-0.123	0.004	0.498	0.363
	2	-0.880	0.684	-0.123	0.004	0.498	0.360
	3	-0.879	0.678	-0.123	0.004	0.497	0.353
	4	-0.881	0.684	-0.123	0.004	0.499	0.356
STXS 3	1	-0.935	0.940	-0.123	0.004	0.554	0.561
	2	-0.938	0.948	-0.123	0.004	0.556	0.565
	3	-0.941	0.956	-0.123	0.004	0.559	0.568
	4	-0.948	0.980	-0.123	0.004	0.566	0.581
STXS 4	1	-0.982	1.452	-0.123	0.004	0.607	1.014
	2	-0.988	1.504	-0.123	0.004	0.614	1.056
	3	-0.995	1.564	-0.123	0.004	0.621	1.106
	4	-1.010	1.669	-0.123	0.004	0.636	1.190
STXS 5	1	-0.988	2.424	-0.123	0.004	0.644	2.170
	2	-0.990	2.545	-0.123	0.004	0.651	2.282
	3	-0.997	2.694	-0.123	0.004	0.663	2.427
	4	-1.005	2.958	-0.123	0.004	0.679	2.700
STXS 6	1	-1.032	4.937	-0.122	0.004	0.696	6.436
	2	-1.041	5.377	-0.122	0.004	0.708	7.147
	3	-1.051	5.788	-0.122	0.004	0.724	7.742
	4	-1.073	6.792	-0.122	0.004	0.755	9.315

Table 10.3 Merged linear and quadratic modifications for $c_{\phi G}$, c_{tG}, and $c_{t\phi}$ in all signal regions.

$c_{\phi G}$

$c_{\phi G}$	**Central Value**	**Uncertainty**	**Stat.**	**Syst.**
	-0.17	[-0.27, 0.27]	[-0.26,0.25]	[-0.08,0.08]

Table 10.4 EFT fit to data for the $c_{\phi G}$ WC, central value and precision.

For $c_{\phi G}$, the fit after removing the high-p_T regions yields a central value of -0.17 with an uncertainty of ± 0.27. The shift in the best-fit point relative to the full fit is small, indicating that the central value remains stable. However, as expected, the overall precision is reduced when the boosted STXS bins are excluded, reflecting the important role of those regions in constraining $c_{\phi G}$. The decomposition of the uncertainty shows that both statistical and systematic components contribute at a similar level, with the statistical term dominating but systematics already playing a non-negligible role. This confirms that the boosted categories provide unique sensitivity to quadratic effects in $c_{\phi G}$, and that removing them leads directly to a measurable loss in precision.

c_{tG}

c_{tG}	**Central Value**	**Uncertainty**	**Stat.**	**Syst.**
	0.21	[-0.24, 0.27]	[-0.22,0.25]	[-0.11,0.11]

Table 10.5 EFT fit to data for the c_{tG} WC, central value and precision.

For c_{tG} compared to the full fit, the central value is relatively stable, but the uncertainty band broadens, showing a clear loss of precision once the high-p_T regions are removed. This behaviour matches the expectation that c_{tG} is particularly sensitive to the higher-p_T bins, where quadratic contributions become more significant. The breakdown of the uncertainty reveals comparable contributions from the statistical and systematic components, with systematics already limiting the precision at a level similar to the data statistics. Overall, this result demonstrates that the high-p_T regions provide unique discriminating power for c_{tG}, and that excluding them reduces the effectiveness of the fit.

$c_{t\phi}$

$c_{t\phi}$	**Central Value**	**Uncertainty**	**Stat.**	**Syst.**
	2.41	[-2.27, 2.38]	[-1.47,1.61]	[-1.73,1.75]

Table 10.6 EFT fit to data for the c_{tG} WC, central value and precision.

For $c_{t\phi}$, the fit after removing the high-p_T regions yields a central value of 2.41 with an uncertainty of $^{+2.38}_{-2.27}$. In contrast to $c_{\phi G}$ and c_{tG}, the precision here does not show a

significant loss when the boosted categories are excluded. This behaviour is consistent with expectations, as the impact of $c_{t\phi}$ is largely rate-like and does not become more pronounced at higher Higgs p_T. The decomposition of the uncertainty shows that the statistical and systematic contributions are of comparable size, with both effects limiting the overall precision. This result confirms that, unlike the other operators, $c_{t\phi}$ does not rely strongly on the high-p_T bins for sensitivity, and the stability of the fit reflects the different nature of its modification to the $t\bar{t}H$ process.

10.4 Discussion

The results presented in this chapter demonstrate that the $t\bar{t}H(b\bar{b})$ channel has advanced into a new regime of sensitivity for constraining dimension-six operators in the SMEFT framework, most notably for c_{tG} and $c_{t\phi}$. While the early Run-2 EFT interpretations already highlighted the potential of $t\bar{t}H$, the present analysis pushes the precision of these measurements further, thanks to both the increased dataset and the explicit improvements of the inclusion of high-p_T boosted and resolved Higgs signal regions.

For c_{tG}, the chromomagnetic operator, the improvement is particularly interesting. The central value remains close to the SM expectation, but the allowed interval has contracted relative to the early Run-2 results, indicating a genuine gain in precision. This enhancement can be traced to the high-p_T STXS bins (STXS 5, 6 and the boosted category), where the quadratic contributions of c_{tG} become increasingly important. By improving the treatment of these signal regions, the analysis adds statistical power to improve the overall precision. The resulting constraints on c_{tG} establish $t\bar{t}H$ as one of the most sensitive direct probes of this operator, and underscore the value of exploiting the boosted Higgs phase space despite its lower event yield.

The situation for $c_{t\phi}$, the Yukawa-like operator, is more nuanced. Here too the precision obtained represents a significant advance over the early Run-2 results, and this constitutes one of the most important outcomes of the present study. Since $c_{t\phi}$ modifies the Higgs-top Yukawa interaction directly, the $t\bar{t}H$ process is essentially unique in its ability to constrain this coefficient. The tighter limits achieved here therefore reflect an exclusive strength of this channel. However, the gain does not originate from the boosted bins in the same way as for c_{tG}. Within the kinematic reach and statistical power of the current dataset, the effect of $c_{t\phi}$ continues to appear largely as a rescaling of the overall $t\bar{t}H$ rate. Although

in principle this operator can induce shape modifications at very high energies, these are not resolved in the Run-2 dataset. The apparent insensitivity of $c_{t\phi}$ to the high-p_T bins is thus a reflection of the limited statistics and energy range probed, rather than a fundamental property of the operator itself. In this sense, the improvement in precision is driven primarily by the increased dataset and refined systematic treatment, rather than by the addition of the boosted categories.

In a broader context, these findings highlight both the promise and the limitations of using single-channels alone to constrain SMEFT operators, such as our $t\bar{t}H(b\bar{b})$ single lepton channel used here. On the one hand, the improvements in c_{tG} and $c_{t\phi}$ demonstrate that this channel can independently deliver meaningful constraints on operators that directly probe the top-Higgs sector. On the other hand, the degeneracies observed in the simultaneous fits underline the necessity of global combinations across multiple production and decay channels in order to fully disentangle correlated operator effects and break the degeneracy. Early Run-2 studies made this point by showing how combined analyses could break degeneracies that are unavoidable in single-channel fits. The present results extend that narrative: $t\bar{t}H(b\bar{b})$ alone has now reached a level of precision where its contribution to global fits is not only complementary, but competitive for two of the three variables.

Looking forward, the lessons from this analysis point clearly to the HL-LHC. With an order-of-magnitude increase in statistics, boosted Higgs categories will become densely populated, and the shape effects of $c_{t\phi}$ at the TeV scale may finally emerge. At the same time, systematic uncertainties, already comparable to statistical errors in the present study, will become a sigficant limitation. Continued advances in theoretical modelling, Monte Carlo generation, and experimental techniques will therefore be essential. Nevertheless, the progress achieved here demonstrates that the $t\bar{t}H(b\bar{b})$ channel is not only a unique probe of the Higgs-top Yukawa interaction, but also a vital component in the precision SMEFT programme at the high-energy frontier.

Chapter 11

Conclusion and Outlook

"When you do things right, people won't be sure you've done anything at all."

— Space God, providing perspective to Bender the robot.

