
**$t\bar{t}H$ Production Measurement
in the $H \rightarrow b\bar{b}$ Decay Channel with ATLAS at 13 TeV
and System Tests for the Pixel Detector**

Dissertation

for the award of the degree
“Doctor rerum naturalium” (Dr. rer. nat.)
of the Georg-August-Universität Göttingen

within the Physics doctoral programme
of the Georg-August University School of Science (GAUSS)

submitted by

Christian Scheulen

from Essen

Göttingen, 2024

Thesis Committee:

Prof. Dr. Arnulf Quadt

II. Institute of Physics, Georg-August-Universität Göttingen

Prof. Dr. Stanley Lai

II. Institute of Physics, Georg-August-Universität Göttingen

Prof. Dr. Steffen Schumann

Institute for Theoretical Physics, Georg-August-Universität Göttingen

Members of the Examination Board:

Reviewer:

Prof. Dr. Arnulf Quadt

II. Institute of Physics, Georg-August-Universität Göttingen

Second Reviewer:

Jun.-Prof. Dr. María Moreno Llácer

Institut de Física Corpuscular (IFIC),
joint centre between CSIC and Universitat de València

Additional Reviewer:

Priv.-Doz. Dr. Judith Katzy

Deutsches Elektronen-Synchrotron,
associated with Humboldt-Universität zu Berlin

Further members of the Examination Board:

Prof. Dr. Stanley Lai

II. Institute of Physics, Georg-August-Universität Göttingen

Prof. Dr. Steffen Schumann

Institute for Theoretical Physics, Georg-August-Universität Göttingen

Prof. Dr. Ansgar Reiners

Institute for Astrophysics, Georg-August-Universität Göttingen

Jun.-Prof. Dr. Thibaud Humair

Deutsches Elektronen-Synchrotron, associated with Georg-August-Universität Göttingen

Date of the oral examination: 6 November 2024

Reference: II.Physik-UniGö-Diss-2024/08

$t\bar{t}H$ Production Measurement in the $H \rightarrow b\bar{b}$ Decay Channel with ATLAS at 13 TeV and System Tests for the Pixel Detector

Abstract

The Yukawa coupling between the Higgs boson and the top quark – the heaviest elementary particle in the Standard Model of particle physics – is one of the fundamental parameters of the theory. With a value close to unity, it may play a unique role in electroweak symmetry breaking. Additionally, it is central to the shape of the Higgs potential and possible models of vacuum decay, while at the same time being highly sensitive to physics processes beyond the Standard Model. The top Yukawa coupling is only directly accessible in processes involving the associated production of a Higgs boson with one or more top quarks.

This thesis presents an analysis of the associated production of a top quark pair with a Higgs boson in the bottom–anti-bottom Higgs boson decay channel and the leptonic top quark decay modes. The analysis uses the full Run 2 dataset of 140 fb^{-1} collected between 2015 and 2018 at a centre-of-mass energy of 13 TeV with the ATLAS detector at the Large Hadron Collider. Since a previous analysis using the same dataset has been published already, a focus is placed on improvements included in the legacy re-analysis presented in this thesis. Herein, special attention is given to the development of permutation-invariant Transformer neural networks to suppress dominant background processes and reconstruct the Higgs boson from its decay products for a differential cross-section measurement conducted as part of the analysis. The observed significance against a background-only hypothesis is 4.6σ compared to an expected significance of 5.4σ . The measured inclusive cross-section of $411^{+101}_{-92} \text{ fb}$ for a Higgs boson mass of 125.09 GeV is consistent with the Standard Model prediction and represents the most precise single-channel measurement of the signal process to date.

The identification of hadronic jets originating from bottom quarks is of central importance to the analysis presented in this thesis since four such jets are expected in the leading-order final state of the signal process. The Pixel detector situated at the core of the ATLAS detector is crucial for identifying the characteristic displaced secondary vertices of these jets. Therefore, a second focus of this thesis lies in the development of automated system tests for the Pixel detector to ensure continued good detector performance for future physics analyses using Run 3 data.

MARCUS BRODY:

[It] is a source of unspeakable power, and it
has to be researched!

MAJ. EATON:

And it will be, I assure you, Dr. Brody, [*He
turns to Indy*] Dr. Jones. We have top men
working on it right now.

HENRY 'INDIANA' JONES:

Who?

MAJ. EATON:

Top. Men.

(from: *Raiders of the Lost Ark*, 1981, Paramount Pictures)

Table of Contents

I. The Standard Model — and how to test it at the LHC	1
1. The Standard Model of Particle Physics	3
1.1. The Standard Model	4
1.2. The Interplay of Top and Higgs	15
2. The CERN Accelerator Complex and the ATLAS Experiment	29
2.1. The Large Hadron Collider	30
2.2. The ATLAS Detector	34
II. Data Acquisition with the ATLAS Pixel Detector	45
3. The Pixel Detector Read-out Chain	47
3.1. Components of the Read-out Chain	47
3.2. Calibrating the Pixel Detector	50
3.3. Pixel Detector Operations in Run 3	53
4. Automated Testing Framework Development	57
4.1. Common Testing Infrastructure	58
4.2. Calibration Tests	59
4.3. Data-Taking Tests	61
III. The $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 Legacy Analysis	63
5. Machine Learning Fundamentals	65
5.1. The Multilayer Perceptron	65