The associated production of a Higgs boson with a top-quark pair was measured in the $H \to b\bar{b}$ decay channel, using the full Run 2 ATLAS dataset of 140 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV. Final states with one or two leptons were analysed, leading to an inclusive cross-section measurement of:

$$\sigma_{t\bar{t}H} = 411 \pm 54 \text{ (stat.)} \, ^{+85}_{-75} \text{ (syst.) fb},$$

for a Higgs boson mass of 125.09 GeV, consistent with the Standard Model prediction of 507^{+35}_{-50} fb. An excess of events over the background expectation was observed with a significance of 4.6 standard deviations (5.4 expected). The cross-section was also measured differentially in bins of p_T^H within the STXS framework, with the highest bin probing Higgs bosons above 450 GeV. These measurements establish the $t\bar{t}H(b\bar{b})$ channel as one of the most precise single-channel probes of the Higgs–top interaction.

This precision is enabled by several key improvements compared to earlier analyses. Looser selection requirements and enhanced b-tagging raised the overall signal acceptance, while a powerful multiclass neural network was used to define control regions enriched in different $t\bar{t}$+jets components. Dedicated simulation and systematic models for the $t\bar{t}+1b$ background, along with data-driven corrections for the $t\bar{t}+1c$ and $t\bar{t}$+light categories. New high p_T extrapolation uncertainties were also implemented providing more robust

statistical treatment for PFlow jets above their respective flavour calibration threshold. These provided greater control of backgrounds and significantly improved sensitivity. These developments not only delivered the most precise inclusive and differential $t\bar{t}H$ measurement in this channel to date, but also enabled competitive SMEFT re-interpretations, particularly constraining the chromomagnetic operator:

$$c_{tG} = 0.20^{+0.16}_{-0.21} = 0.20^{+0.13}_{0.20}(\text{stat.})\ ^{+0.07}_{-0.07}(\text{syst.}),$$

and the Yukawa-like operator:

$$c_{t\phi} = 2.55^{+2.22}_{-2.15} = 2.55^{+1.46}_{-1.37}(\text{stat.})\ ^{+1.61}_{-1.60}(\text{syst.}),$$

which sees an improvement in its precision by a factor of 2. At the same time, the analysis highlights important limitations. The $t\bar{t}H(b\bar{b})$ measurement is dominated by systematic uncertainties, especially those from $t\bar{t}H$ signal modelling and $t\bar{t}$+jets backgrounds. In the EFT interpretation, the gain in c_{tG} precision originates from boosted and high-p_T bins, where quadratic contributions become relevant, whereas $c_{t\phi}$ remains largely constrained through rate effects due to the limited statistical reach of Run 2 data. This shows that while $t\bar{t}H(b\bar{b})$ alone provides meaningful constraints, it does not provide the high energy statistical power needed to resolve the unique effects of $c_{t\phi}$'s governing operator. Not to mention the observed degeneracy between operators, underlining the need for global combinations across multiple production and decay modes to fully disentangle correlated effects.

Looking forward, the HL-LHC offers a clear path to extend these results. With an order-of-magnitude increase in statistics, boosted Higgs categories will become well populated, allowing sensitivity to the high-energy shape effects of operators like $c_{t\phi}$. To match this statistical reach, advances in theoretical modelling, Monte Carlo generation, and experimental techniques will be essential, as systematic uncertainties are already a limiting factor. These improvements should hopefully lead us to new physics at the energy frontier.

Appendix A

Auxiliary Plots

A.1 Pre-Fit Modelling

For the di-lepton channel, the background composition is shown in Figure A.1. The same general pattern is observed, with the $t\bar{t}$+light and $t\bar{t}+ \geq 2b$ regions exhibiting high purity. Here, the $t\bar{t}+ \geq 1c$ process is more cleanly separated than in the single lepton case, benefiting from the additional constraints of the di-lepton final state. As in the single lepton channel, the $t\bar{t}+ \geq 2b$ background remains the dominant contribution in the signal regions, with $t\bar{t} + 1b$ and other processes becoming more relevant in the low- and high-p_T^H bins, respectively.

For the di-lepton channel, five control regions are used, as illustrated in Figure A.3. Again, the ratio of data to MC predictions indicates a very good level of pre-fit agreement within the defined uncertainties. The six resolved signal regions are shown in Figure A.2, with similarly strong pre-fit agreement. No boosted signal region is defined in the di-lepton channel due to limited statistics.

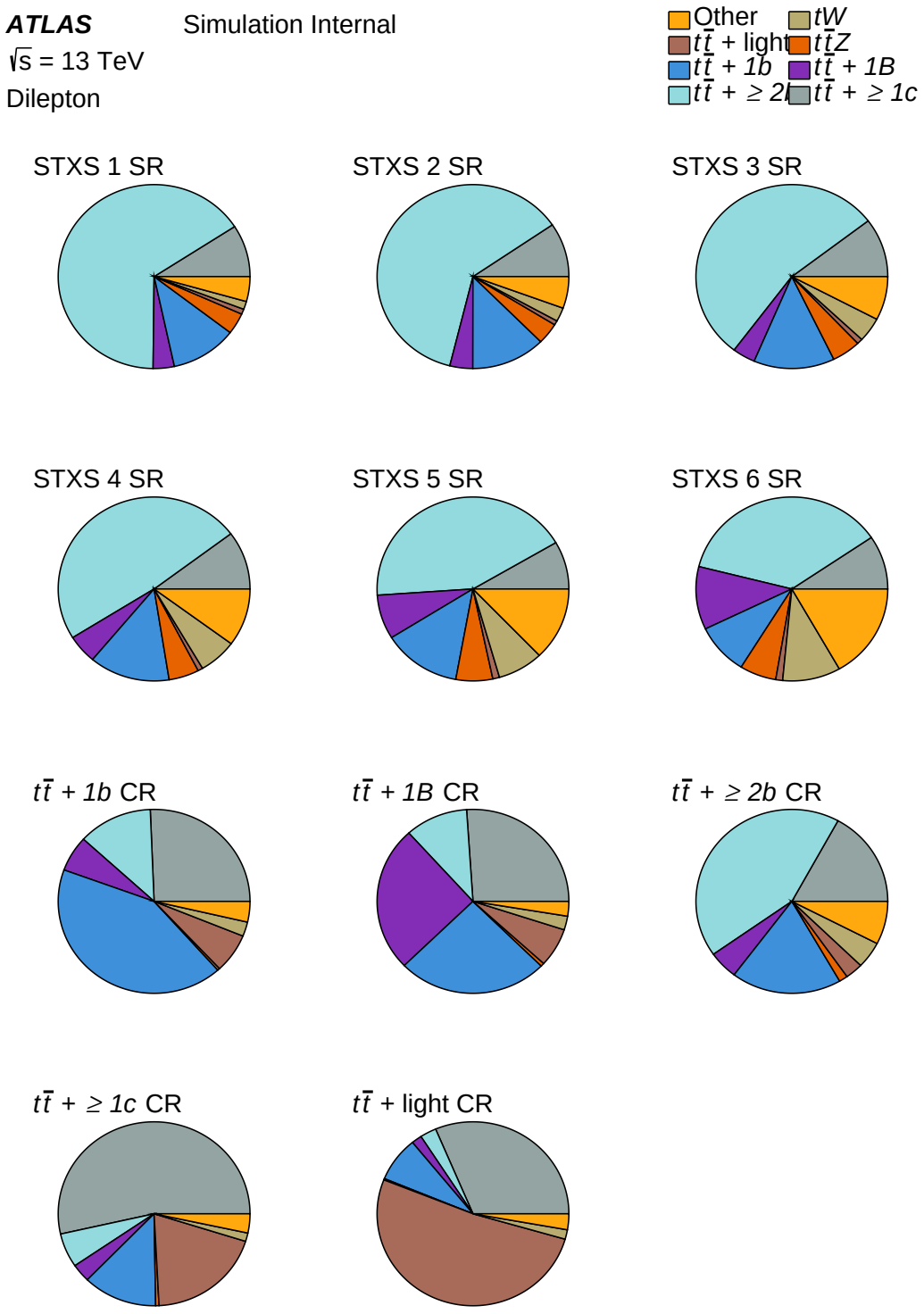
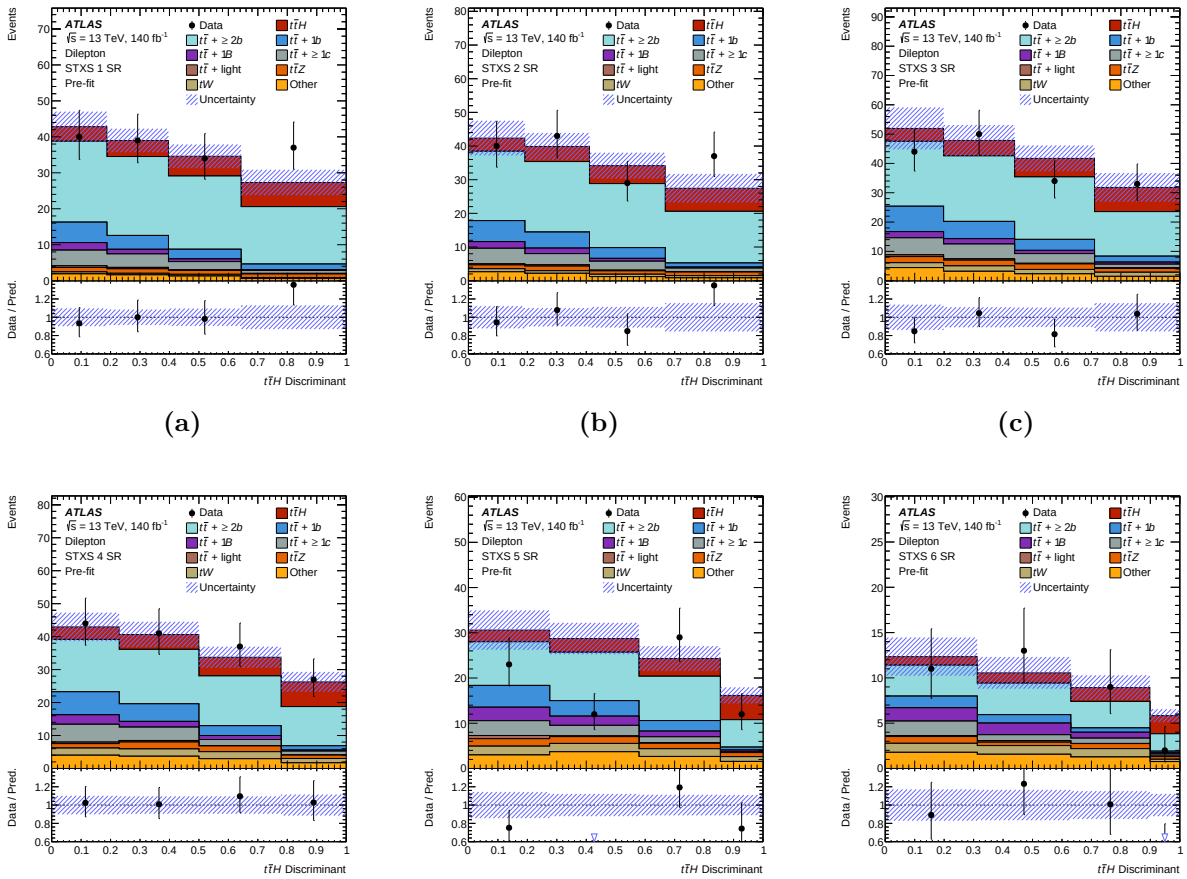