Table of Contents

5.2. Learning via Backpropagation and Gradient Descent	67
5.3. The Transformer Architecture	71
6. The $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 Legacy Analysis	77
6.1. Comparison with First Full Run 2 Analysis Setup	77
6.2. Data and Simulation Samples	79
6.3. Object & Event Selection	87
6.4. Profile Likelihood Fits	91
6.5. Analysis Strategy	94
7. Multivariate Analysis Techniques for $t\bar{t}(H \rightarrow b\bar{b})$	101
7.1. Network Architecture	101
7.2. Network Training	107
7.3. Transformer Performance	110
7.4. Model Comparison	114
8. Results	117
8.1. Systematic Uncertainties	117
8.2. Pre-Fit Modelling	119
8.3. $t\bar{t}H$ Measurement Results	124
Summary and Conclusions	137
Bibliography	139
Appendices	165
A. Additional MVA Material	167
A.1. Network Training	167
A.2. Transformer Performance	174
A.3. Additional Layer Schematics	176
B. Terms & Acronyms	179
C. Declaration on the Usage of Generative AI	191

Acknowledgements

While one severely over-cafeinated person stands as the author on every submitted thesis, it takes a village to complete a PhD – thanks to a whole network of people, the same person is barely hanging on to sanity. In the following, I want to thank the group of people who have helped me hold on to mine most of the time.

I would like to begin by thanking Arnulf Quadt for making my journey towards a PhD possible. You took me into your working group after an inspiring summer internship in 2018 and pushed me to pursue a diverse range of topics during the past four years. I am incredibly grateful for the things I learned and the opportunities I was afforded as a result of this. I would also like to thank María Moreno Llácer for agreeing to be the second reviewer of this thesis after having already followed the progress of the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis as HTop convener. Equally, I would like to thank Judith Katzy for agreeing to act as an additional reviewer of this thesis. Additionally, I am thankful to Stan Lai and Steffen Schumann as the other members of my thesis advisory committee and for agreeing to serve on my defence committee. Similarly, I am thankful to Ansgar Reiners and Thibaud Humair for agreeing to be part of said committee. I also want to thank the German Academic Scholarship Foundation (*Studienstiftung des Deutschen Volkes*) for supporting me with a scholarship during the majority of my doctoral studies.

I am very grateful to Elizaveta Shabalina for my day-to-day supervision within the $t\bar{t}(H \rightarrow b\bar{b})$ analysis. You were always there for discussions and suggestions, your feedback was invaluable, and I am incredibly grateful for your no-nonsense approach. Similarly, I would like to thank Marcello Bindi and Tobias Bisanz for their day-to-day supervision of my qualification task and for always taking the time to answer my numerous questions about the Pixel detector and its read-out system. Extra special thanks to Liza are also due for initiating my next chapter past the PhD.

I want to say thanks to the whole $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis team for the last four years of working together towards a common goal. Thank you all for the invaluable discussions in the course of the analysis, and for answering any questions I had during the write-up of this thesis. I am looking forward to what HTop still has in

Acknowledgements

store for Run 3. At the same time, I am immensely grateful to the Pixel Ops team for letting me anywhere near real-life hardware during my stay at CERN. You showed me another thrilling aspect of experimental physics when making yet another set of slides seemed unbearable, and I will never look at a green wire in the same way again... Also, I am looking forward to finally joining a Pixel BBQ with somebody else on call.

Major thanks are due to everybody who proofread parts of this dissertation to disentangle my logical leaps and hunt down every single last typoe¹. This includes, in alphabetical order like all of the best papers: Marcello Bindi, Stephen Eggebrecht, Alex Froch, Jörn Große-Knetter, Steffen Korn, Baptiste Ravina, Elizaveta Shabalina, and Yusong Tian. I also want to thank Marike Schwickardi for explaining how to use the GWDG printers while I was in a near-catatonic state following the electronic submission.

I want to thank the members of the II. Institute of Physics for providing a sometimes much-needed distraction from work. At this point, a special shout-out to Ishan is due for getting me out of the house in early 2021 after not having met another living soul except for my ex-flatmates for half a year. No thanks are due to them, and I wish never to have to move to a new city in the middle of a pandemic again. However, I am grateful to Alexej, Michi, and Sophie for making our flat in Göttingen feel like a home. Another special shout-out is due to coffee, without which this thesis could never have been finished on time, and to Jörn, who was conscientiously brewing it break after break.

I would especially like to thank Yusong. You accompanied me for a part of our journey over these last years, for which I am incredibly thankful.

Last but not least, I want to give a huge thank you to my family for their unwavering support during – but also already long before – I embarked on these last four years. Ich drücke Euch alle!

¹ Except in this part, of course.

Contributions by the Author

In the same way that no PhD thesis is written in isolation, every PhD project in modern science is an inherently collaborative effort. This goes doubly for experimental particle physics, where the names of collaborations stand as the authors of papers instead of individual people. Therefore, a non-exhaustive summary of the author's contributions is provided in the following. At the same time, the chapters of this thesis will endeavour to give a complete account of the collaborative work undertaken.

The main topic of this thesis concerns the ATLAS $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis published as Reference [1]. The results are the work of a large analysis team spread over many institutes. The author was responsible for the development and debugging of the analysis software framework used to process Monte-Carlo simulations and data samples. He also performed the central sample production and carried out sample validations. While J. Clavijo carried out the original Transformer development in the analysis, the author performed two retraining campaigns, implemented improvements in network architecture and training procedures, and debugged and validated the models. The results of these efforts are contained in Chapter 7, where a comparison to an alternative model developed by N. Kumari is also included. Additionally, the author conducted studies of the analysis strategy in the single-lepton boosted channel, implemented muon-in-jet corrections, validated samples for the tW background, and rederived the dilepton fake background estimate.