Figure A.1 Composition of the background processes in the signal and control regions of the di-lepton channel. The pie charts display the relative contributions of all processes, with one chart per analysis region. Each control region is enriched in the corresponding output class of the transformer network, while the signal regions are further split according to p_T^H.



(a) **(b)** **(c)**

(d) j **(e)** **(f)**

Figure A.2 Pre-fit distributions of the six signal regions in the dilepton channel: (a) STXS 1, (b) STXS 2, (c) STXS 3, (d) STXS 4, (e) STXS 5, (f) STXS 6. Each region is constructed to be enriched in the relevant flavour component of the additional jets in the dominant $t\bar{t}$+jets background. The uncertainty band includes both statistical and systematic contributions, excluding the free-floating normalisations of the individual $t\bar{t}$+jets flavour components. All discriminant scores are rescaled to the interval $[0, 1]$ using a logistic function.

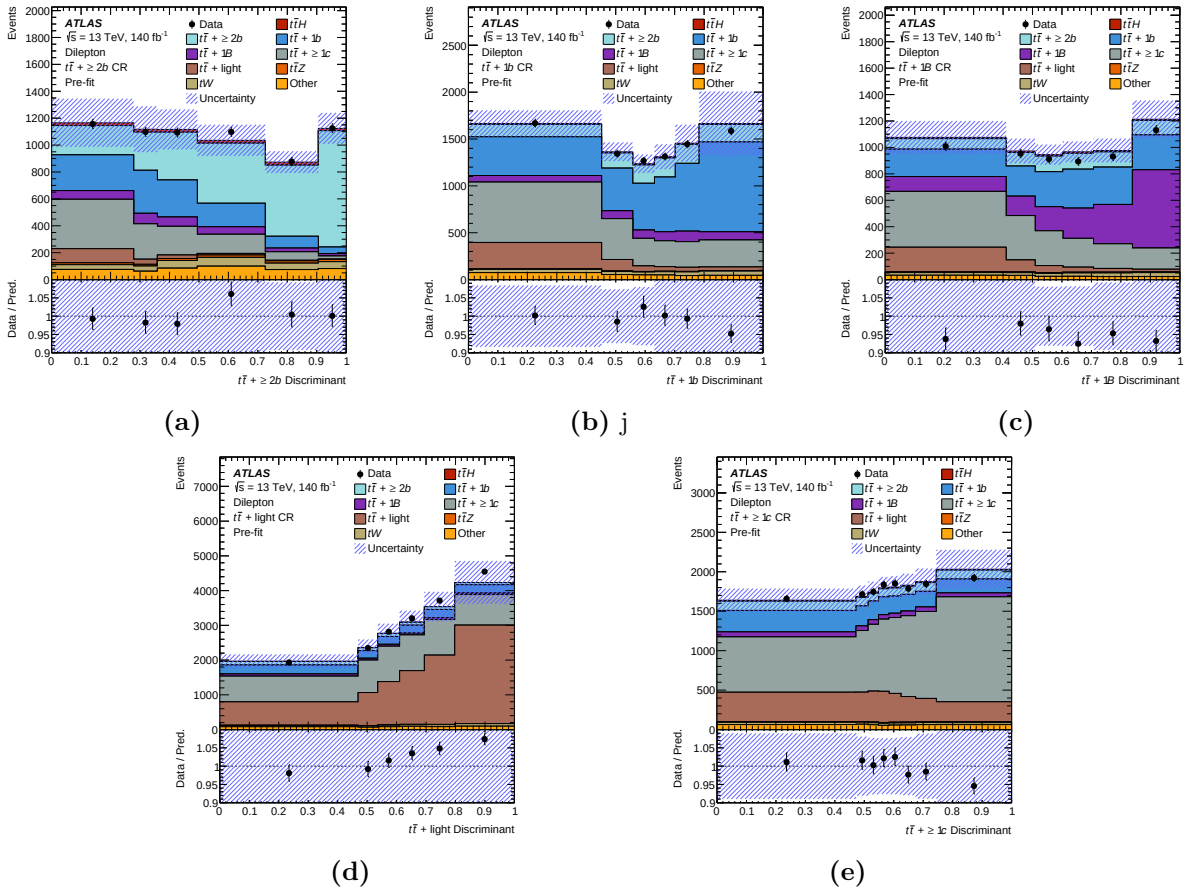


Figure A.3 Pre-fit distributions of the six control regions in the di-lepton channel: (a) $t\bar{t}+\geq 2b$, (b) $t\bar{t}+1b$, (c) $t\bar{t}+1B$, (d) $t\bar{t}$+light, (e) $t\bar{t}+\geq 1c$. Each region is constructed to be enriched in the corresponding flavour component of the additional jets in the dominant $t\bar{t}$+jets background. The uncertainty band includes both statistical and systematic contributions, excluding the free-floating normalisations of the individual $t\bar{t}$+jets flavour components. All discriminant scores are rescaled to the interval $[0, 1]$ using a logistic function.