Other contributions to the analysis include, but are not limited to, the following: comparisons of data and Monte-Carlo samples, the systematics model development, and a majority of the fitting studies in the early analysis stages were performed by L. Ferencz, and by L. Evans in the later stages of the analysis, while many other members of the team conducted additional fit studies. The fake background estimation in the single-lepton channel was performed by K. Solovieva and A. Froch, who also completed the initial H_T reweighting efforts of J. Clavijo. A. Borbély performed a retraining on the deep neural network used in the boosted single-lepton channel, while E. Maričić carried out studies for a dilepton boosted region which was not

adopted for the final analysis. Y. Tian developed $t\bar{t} + b\bar{b}$ b -tagging Monte-Carlo scale factor corrections, and Z. Rozario developed high- p_T b -tagging extrapolations. L. Carvalho derived the W + jets background estimates. A. Held developed the reco bin edge corrections and carried out the compatibility studies with the previous analysis round.

The qualification task of the author to become a full-fledged member of the ATLAS collaboration consisted of the development of automated system tests for the Pixel detector calibration read-out chain, detailed in Section 4.2 of this thesis. A. Ghosal performed the corresponding development of data-taking tests as a complementary qualification task. Both tasks were seeded by previous preliminary developments by the Pixel operations team. The final qualification task reports are contained in References [2, 3]. The author took over the further development of the overall test development after the conclusion of his qualification task, and refactored large parts of the data-taking tests to be more modular before integrating them into the continuous integration system. Chapter 4 gives a complete account of the current state of this ongoing project. Meanwhile, Chapter 3 provides context on the Pixel read-out developments during Run 3 and the Pixel detector operations for the same period, which the author contributed to as an on-call expert during his stay at CERN.

To convey the work undertaken as part of this thesis, it is structured into several parts as follows. Part I contains background information on Standard Model physics in Chapter 1, and introduces the LHC accelerator complex and the ATLAS detector in Chapter 2. Part II consists of an introduction to the data acquisition system of the Pixel detector and Run 3 operations in Chapter 3 followed by a description of the system tests developed as part of this thesis in Chapter 4. Part III outlines the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis, starting with a short introduction of relevant machine learning concepts in Chapter 5, followed by a description of the samples, object and event selection, and general analysis strategy in Chapter 6. Chapter 7 presents studies on the multivariate analysis techniques employed in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis. Subsequently, Chapter 8 outlines the systematics model of the analysis, followed by a presentation and comparison of its results. The thesis concludes with a [summary](#) of the work performed.

Part I.

The Standard Model — and how to test it at the LHC

The Standard Model of Particle Physics

To explain the phenomena observed in experimental particle physics experiments, the Standard Model [4–13] was formulated in the 1960s and 1970s as the most complete framework describing our universe to date. It encompasses three of the four fundamental forces governing nature with unrivalled precision using the common language of renormalisable quantum field theories.

Of course, unlike some of its virtual particles, the Standard Model did not emerge from the vacuum¹. Instead, it builds on the foundations laid by such earlier models as covariant Quantum Electrodynamics (QED) [14–22], the Yukawa interaction [23], and Fermi theory [24], which were formulated to describe the electromagnetic force, the strong force between hadrons, and the weak force observed in radioactive β -decays, respectively. In turn, these are based on the paradigm-shifting formulation of quantum mechanics [25–30] and special relativity [31, 32] – the list could go on. On the experimental side, these models were met by discoveries such as those of particle spin [33] and of anti-particles [34], but also by the development of experimental techniques such as the cloud chamber [35] and the cyclotron [36], which enabled particle physics experiments in the first place. In general, the history of modern particle physics – and that of science overall – is one of the continuous experimental verification and falsification of theoretical models in the hope of slowly converging on the true nature of reality. The work carried out for this thesis in one corner of the Standard Model will hopefully do its small part in this ongoing story.

In the following, the Standard Model will be delineated in Section 1.1 with a short outline of its shortcomings and possible solutions involving physics beyond the Standard Model. Afterwards, Section 1.2 will take a closer look at the physics of the top quark, the Higgs boson, and their interaction with each other.

¹ It has also stood the test of time (and of many experiments) better than the average virtual particle.