A.2 Post-Fit Modelling

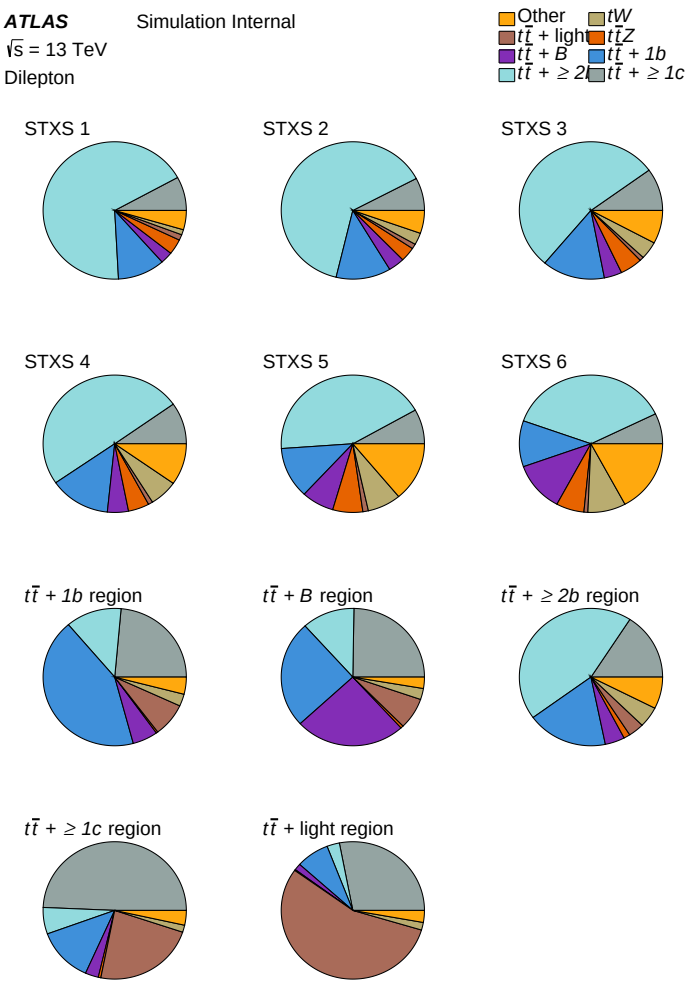
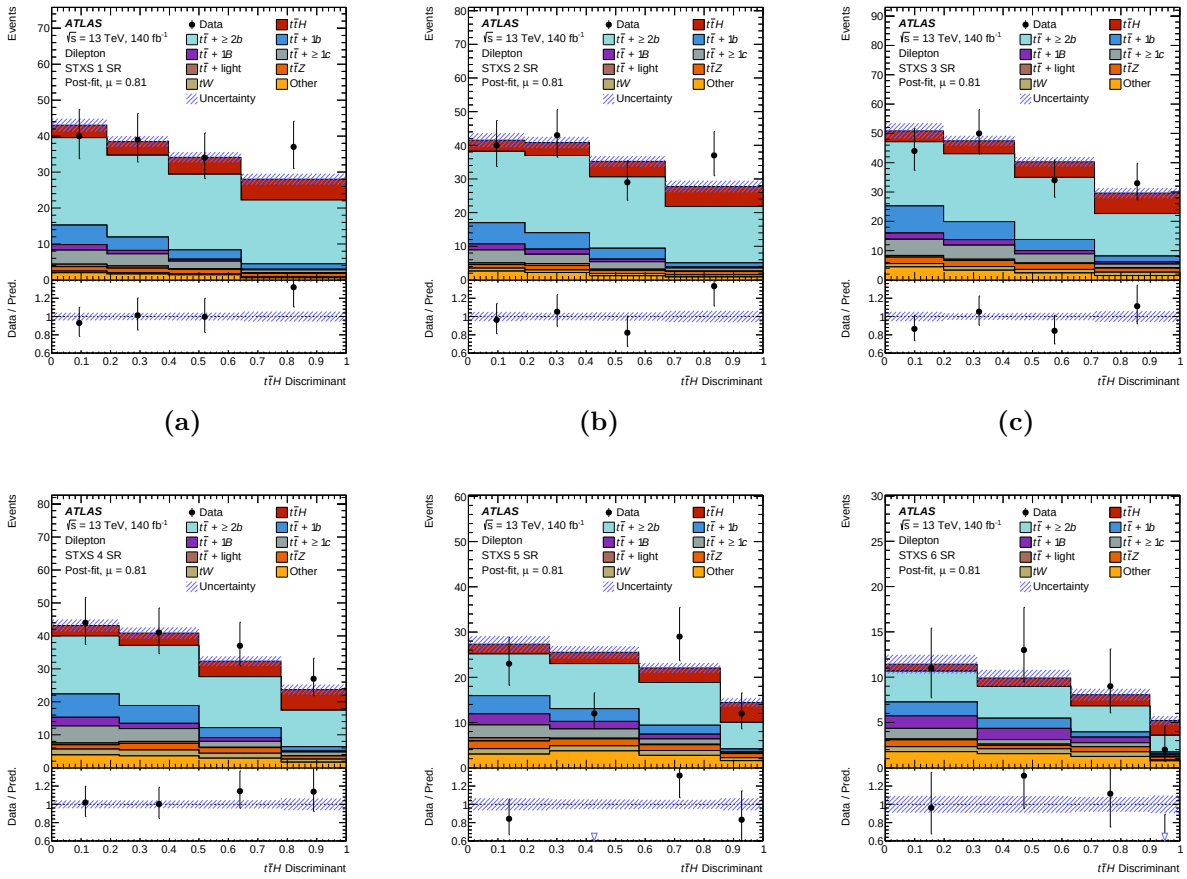


Figure A.4 Composition of the signal and control regions of the di-lepton channel. The pie charts indicate the relative contributions of all processes, with one chart per analysis region.



(a) **(b)** **(c)**

(d) j **(e)** **(f)**

Figure A.5 Post-fit distributions of the seven signal regions in the Di-lepton channel: (a) STXS 1, (b) STXS 2, (c) STXS 3, (d) STXS 4, (e) STXS 5, (f) STXS 6.

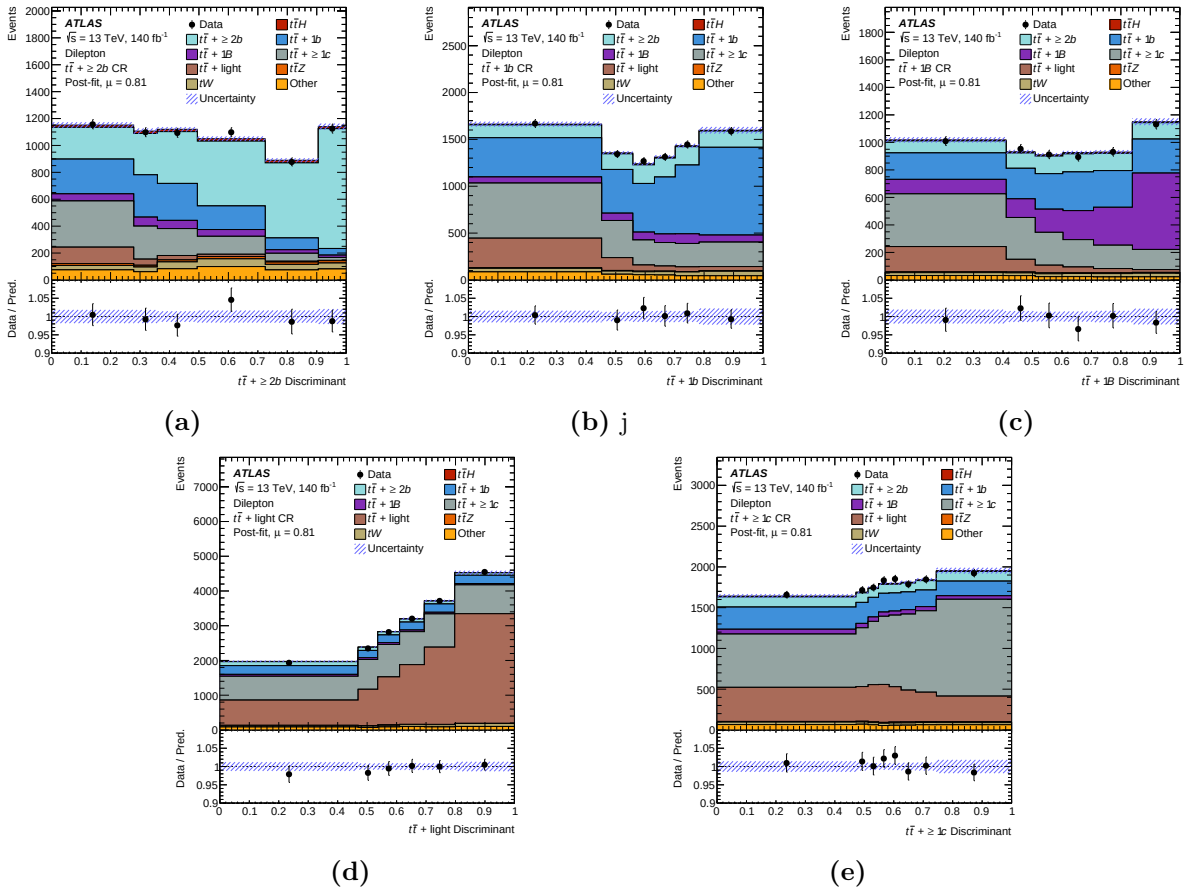


Figure A.6 Post-fit distributions of the six control regions in the di-lepton channel: (a) $t\bar{t}+\geq 2b$, (b) $t\bar{t}+1b$, (c) $t\bar{t}+1B$, (d) $t\bar{t}$+light, (e) $t\bar{t}+\geq 1c$.

A.3 STXS Ranking Plots

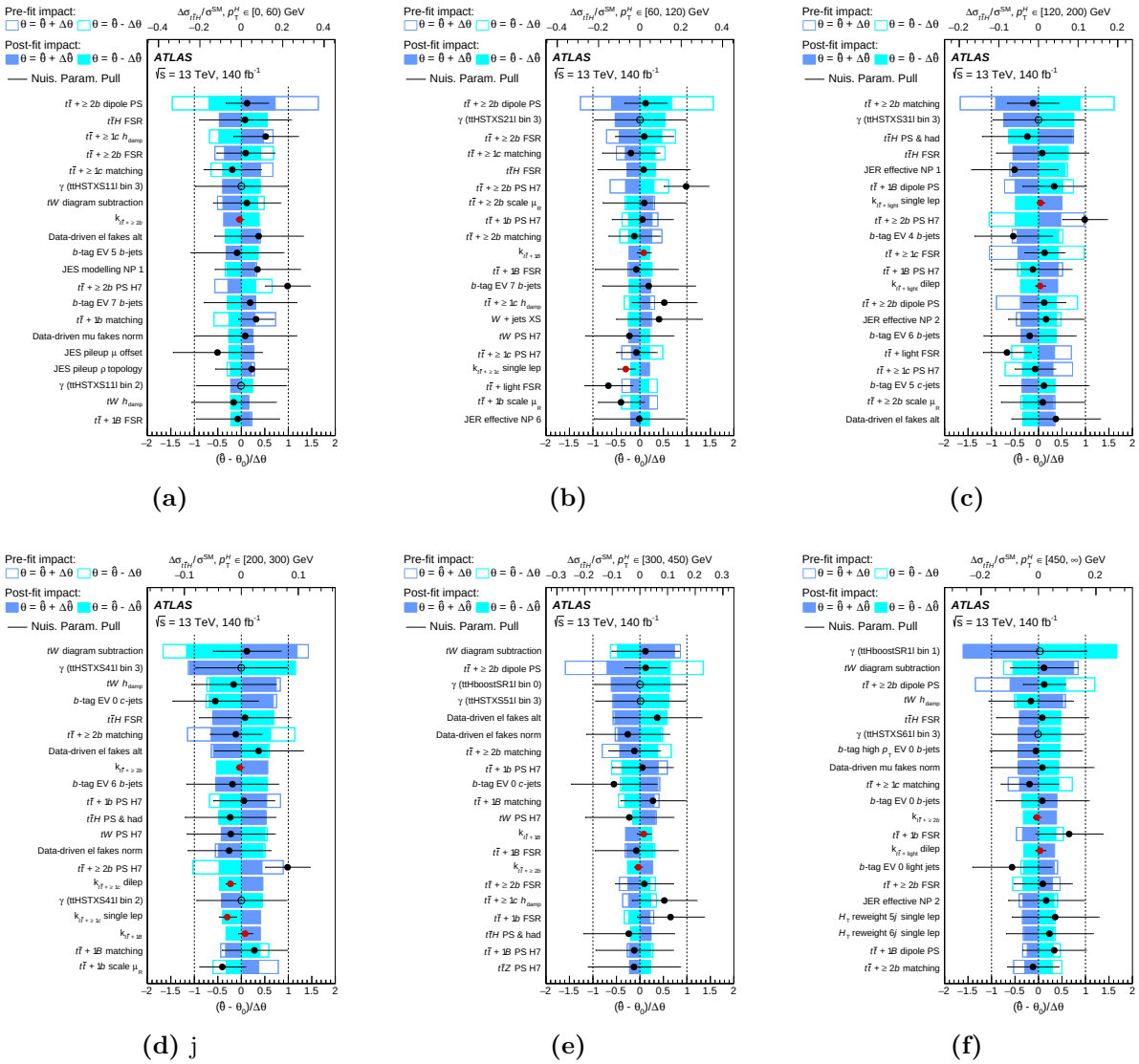


Figure A.7 Ranking of the 20 modelling and experimental systematic uncertainties with the largest post-fit impact in the inclusive STXS regions: (a) STXS 1, (b) STXS 2, (c) STXS 3, (d) STXS 4, (e) STXS 5, (f) STXS 6.

Bibliography

[1] Emmy Noether, *Invariant variation problems*,
Transport Theory and Statistical Physics **1** (1971) 186, ISSN: 1532-2424,
URL: http://dx.doi.org/10.1080/00411457108231446 (cit. on p. 4).

[2] Sheldon L Glashow, *The renormalizability of vector meson interactions*,
Nuclear Physics **10** (1959) 107 (cit. on p. 5).

[3] Abdus Salam and John Clive Ward, *Weak and electromagnetic interactions*,
Il Nuovo Cimento (1955-1965) **11** (1959) 568 (cit. on p. 5).

[4] Abdus Salam and John Clive Ward, *Weak and electromagnetic interactions*,
Il Nuovo Cimento (1955-1965) **11** (1959) 568 (cit. on p. 5).

[5] A Salam and JC Ward, *Electromagnetic and weak interactions*,
Physics Letters **13** (1964) 168 (cit. on p. 5).

[6] Clifford Cheung, *TASI Lectures on Scattering Amplitudes*, 2017,
arXiv: 1708.03872 [hep-ph], URL: https://arxiv.org/abs/1708.03872
(cit. on p. 5).

[7] F. Englert and R. Brout,
Broken Symmetry and the Mass of Gauge Vector Mesons,
Phys. Rev. Lett. **13** (9 1964) 321,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321
(cit. on p. 6).

[8] Peter W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*,
Phys. Rev. Lett. **13** (16 1964) 508,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508
(cit. on p. 6).

[9] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble,
Global Conservation Laws and Massless Particles,
Phys. Rev. Lett. **13** (20 1964) 585,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.585
(cit. on p. 6).

[10] S.M. Bilenky and J. Hošek, *Glashow-Weinberg-Salam theory of electroweak interactions and the neutral currents*, Physics Reports **90** (1982) 73 (cit. on p. 7).

[11] Hideki Yukawa, *On the Interaction of Elementary Particles. I*,
Proc. Phys. Math. Soc. Jap. **17** (1935) 48 (cit. on p. 8).

[12] Wikipedia contributors, *Standard Model of Elementary Particles*, 2025,
URL: https://en.wikipedia.org/wiki/File:
Standard_Model_of_Elementary_Particles.svg (cit. on p. 9).

[13] Particle Data Group et al., *Review of Particle Physics*,
Progress of Theoretical and Experimental Physics **2020** (2020) 083C01,
ISSN: 2050-3911, eprint: https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf,
URL: https://doi.org/10.1093/ptep/ptaa104 (cit. on p. 11).

[14] C T Potter et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*, en, 2013,
URL: http://cds.cern.ch/record/1559921 (cit. on p. 12).

[15] Nima Arkani-Hamed, Savas Dimopoulos, and Gia Dvali,
The Hierarchy problem and new dimensions at a millimeter,
Phys. Lett. B **429** (1998) 263, arXiv: hep-ph/9803315 [hep-ph] (cit. on p. 13).

[16] Morad Aaboud et al., *Search for the standard model Higgs boson produced in association with top quarks and decaying into abb pair in pp collisions at s= 13 TeV with the ATLAS detector*,
Physical Review D **97** (2018) 072016 (cit. on pp. 13, 14, 101, 103, 104, 106, 144).

[17] CMS Collaboration, *Observation of $t\bar{t}H$ Production*,
Phys. Rev. Lett. **120** (2018) 231801, arXiv: 1804.02610 [hep-ex]
(cit. on p. 13).

[18] CMS collaboration et al., *Search for ttH production in the H→bb decay channel with leptonic tt decays in proton-proton collisions at s= 13 TeV*,
Journal of High Energy Physics **2019** (2019) 26 (cit. on p. 14).