1.1. The Standard Model

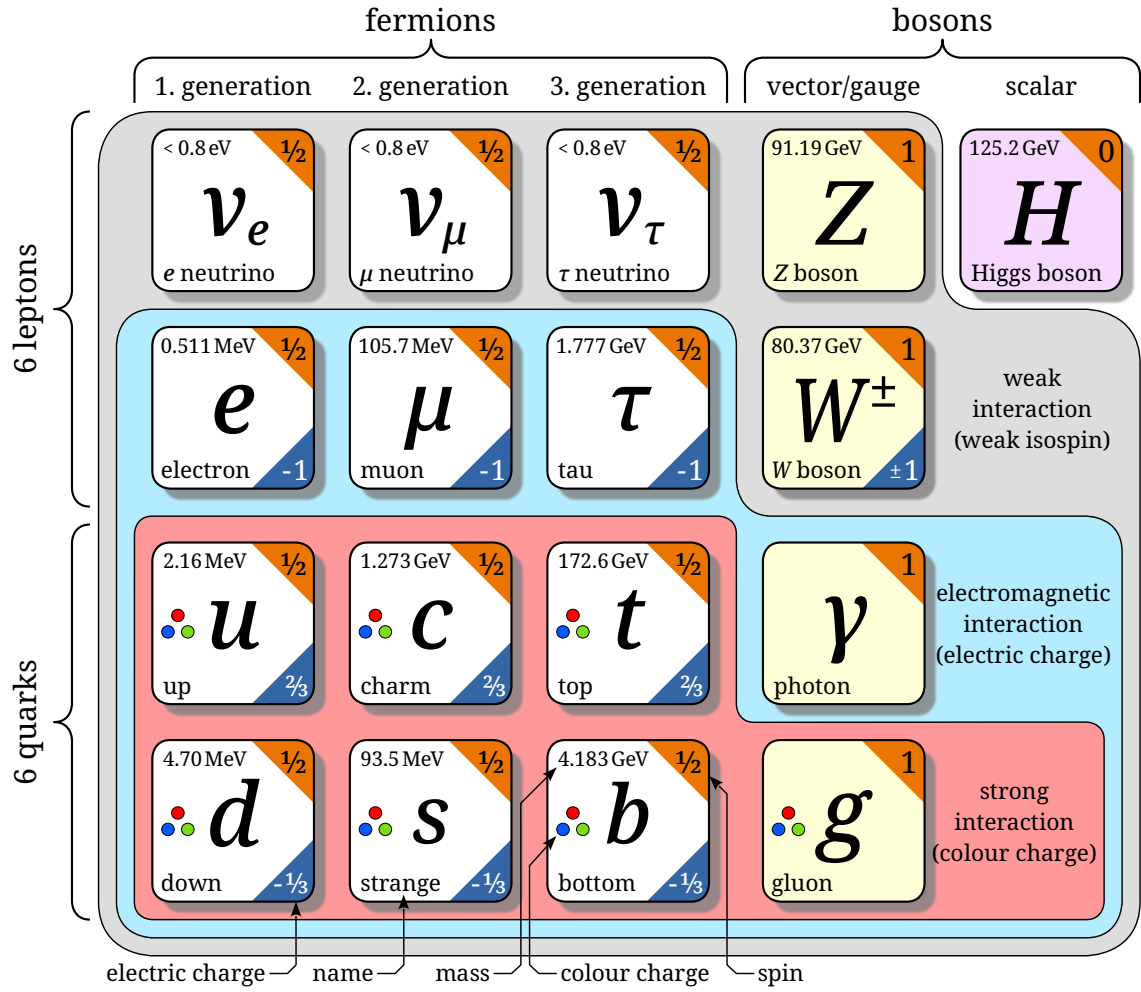


Figure 1.1: Graphic of the fundamental Standard Model particles with their properties. The particles can be divided into fermions, vector bosons mediating the three fundamental forces included in the Standard Model, and the scalar Higgs boson. Coloured frames relate force mediators to the respective fermions with non-zero couplings. The couplings between bosons are not shown. The mass values are taken from Reference [37].

The fundamental particles postulated by the Standard Model (SM) organise themselves into two principal groups: the twelve spin- $\frac{1}{2}$ fermions² obeying Fermi-Dirac statistics, and the integer-spin bosons obeying Bose-Einstein statistics. While the fermions make up all visible matter in the universe, spin-1 gauge bosons mediate the interactions between them. The electromagnetic force is mediated via massless photons, the strong force described by Quantum Chromodynamics (QCD) via

² Additionally, each fermion has an anti-fermion ‘twin’ with some of its quantum numbers, such as the electric charge, inverted. For brevity, *fermion* is used to refer to both fermions and anti-fermions, unless explicitly stated.

massless gluons, and the weak force via massive electrically charged $W^\pm$ bosons and neutral Z bosons. The electromagnetic and the weak force can be combined to form a single Quantum Field Theory (QFT) via *electroweak unification* [4–6] as described in Section 1.1.2. Lastly, the SM contains one additional particle: the massive scalar³ Higgs boson linked to *spontaneous symmetry breaking* of the electroweak gauge symmetry. The Higgs field gives rise to the particle masses via the Brout-Englert-Higgs mechanism [38–43] outlined in Section 1.1.3. A summary of all fundamental SM particles and their properties is provided in Figure 1.1.

Which fermions couple to which force mediator is determined by a separate quantum number – the charge of the interaction – for each of the fundamental forces. Six of the fermions carry a colour charge – the quantum number linked to QCD – and thus couple to the strong force, while the other six do not. The former group is therefore referred to as quarks, while the latter is called leptons. Meanwhile, the coupling to the electromagnetic force is determined by the electric charge Q . Finally, the charge associated with the weak force is the *weak isospin* T , which behaves differently depending on the *chirality* of the fermions. As a result, only the left-handed fermions and right-handed anti-fermions, organised in weak isospin doublets, interact with the W bosons⁴. The right-handed fermions and left-handed anti-fermions instead form weak isospin singlets, and thus do not couple to the W bosons. Additionally, the fermions can be classified into three *generations*. Corresponding particles in different generations have identical quantum numbers but increase in mass with increasing generation. The quark pair and the lepton pair in each generation form the weak isospin doublets.

As a QFT, the SM treats particles as excitations of fundamental quantum fields described in a Lagrangian formalism. For free fermions without particle interactions, the Lagrangian based on the Dirac equation has the form

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \quad (1.1)$$

using the Einstein summation convention and the four-component spinor fields ψ , the fermion mass m , the Dirac matrices⁵ γ^μ , and the adjoint spinor fields $\bar{\psi} = \psi^\dagger \gamma^0$.