[19] ATLAS Collaboration,
Measurement of Higgs boson decay into b-quarks in associated production with a top-quark pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,
JHEP **06** (2022) 097 (cit. on p. 14).

[20] Howard Georgi, *An effective field theory for heavy quarks at low energies*,
Physics Letters B **240** (1990) 447 (cit. on p. 15).

[21] Gino Isidori, Felix Wilsch, and Daniel Wyler,
The Standard Model effective field theory at work, 2023,
arXiv: 2303.16922 [hep-ph], URL: https://arxiv.org/abs/2303.16922
(cit. on p. 15).

[22] The ATLAS Collaboration,
The ATLAS Experiment at the CERN Large Hadron Collider,
Journal of Instrumentation **3** (2008) S08003,
URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08003 (cit. on p. 18).

[23] The CMS Collaboration, *The CMS experiment at the CERN LHC*,
Journal of Instrumentation **3** (2008) S08004,
URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08004 (cit. on p. 18).

[24] The LHCb Collaboration, *The LHCb Detector at the LHC*,
Journal of Instrumentation **3** (2008) S08005,
URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08005 (cit. on p. 18).

[25] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*,
Journal of Instrumentation **3** (2008) S08002,
URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08002 (cit. on p. 18).

[26] E. Boltezar et al., *Performance of the New CERN 50 MeV Linac*,
IEEE Transactions on Nuclear Science **26** (1979) 3674 (cit. on p. 19).

[27] D J Warner, *Project study for a new 50 MeV linear accelerator for the C. P. S*,
tech. rep., CERN, 1973, URL: https://cds.cern.ch/record/414071
(cit. on p. 19).

[28] *The second stage CPS improvement study: 800 MeV booster synchrotron*,
Geneva: CERN, 1967, URL: https://cds.cern.ch/record/109292
(cit. on p. 19).

[29] *CERN Rapport annuel 1988*, tech. rep., CERN, 1989,
URL: https://cds.cern.ch/record/1516883 (cit. on p. 19).

[30] *CERN's 25 GeV proton synchroton*, Geneva: CERN, 1960, URL: https://cds.cern.ch/record/445034 (cit. on p. 19).

[31] *The 300 GeV programme*, French version available : Le programme 300 GeV, Geneva: CERN, 1972, URL: https://cds.cern.ch/record/104068 (cit. on p. 19).

[32] Julie Haffner, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, (2013), General Photo, URL: https://cds.cern.ch/record/1621894 (cit. on p. 19).

[33] L Arnaudon et al., *Linac4 Technical Design Report*, tech. rep., revised version submitted on 2006-12-14 09:00:40: CERN, 2006, URL: https://cds.cern.ch/record/1004186 (cit. on p. 20).

[34] Werner Herr and B Muratori, *Concept of luminosity*, (2006), URL: https://cds.cern.ch/record/941318 (cit. on p. 20).

[35] Joao Pequenao, "Computer generated image of the whole ATLAS detector", 2008, URL: https://cds.cern.ch/record/1095924 (cit. on p. 22).

[36] *ATLAS magnet system: Technical Design Report, 1*, Technical design report. ATLAS, Geneva: CERN, 1997, URL: https://cds.cern.ch/record/338080 (cit. on p. 23).

[37] A. Yamamoto et al., *The ATLAS central solenoid*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **584** (2008) 53 (cit. on p. 23).

[38] H.H. Ten Kate, *Superconducting magnet system for the ATLAS detector at CERN*, IEEE Transactions on Applied Superconductivity **9** (1999) 841 (cit. on p. 23).

[39] Goodson J, "Computer Generated image of the ATLAS magnet system", URL: http://www.jetgoodson.com/images/thesisImages/magnetSystems.png (cit. on p. 24).

[40] *ATLAS inner detector: Technical Design Report, 1*, Technical design report. ATLAS, Geneva: CERN, 1997, URL: https://cds.cern.ch/record/331063 (cit. on p. 24).

[41] Giulia Ripellino, *The alignment of the ATLAS Inner Detector in Run-2*, PoS **LHCP2016** (2016) 196 (cit. on p. 24).

[42] Norbert Wermes and G Hallewel, *ATLAS pixel detector: Technical Design Report*, Technical design report. ATLAS, Geneva: CERN, 1998, URL: https://cds.cern.ch/record/381263 (cit. on p. 24).

[43] G Aad et al., *ATLAS pixel detector electronics and sensors*, Journal of Instrumentation **3** (2008) P07007, URL: https://dx.doi.org/10.1088/1748-0221/3/07/P07007 (cit. on p. 24).

[44] M Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, tech. rep., 2010, URL: https://cds.cern.ch/record/1291633 (cit. on p. 24).

[45] A Miucci, *The ATLAS Insertable B-Layer project*, Journal of Instrumentation **9** (2014) C02018 (cit. on p. 24).

[46] B. Abbott et al., *Production and integration of the ATLAS Insertable B-Layer*, Journal of Instrumentation **13** (2018) T05008 (cit. on p. 24).

[47] The ATLAS collaboration, *Operation and performance of the ATLAS semiconductor tracker*, Journal of Instrumentation **9** (2014) P08009, ISSN: 1748-0221, URL: http://dx.doi.org/10.1088/1748-0221/9/08/P08009 (cit. on p. 25).

[48] V. A. MITSOU and, "THE ATLAS TRANSITION RADIATION TRACKER" 497 (cit. on p. 25).

[49] Adrian Vogel, *ATLAS Transition Radiation Tracker (TRT): Straw tube gaseous detectors at high rates*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **732** (2013) 277 (cit. on p. 25).

[50] Joao Pequenao, "Computer generated image of the ATLAS inner detector", 2008, URL: https://cds.cern.ch/record/1095926 (cit. on p. 26).

[51] *ATLAS calorimeter performance: Technical Design Report*, Technical design report. ATLAS, Geneva: CERN, 1996, URL: https://cds.cern.ch/record/331059 (cit. on p. 26).

[52] *ATLAS liquid-argon calorimeter: Technical Design Report*, Technical design report. ATLAS, Geneva: CERN, 1996, URL: https://cds.cern.ch/record/331061 (cit. on p. 27).

[53] *ATLAS tile calorimeter: Technical Design Report*, Technical design report. ATLAS, Geneva: CERN, 1996, URL: https://cds.cern.ch/record/331062 (cit. on p. 27).

[54] Joao Pequenao, "Computer Generated image of the ATLAS calorimeter", 2008, URL: https://cds.cern.ch/record/1095927 (cit. on p. 28).

[55] *ATLAS muon spectrometer: Technical Design Report*, Technical design report. ATLAS, Geneva: CERN, 1997, URL: https://cds.cern.ch/record/331068 (cit. on p. 28).

[56] Joao Pequenao, "Computer generated image of the ATLAS Muons subsystem", 2008, URL: https://cds.cern.ch/record/1095929 (cit. on p. 29).

[57] Michele Livan, *Monitored drift tubes in ATLAS*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **384** (1996) 214 (cit. on p. 29).

[58] T. Argyropoulos et al., "Cathode strip chambers in ATLAS : Installation, commissioning and in situ performance", *2008 IEEE Nuclear Science Symposium Conference Record*, 2008 2819 (cit. on p. 30).

[59] D. Boscherini, *Performance and operation of the ATLAS Resistive Plate Chamber system in LHC Run-1*, Journal of Instrumentation **9** (2014) C12039 (cit. on p. 30).

[60] Koichi Nagai, *Thin gap chambers in ATLAS*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **384** (1996) 219 (cit. on p. 30).

[61] G. Aad et al., *Performance of the ATLAS muon trigger in pp collisions at $$\sqrt{s}=8$$TeV*, The European Physical Journal C **75** (2015) 120 (cit. on p. 31).

[62] *ATLAS level-1 trigger: Technical Design Report*, Technical design report. ATLAS, Geneva: CERN, 1998, URL: https://cds.cern.ch/record/381429 (cit. on p. 31).

[63] Peter Jenni, Marzio Nessi, Markus Nordberg, and Kenway Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*, Technical design report. ATLAS, Geneva: CERN, 2003, URL: https://cds.cern.ch/record/616089 (cit. on p. 32).

[64] Atlas Collaboration et al., *Operation of the ATLAS trigger system in Run 2*, Journal of Instrumentation **15** (2020) P10004 (cit. on p. 32).

[65] Laura Gonella, *The ATLAS ITk detector system for the Phase-II LHC upgrade*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1045** (2023) 167597 (cit. on pp. 34, 35, 37).