Continuous symmetries in the theory give rise to conserved currents in accordance with Noether’s theorem [44], highlighting the special role of symmetries in theoretical physics. The formulation of the SM ensures invariance of the Lagrangian under gauge transformations of the form

$$\psi \rightarrow \psi' = \exp(i\theta_a G_a)\psi, \quad (1.2)$$

where ψ is a fermion field, and $\exp(i\theta_a G_a)$ with implicit summation over a represents the gauge transformation using the generators G_a of the Lie group associated

³ Meaning it has a spin of 0 and transforms symmetrically under parity transforms.

⁴ The story gets more complicated for Z bosons due to electroweak unification.

⁵ The Dirac matrices satisfy $\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}$ for the Minkowski metric $g = \text{diag}(1, -1, -1, -1)$.

with the gauge symmetry under consideration. For local gauge transformations, coordinate-dependent phase factors $\theta_a(x)$ are used. In this case, ∂_μ in Equation (1.1) is replaced by the *covariant derivative*,

$$D_\mu = \partial_\mu + iG_a A_{a,\mu},$$

for invariance under local gauge transformations. $A_{a,\mu}$ are the gauge fields, which themselves transform as $A_{a,\mu} \rightarrow A'_{a,\mu} = A_{a,\mu} - \partial_\mu \theta_a$ under the gauge transformation.

Overall, the SM Lagrangian is based on the gauge symmetry

$$\text{SU}(3)_C \times \text{SU}(2)_L \times \text{U}(1)_Y,$$

where $\text{SU}(n)$ refer to special unitary and $\text{U}(n)$ to unitary groups of degree n . Additionally, the subscript L signifies that the gauge group only acts on the left-handed fermion doublets, while C and Y refer to the quantum numbers of the colour charge and the weak hypercharge⁶, respectively. Thus, $\text{SU}(3)_C$ refers to the non-Abelian gauge group with eight generators underlying QCD. Meanwhile, $\text{SU}(2)_L \times \text{U}(1)_Y$ represents the combination of a non-Abelian group with three generators and an Abelian group with a single generator underlying the electroweak interaction.

1.1.1. Quantum Chromodynamics

When a veritable *subnuclear zoo*⁷ of hadrons surfaced in the bubble chamber experiments of the late 1940s and the 1950s, a taxonomy of all these new particles was desired. First attempts at this had already been made via the *strong isospin* [45–47] to explain the similarities between proton and neutron interactions. With the addition of *strangeness* [48–51], the approximate $\text{SU}(3)$ symmetry of the *eightfold way* [52, 53] for light hadrons emerged. Later supported by Deep Inelastic Scattering (DIS) experiments [54, 55], the quark model [56, 57] was proposed as a more fundamental theory in 1964. QCD was then formulated as the QFT underlying the strong interaction with three quantum number charges, commonly referred to by analogy to the colours *red*, *green*, and *blue* [58–61].

Lone quarks and gluons are not found experimentally due to *colour confinement* [60–62]: overall, only colourless objects can exist so that quarks form into gluon-mediated baryons (qqq states where quarks of different colours make an overall ‘white’ hadron), mesons ($q\bar{q}$ states, where the colours of the quark and the anti-colour of the anti-quark cancel out), and more exotic states such as tetra- and pentaquarks. Because of colour confinement and hadronisation, quarks generated in high-energy particle collisions form particle showers, the so-called *jets*. The existence of gluons was finally confirmed via planar three-jet events at PETRA in 1979 [63–66].

⁶ $Y = 2(Q - T_3)$ with electric charge Q and third weak isospin component T_3 .

⁷ The term was coined by Oppenheimer in 1956.

The QCD Lagrangian without the free-fermion terms of Equation (1.1) has the structure

$$\mathcal{L}_{\text{QCD}} = -\bar{\psi}_i \left(\gamma^\mu g_s \frac{1}{2} \lambda_{a,ij} G_{a,\mu} \right) \psi_j - \frac{1}{4} \mathcal{G}_{a,\mu\nu} \mathcal{G}_a^{\mu\nu}, \quad (1.3)$$

where implicit summation over the four space-time dimensions for μ and ν , the three colour charges for i and j , and the eight gluon fields $G_a(x)$ is performed. Meanwhile, the fermion types are neglected in Equation (1.3). The QCD coupling strength is given by $g_s = \sqrt{4\pi\alpha_s}$ and λ_a are the eight Gell-Mann matrices corresponding to the SU(3) generators up to a factor of two. Using the SU(3) structure functions⁸ f_{abc} , the gluon field strength tensor $\mathcal{G}_{a,\mu\nu}$ expands as

$$\mathcal{G}_{a,\mu\nu} = \partial_\mu G_{a,\nu} - \partial_\nu G_{a,\mu} + g_s f_{abc} G_{b,\mu} G_{c,\nu}. \quad (1.4)$$

QCD is a Yang-Mills gauge theory [67]. As it is non-Abelian, the generators do not commute with each other, resulting in non-vanishing structure functions and self-interacting gluons, which carry colour charges. With increasing energy, particle self-interactions give rise to poles. Fortunately, these infinities can be removed by *renormalisation* [7–9, 68] to some characteristic energy called the *renormalisation scale* μ_R close to the energy ranges of the processes under investigation. This procedure effectively shifts the energy dependence into the QCD coupling constant α_s to create a *running coupling* $\alpha_s = \alpha_s(\mu_R^2)$. Overall, the strong coupling becomes weaker with increasing interaction energy in a process called *asymptotic freedom* [12, 13, 62]. As a result, perturbative QCD calculations with fermion–gluon, and three- and four-boson gluon–gluon vertices are possible at high enough energy scales.