[66] D Bortoletto, *ATLAS ITk tracking and readout performance*, Nucl. Instrum. Methods Phys. Res., A **1048** (2022) 167912, URL: https://cds.cern.ch/record/2883000 (cit. on p. 35).

[67] Andreas Gisen et al., *Planar n-in-n quad module prototypes for the ATLAS ITk upgrade at HL-LHC*, Journal of Instrumentation **12** (2017) C12032 (cit. on p. 36).

[68] Stefano Terzo, Sebastian Grinstein, Maria Manna, Giulio Pellegrini, and David Quirion, *A new generation of radiation hard 3D pixel sensors for the ATLAS upgrade*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **982** (2020) 164587 (cit. on p. 36).

[69] Howard Altman, *Captured Russian Weapons Are Packed With U.S. Microchips*, The War Zone (2022), URL: https://www.twz.com/captured-russian-weapons-are-packed-with-foreign-microchips (cit. on p. 37).

[70] Allison Wang, Rajeev Tomer, and Jack Tench, *InfluxDB: Real-time insights at any scale*, 2023, URL: https://www.influxdata.com/ (cit. on p. 39).

[71] Jacob VanderPlas et al., *Altair: Interactive statistical visualizations for python*, Journal of open source software **3** (2018) 1057 (cit. on p. 40).

[72] T. Cornelissen et al., "The new ATLAS track reconstruction (NEWT)", *Journal of Physics: Conference Series*, vol. 119, 3, 2008 032014, URL: https://dx.doi.org/10.1088/1742-6596/119/3/032014 (cit. on p. 44).

[73] The ATLAS Collaboration, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv: 1704.07983 [hep-ex] (cit. on p. 44).

[74] Azriel Rosenfeld and John L. Pfaltz, *Sequential Operations in Digital Picture Processing*, J. ACM **13** (1966) 471, ISSN: 0004-5411, URL: https://doi.org/10.1145/321356.321357 (cit. on p. 44).

[75] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987) 444, ISSN: 0168-9002, URL: https://www.sciencedirect.com/science/article/pii/0168900287908874 (cit. on p. 45).

[76] D Wicke, *A New Algorithm for Solving Tracking Ambiguities*, (1998), URL: https://cds.cern.ch/record/2625731 (cit. on p. 45).

[77] Heather M Gray, *The Charged Particle Multiplicity at Center of Mass Energies from 900 GeV to 7 TeV measured with the ATLAS Experiment at the Large Hadron Collider*, 2010 (cit. on p. 46).

[78] The ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, The European Physical Journal C **77** (2017) 1 (cit. on p. 46).

[79] The ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $$\sqrt{s} = 13$$ $$\text {TeV}$$*, The European Physical Journal C **79** (2019) 639 (cit. on pp. 47–49).

[80] The ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, The European Physical Journal C **77** (2017) 490 (cit. on p. 47).

[81] *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-022: CERN, 2017, URL: https://cds.cern.ch/record/2298955 (cit. on p. 48).

[82] J. Illingworth and J. Kittler, *A survey of the hough transform*, Computer Vision, Graphics, and Image Processing **44** (1988) 87 (cit. on p. 49).

[83] ATLAS Collaboration atlas. publications@ cern. ch et al., *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at s= 13 TeV*, The European Physical Journal C **81** (2021) 578 (cit. on pp. 50, 128).

[84] ATLAS Collaboration et al., *Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, arXiv preprint arXiv:2211.16178 (2022) (cit. on p. 50).

[85] Matteo Cacciari, Gavin P Salam, and Gregory Soyez, *The anti-kt jet clustering algorithm*, Journal of High Energy Physics **2008** (2008) 063 (cit. on pp. 51, 52).

[86] Gennaro Corcella et al., *HERWIG 6.5 release note*, arXiv preprint hep-ph/0210213 (2002) (cit. on pp. 52, 112).

[87] Morad Aaboud et al., *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex] (cit. on p. 52).

[88] *Jet reclustering and close-by effects in ATLAS run II*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2017-062: CERN, 2017, URL: https://cds.cern.ch/record/2275649 (cit. on p. 54).

[89] ATLAS Collaboration atlas. publications@ cern. ch et al., *Jet energy scale and resolution measured in proton–proton collisions at s= 13 TeV with the ATLAS detector*, The European Physical Journal C **81** (2021) 689 (cit. on pp. 55, 77, 129).

[90] *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-045: CERN, 2015, URL: https://cds.cern.ch/record/2064383 (cit. on p. 56).

[91] The ATLAS collaboration, *ATLAS b-jet identification performance and efficiency measurement with tt events in pp collisions at s= 13 TeV*, The European Physical Journal C **79** (2019) 970 (cit. on p. 56).

[92] Martino Tanasini, *New ATLAS b-tagging algorithm for Run 3*, PoS **ICHEP2022** (2022) 1090 (cit. on p. 57).

[93] *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-013: CERN, 2017,
URL: https://cds.cern.ch/record/2273281 (cit. on p. 58).

[94] *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, tech. rep.,
All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-003: CERN, 2017,
URL: https://cds.cern.ch/record/2255226 (cit. on p. 58).

[95] *Secondary vertex finding for jet flavour identification with the ATLAS detector*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-011: CERN, 2017,
URL: https://cds.cern.ch/record/2270366 (cit. on p. 58).

[96] *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*, tech. rep.,
All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2018-025: CERN, 2018,
URL: https://cds.cern.ch/record/2645405 (cit. on pp. 58, 110).

[97] Georges Aad et al.,
ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset,
The European Physical Journal C **83** (2023) 681 (cit. on pp. 59, 106).

[98] *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2016-012: CERN, 2016,
URL: https://cds.cern.ch/record/2160731 (cit. on pp. 59, 61).

[99] Georges Aad et al.,
ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset,
The European Physical Journal C **83** (2023) 681 (cit. on p. 59).

[100] *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s}$= 13 TeV*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-027: CERN, 2015, URL: https://cds.cern.ch/record/2037904 (cit. on pp. 63, 129).

[101] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: 0811.4622 [hep-ph] (cit. on p. 65).

[102] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics*, vol. 8, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, 1996, ISBN: 9780521545891 (cit. on p. 66).

[103] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438, [Yad. Fiz. 15 (1972) 781] (cit. on p. 67).

[104] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics*, Sov. Phys. JETP **46** (1977) 641, [Zh. Eksp. Teor. Fiz. 73 (1977) 1216] (cit. on p. 67).

[105] Guido Altarelli and Giorgio Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. B **126** (1977) 298, ISSN: 0550-3213 (cit. on p. 67).

[106] Leif Lönnblad, *Parton Showers*, Presentation at IPPP Durham Workshop, IPPP Durham, UK. PDF available online, 2005, URL: https://conference.ippp.dur.ac.uk/event/667/contributions/3846/attachments/3273/3566/lonnblad.pdf (cit. on pp. 68, 129).

[107] Torbjörn Sjöstrand, *Jet fragmentation of multiparton configurations in a string framework*, Nucl. Phys. B **248** (1984) 469 (cit. on p. 68).

[108] Gurpreet S. Chahal and Frank Krauss, *Cluster Hadronisation in Sherpa*, SciPost Phys. **13** (2022) 019 (cit. on p. 68).

[109] Yakov I. Azimov, Yuri L. Dokshitzer, Valery A. Khoze, and S. I. Troian, *The String Effect and QCD Coherence*, Phys. Lett. B **165** (1985) 147 (cit. on p. 69).

[110] Torbjörn Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph] (cit. on pp. 69, 112, 130).

[111] S. Agostinelli et al., *GEANT4 – a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 69).

[112] P. J. Clark for the ATLAS Collaboration, *The ATLAS Detector Simulation*, Nucl. Phys. B Proc. Suppl. **215** (2011) 85, Proceedings of the 12th Topical Seminar on Innovative Particle and Radiation Detectors (IPRD10), ISSN: 0920-5632, URL: http://www.sciencedirect.com/science/article/pii/S092056321100212X (cit. on p. 69).

[113] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C70 (2010) 823, CERN, 2010 (cit. on p. 70).

[114] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, tech. rep. ATL-PHYS-PUB-2010-013, CERN, 2010, URL: http://cds.cern.ch/record/1300517 (cit. on p. 70).

[115] ATLAS Collaboration and K. Edmonds et al., *The Fast ATLAS Track Simulation (FATRAS)*, tech. rep. ATL-SOFT-PUB-2008-001, CERN, 2008, URL: http://cds.cern.ch/record/1091969 (cit. on p. 70).