The recasting of protons as composite particles made up of *partons* [70, 71] – the quarks and gluons – has implications for hadron collider experiments: instead of the protons colliding at a centre-of-mass (CoM) energy $\sqrt{s}$, two individual partons carrying proton momentum fractions⁹ x_1 and x_2 form the initial state of the perturbative partonic cross-section calculations. Meanwhile, the remaining constituents of the protons act as spectator particles to the process of interest. However, cross-sections for the full hadron collisions are needed to compare the theory to experimental data. Fortunately, the partons can be treated as collinearly moving quasi-free particles at high enough collision energies because of asymptotic freedom. As a result, it is possible to disentangle the high-energy interaction and the parton dynamics in the bound hadron states using the *factorisation theorem* [73–76]. The differential cross-section for a proton–proton collision into the final state X then turns into

$$\sigma_{pp \rightarrow X} = \sum_{a,b} \int dx_1 dx_2 f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2) \cdot \sigma_{ab \rightarrow X}(\hat{s}, \mu_R^2, \mu_F^2),$$

where a and b sum over all partons, i.e. the quarks, anti-quarks, and gluons. The

⁸ Defined via the group generator commutation relations: $[\lambda_a, \lambda_b] = 2i \sum_c f_{abc} \lambda_c$.

⁹ Also called the Bjorken x from its role in Bjorken scaling [72].

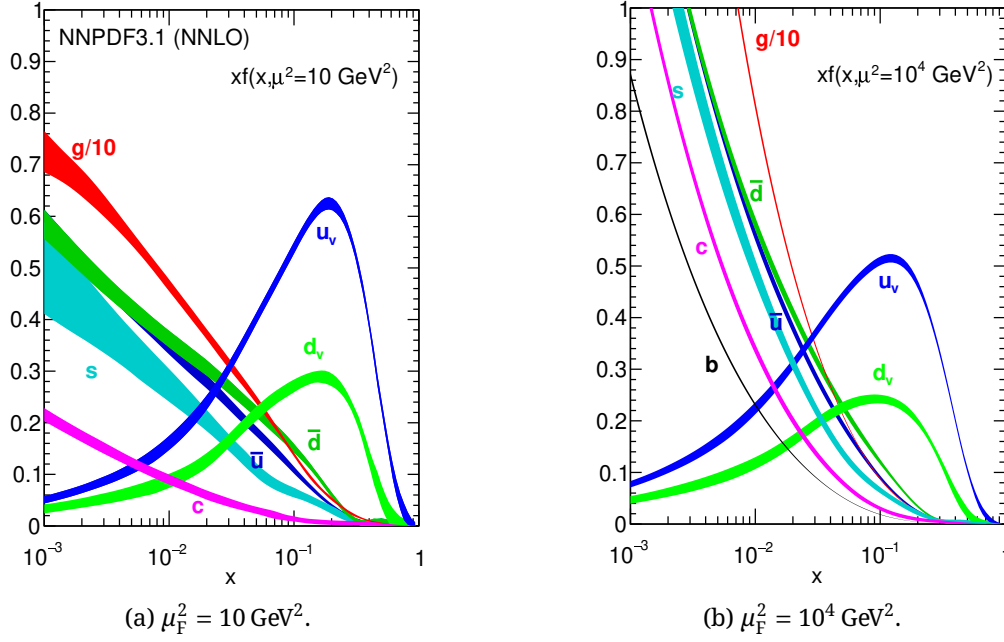


Figure 1.2: Semi-logarithmic plots of the NNPDF3.1 proton Parton Distribution Functions at two different factorisation scales μ_F^2 . The difference between the u and d valence quarks, and the remaining partons is clearly visible. Additionally, the momentum fraction x carried by sea quarks and gluons increases with increasing μ_F . The plots are taken from Reference [69].

Parton Distribution Functions (PDFs) $f_c(x, \mu_F^2)$ give the probability of finding a parton of type c in the momentum fraction range between x and $x + dx$. The PDFs depend on the *factorisation scale* μ_F marking the transition between perturbative high-energy QCD and non-perturbative QCD parametrised by the PDFs. The partonic cross-section $\sigma_{ab \rightarrow X}(\hat{s}, \mu_R^2, \mu_F^2)$ taking place at an energy $\sqrt{\hat{s}} = \sqrt{x_1 x_2 s}$ can be calculated perturbatively at the renormalisation scale μ_R . Meanwhile, the PDFs lie in the non-perturbative QCD regime and have to be determined experimentally. With the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations [77–79], which describe the scale dependence of the PDFs, the distributions can be translated to different energy scales once they are known at one scale.

In addition to the three so-called *valence quarks* expected from the uud quark content of the proton, a cloud of virtual *sea quarks* and gluons is also present and has to be considered in the PDFs. Because of the equivalence of the PDFs and the structure functions for DIS [75, 76], data from DIS experiments formed a crucial ingredient for the first PDFs. Today, global PDF fits for the Monte-Carlo (MC) generators used to simulate hadron collision processes include DIS results from HERA experiments, data from fixed-target Drell-Yan [80] experiments, and the collider experiments at the TEVATRON and LHC under utilisation of the DGLAP equations [69]. Plots of the NNPDF3.1 PDFs [69] produced by the NNPDF collaboration are shown as an example in Figure 1.2.

1.1.2. Electroweak Unification

Fermi theory [24], which explains radioactive β -decay via four-fermion interactions, is not renormalisable like QED and QCD. Originally suggested as a solution to the θ - τ puzzle [81], the observation of parity violation in the weak interaction underlying β -decay in 1957 [82] additionally fuelled the search for a better description of the interaction. Parity violation is allowed in interactions with a $V - A$ (vector – axial) structure, which might be mediated by electrically charged heavy gauge bosons. Therefore, it was hypothesised that the weak charged currents might behave in such a way [83, 84]. As a result, Fermi theory would represent an Effective Field Theory (EFT) only valid at low energy scales. With the development of spontaneous symmetry breaking, electroweak unification [4–6] could then be formulated as a renormalisable gauge invariant theory with an $SU(2)_L \times U(1)_Y$ symmetry.

In the parity-violating $V - A$ structure, the fermion spinors are split into left-handed and right-handed components¹⁰. The left-handed fermions (and similarly the right-handed anti-fermions) form weak isospin doublets ρ_L as

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, \quad \begin{pmatrix} c \\ s \end{pmatrix}_L, \quad \begin{pmatrix} t \\ b \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L,$$

where the quark spinors are represented by the letters used in Figure 1.1, while ℓ refers to the charged lepton spinors with their respective neutrino spinors ν_ℓ . Similar to the spinors, $\bar{\rho}_L = \rho_L^\dagger \gamma^0$.

Using the weak isospin doublets ρ_L and singlets ψ_R , the electroweak Lagrangian before symmetry breaking consists of (without the free-fermion terms):

$$\begin{aligned} \mathcal{L}_{EW} = & -\bar{\rho}_L \left(\gamma^\mu g_W \frac{1}{2} \sigma_a W_{a,\mu} + \gamma^\mu g' \frac{Y}{2} B_\mu \right) \rho_L - \bar{\psi}_R \left(\gamma^\mu g' \frac{Y}{2} B_\mu \right) \psi_R \\ & - \frac{1}{4} \mathcal{W}_{a,\mu\nu} \mathcal{W}_a^{\mu\nu} - \frac{1}{4} \mathcal{B}_{\mu\nu} \mathcal{B}^{\mu\nu}, \end{aligned} \quad (1.5)$$

with the couplings g_W and g' for the $SU(2)_L$ and $U(1)_Y$ gauge groups, respectively. The summation is analogous to Equation (1.3) except that a here indexes the three fields $W_a(x)$ of $SU(2)$, which has the Pauli matrices σ_a as generators. The field strength tensors can be defined analogously to Equation (1.4). However, the field B of the Abelian $U(1)$ group commutes with itself, resulting in no self-interaction. Before symmetry breaking, the W_a gauge fields do not have mass terms yet¹¹ and therefore cannot correspond to the physical weak bosons.

In the weak isospin doublets of Equation (1.5), the fermion fields are given in a weak eigenstate basis, which is not identical to that of the mass eigenstates. As a result,

¹⁰ This can be accomplished via the projection operators as $\psi_L = \frac{1}{2}(1 - \gamma^5)\psi$ and $\psi_R = \frac{1}{2}(1 + \gamma^5)\psi$ where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$.

¹¹ Manually added mass terms would not be invariant under gauge transforms.

the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix $\mathbf{V}_{\text{CKM}}$ [85, 86] was introduced to express the down-type weak eigenstates (dashed) in terms of the mass eigenstates (undashed):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \overbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}^{\mathbf{V}_{\text{CKM}}} \begin{pmatrix} d \\ s \\ b \end{pmatrix},$$

$$\mathbf{V}_{\text{CKM}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (1.6)$$

where the parametrisation in Equation (1.6) is in terms of the sines s_{ij} and cosines c_{ij} of three mixing angles θ_{ij} with $0 < \theta_{ij} < \frac{\pi}{2}$ and (experimentally) $s_{13} < s_{23} < s_{12} < 1$. As a result, diagonal entries $|V_{ii}| \sim 1$, while off-diagonal elements – especially V_{td} and V_{ub} – are Cabibbo-suppressed. The CKM matrix also provides an explanation for the CP violation observed in neutral kaon decay experiments in 1964 [87] via the CP -violating phase δ chosen to only appear in conjunction with small matrix elements in the parametrisation [88]. The prediction of a third quark generation required for CP violation in the CKM matrix was confirmed through the discovery of the b -quark via the Y -meson in 1977 [89], while the tau lepton had already been discovered in 1975 [90] as the first third-generation fermion to be found experimentally. Because of the quark mixing, the weak charged current is the only possibility for a quark to decay into another lighter quark in the SM.

1.1.3. Electroweak Symmetry Breaking and the Brout-Englert-Higgs Mechanism

The electroweak gauge fields do not correspond to the force mediators encountered at hadron colliders. Instead, the physical bosons are the result of spontaneous symmetry breaking caused by the scalar Higgs field. In the SM, the Higgs model consists of two complex scalar fields structured in an additional weak isospin doublet,

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

As for the other isospin doublets, the upper and lower components differ by one unit of electric charge. The Lagrangian for these fields (without any couplings to the electroweak sector) is

$$\mathcal{L}_H = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - \overbrace{\left(\mu^2 \phi^\dagger \phi + \lambda^2 (\phi^\dagger \phi)^2 \right)}^{V(\phi)}, \quad (1.7)$$

where $\lambda^2 > 0$ to create a minimum in the Higgs potential $V(\phi)$. The potential has global minima at $\phi^\dagger \phi = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda^2}$ for $\mu^2 < 0$ leading to spontaneous symmetry breaking. The constant v is referred to as the vacuum expectation value, since the fields can be expanded about the minimum such that only the neutral scalar field ϕ_3 has a non-zero vacuum expectation value giving rise to the physical scalar boson field h as $\phi_3(x) = v + h(x)$. Meanwhile, the other three degrees of freedom turn into massless *Goldstone bosons* [91, 92] during the breaking of the symmetry. A schematic for an unbroken and a broken symmetry with a single complex field is shown in Figure 1.3.