[116] ATLAS Collaboration, *AtlFast3: the next generation of fast simulation in ATLAS*, Computing and Software for Big Science **6** (2022) 7 (cit. on p. 70).

[117] *Simulation-based extrapolation of b-tagging calibrations towards high transverse momenta in the ATLAS experiment*, tech. rep., All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2021-003: CERN, 2021, URL: https://cds.cern.ch/record/2753444 (cit. on p. 73).

[118] Yan Ke, Ilaria Luise, Quentin Buat, and Giacinto Piacquadio, *Recent Developments on the Statistical Treatment of Flavour Tagging Uncertainties in ATLAS*, PoS **LHCP2022** (2023) 322 (cit. on p. 73).

[119] Ashish Vaswani et al., *Attention is all you need*, Advances in neural information processing systems **30** (2017) (cit. on p. 86).

[120] Wouter Verkerke and David P. Kirkby, *The RooFit toolkit for data modelling*, (2003), arXiv: physics/0306116 [physics.data-an], URL: https://arxiv.org/abs/physics/0306116 (cit. on pp. 91, 133).

[121] Lorenzo Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv: 1009.1003 [physics.data-an], URL: https://arxiv.org/abs/1009.1003 (cit. on pp. 91, 133).

[122] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, tech. rep., ATL-SOFT-PUB-2012-001, ATL-COM-SOFT-2012-030: New York University, 2012, URL: https://cds.cern.ch/record/1456844 (cit. on pp. 91, 133).

[123] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343 (cit. on pp. 91, 133).

[124] Rene Brun and Fons Rademakers, *ROOT – An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81, ISSN: 0168-9002 (cit. on pp. 91, 133).

[125] Simon J.D. Prince, *Understanding Deep Learning*, The MIT Press, 2023, URL: http://udlbook.com (cit. on pp. 92, 100).

[126] LHC Higgs Cross Section Working Group, ed., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, CERN Yellow Reports: Monographs, 869 pages, 295 figures, 248 tables and 1645 citations. Working Group web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG, Geneva: CERN, 2017, URL: https://cds.cern.ch/record/2227475 (cit. on p. 104).

[127] Nicolas Berger et al., *Simplified Template Cross Sections - Stage 1.1*, arXiv preprint (2019), arXiv: 1906.02754 [hep-ph], URL: https://arxiv.org/abs/1906.02754 (cit. on pp. 104, 126).

[128] Simone Alioli, Paolo Nason, Carlo Oleari, and Elia Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph] (cit. on p. 112).

[129] Richard D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on pp. 112, 114).

[130] Tomáš Jezo and Paolo Nason, *On the Treatment of Resonances in Next-to-Leading Order Calculations Matched to a Parton Shower*, JHEP **12** (2015) 065, arXiv: 1509.09071 [hep-ph] (cit. on p. 112).

[131] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph] (cit. on p. 112).

[132] ATLAS collaboration et al., *Measurement of the associated production of a top-antitop-quark pair and a Higgs boson decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, arXiv preprint arXiv:2407.10904 (2024) (cit. on p. 118).

[133] Albert Gyorgy Borbely, *Measurement of the associated production of a top quark pair and a Higgs boson (t tH) with boosted topologies*, PhD thesis: University of Glasgow, 2023 (cit. on p. 119).

[134] Atlas Collaboration et al., *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, arXiv preprint arXiv:2212.09379 (2022) (cit. on p. 128).

[135] Atlas Collaboration et al., *Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data*, arXiv preprint arXiv:1908.00005 (2019) (cit. on p. 128).

[136] John M. Campbell and R. Keith Ellis, *$tW^{\pm}$ production and decay at NLO*, JHEP **07** (2012) 052, arXiv: 1204.5678 [hep-ph] (cit. on p. 132).

[137] M. Aliev et al., *HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR*, Comput. Phys. Commun. **182** (2011) 1034, arXiv: 1007.1327 [hep-ph] (cit. on p. 132).

[138] P. Kant et al., *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, Comput. Phys. Commun. **191** (2015) 74, arXiv: 1406.4403 [hep-ph] (cit. on p. 132).

[139] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Uncertainties on α_s in global PDF analyses and implications for predicted hadronic cross sections*, Eur. Phys. J. C **64** (2009) 653, arXiv: 0905.3531 [hep-ph] (cit. on p. 132).

[140] Stefano Frixione, Eric Laenen, Patrick Motylinski, Chris White, and Bryan R. Webber, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv: 0805.3067 [hep-ph] (cit. on p. 132).

[141] ATLAS Collaboration, *Multi-boson simulation for 13 TeV ATLAS analyses*, tech. rep., CERN, 2016, URL: https://cds.cern.ch/record/2119986 (cit. on p. 132).

[142] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph] (cit. on p. 132).

[143] Fabio Maltoni, Eleni Vryonidou, and Cen Zhang, *Higgs production in association with a top-antitop pair in the Standard Model Effective Field Theory at NLO in QCD*, JHEP **10** (2016) 123, arXiv: 1607.05330 [hep-ph] (cit. on pp. 156, 157).

[144] G. Aad et al. ATLAS collaboration, *Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. D **90** (2014) 112015, eprint: 1408.7084 (cit. on p. 156).

[145] V. Khachatryan et al. CMS collaboration, *Observation of the diphoton decay of the Higgs boson and measurement of its properties*, Eur. Phys. J. C **74** (2014) 3076, eprint: 1407.0558 (cit. on p. 156).

[146] CMS Collaboration, *First results on Higgs to gammagamma at 13 TeV*, tech. rep. CMS-PAS-HIG-15-005, CERN, 2016 (cit. on p. 156).

[147] G. Aad et al. ATLAS collaboration, *Observation and measurement of Higgs boson decays to WW with the ATLAS detector*, Phys. Rev. D **92** (2015) 012006, eprint: 1412.2641 (cit. on p. 156).

[148] S. Chatrchyan et al. CMS collaboration, *Measurement of Higgs boson production and properties in the WW decay channel with leptonic final states*, JHEP **01** (2014) 096, eprint: 1312.1129 (cit. on p. 156).

[149] G. Aad et al. ATLAS collaboration, *Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. D **91** (2015) 012006, eprint: 1408.5191 (cit. on p. 156).

[150] S. Chatrchyan et al. CMS collaboration, *Measurement of the properties of a Higgs boson in the four-lepton final state*, Phys. Rev. D **89** (2014) 092007, eprint: 1312.5353 (cit. on p. 156).

[151] G. Aad et al. ATLAS collaboration, *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, JHEP **04** (2015) 117, eprint: 1501.04943 (cit. on p. 156).

[152] S. Chatrchyan et al. CMS collaboration, *Evidence for the 125 GeV Higgs boson decaying to a pair of leptons*, JHEP **05** (2014) 104, eprint: 1401.5041 (cit. on p. 156).

[153] V. Khachatryan et al. CMS collaboration, *Search for the associated production of the Higgs boson with a top-quark pair*, JHEP **09** (2014) 087, eprint: 1408.1682 (cit. on p. 156).

[154] G. Aad et al. ATLAS collaboration, *Search for H → 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*, Phys. Lett. B **740** (2015) 222, eprint: 1409.3122 (cit. on p. 156).

[155] G. Aad et al. ATLAS collaboration, *Search for the associated production of the Higgs boson with a top quark pair in multilepton final states with the ATLAS detector*, Phys. Lett. B **749** (2015) 519, eprint: 1506.05988 (cit. on p. 156).

[156] G. Aad et al. ATLAS collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into bb in pp collisions at s = 8 TeV with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 349, eprint: 1503.05066 (cit. on p. 156).

[157] CMS collaboration, *Search for ttH production in multilepton final states at sqrt(s) = 13 TeV*, tech. rep. CMS-PAS-HIG-15-008, CERN, 2016 (cit. on p. 156).

[158] CMS collaboration, *Search for ttH production in the H→ bb decay channel with $\sqrt{s}$ = 13 TeV pp collisions at the CMS experiment*, tech. rep. CMS-PAS-HIG-16-004, CERN, 2016 (cit. on p. 156).

[159] The ATLAS collaboration,
Interpretations of the ATLAS measurements of Higgs boson production and decay rates and differential cross-sections in pp collisions at $\sqrt{s}$= 13 TeV,
Journal of High Energy Physics **2024** (2024) 97 (cit. on pp. 158, 160).

[160] Andy Buckley, Chris White, and Martin White, *Practical Collider Physics*,
Institute of Physics Publishing, 2022, ISBN: 978-0-7503-2442-7 (cit. on p. 163).

List of Figures

List of Tables