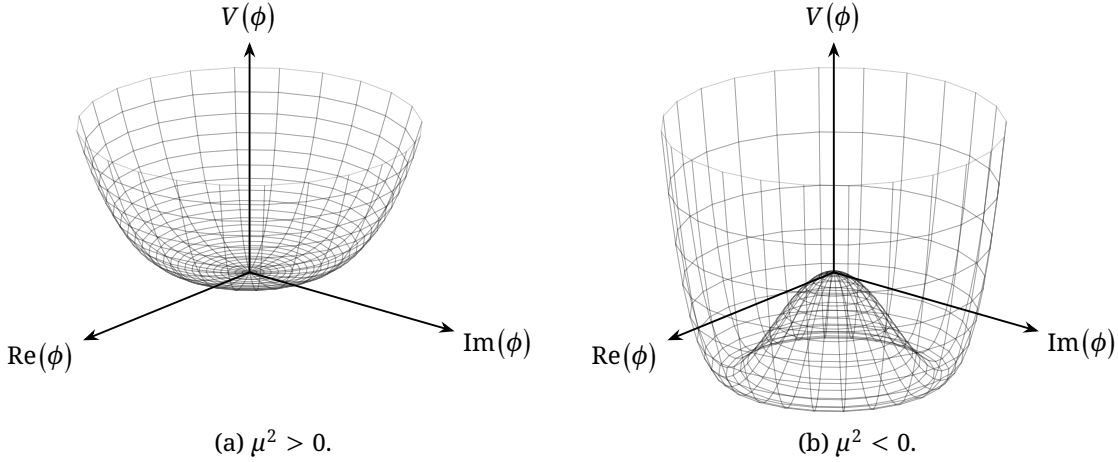


Figure 1.3: Plots of the Higgs potential $V(\phi)$ from Equation (1.7) simplified to a single complex field for (a) $\mu^2 > 0$ and (b) $\mu^2 < 0$. In the latter case, spontaneous symmetry breaking occurs due to the ring of global minima at non-zero field values in the ‘Mexican hat’ potential.

During Electroweak Spontaneous Symmetry Breaking (EWSB), the Goldstone bosons turn into the longitudinal spin states of the massive weak gauge bosons [38–43, 93] when transforming into the so-called *unitary gauge* [94], where $\phi = \begin{pmatrix} 0 \\ v+h \end{pmatrix}$. When subjecting the Higgs Lagrangian of Equation (1.7) to the EW gauge symmetry via a covariant derivative, the following Lagrangian (up to a constant factor) emerges after gauging away the Goldstone bosons:

$$\begin{aligned}
 \mathcal{L}_{\text{EW}+H} = & \underbrace{(\partial_\mu h)(\partial^\mu h) - \lambda^2 v^2 h^2}_{\text{massive scalar field } h} - \underbrace{(\lambda^2 v h^3 + \frac{\lambda^2}{4} h^4)}_{h \text{ self-interactions}} + \underbrace{\frac{1}{4} g_W^2 v^2 W_\mu^- W^{+\mu} + \frac{1}{8} g_Z^2 v^2 Z_\mu Z^\mu}_{\text{weak boson masses}} \\
 & + \underbrace{\frac{1}{2} g_W^2 v h W_\mu^- W^{+\mu} + \frac{1}{4} g_W^2 h^2 W_\mu^- W^{+\mu}}_{h-W \text{ couplings}} + \underbrace{\frac{1}{4} g_Z^2 v h Z_\mu Z^\mu + \frac{1}{8} g_Z^2 h^2 Z_\mu Z^\mu}_{h-Z \text{ couplings}} \\
 & - \underbrace{\left(\frac{1}{4} \mathcal{W}_{a,\mu\nu} \mathcal{W}_a^{\mu\nu} + \frac{1}{4} \mathcal{B}_{\mu\nu} \mathcal{B}^{\mu\nu} \right)}_{\text{EW field strength tensors}}. \tag{1.8}
 \end{aligned}$$

As a result of EWSB, the physical Higgs boson obtains a mass $m_H = \sqrt{2}\lambda v$, while the physical weak gauge bosons obtain the masses $m_W = \frac{1}{2}g_W v$ and $m_Z = \frac{1}{2}g_Z v$ (with $g_Z^2 = g_W^2 + g'^2$), which are linear in the vacuum expectation value v . The W bosons additionally obtain unit electric charges from the charged complex Higgs field.

The physical gauge boson fields in Equation (1.8) are generated from linear combinations of the massless EW fields as

$$W^\pm = \frac{1}{\sqrt{2}}(W_1 \mp iW_2), \quad A = \cos \theta_W B + \sin \theta_W W_3, \quad Z = -\sin \theta_W B + \cos \theta_W W_3,$$

where θ_W is the Weinberg angle, which satisfies $\tan \theta_W = \frac{g'}{g_W}$. As a result, the gauge boson masses and couplings are related as $m_W = m_Z \cos \theta_W$ and $g_W = g_Z \cos \theta_W$. Because of the mixing of the photon field A and the Z boson field Z , the gauge boson interactions arising from the field strength tensors are modified as well. Consequently, the photon couples to the weak bosons unlike the unbroken B field. Additionally, triple and quartic couplings of the Higgs field to the vector boson fields and to itself can be identified.

The couplings of the fermions are affected as well by EWSB:

$$\begin{aligned} \mathcal{L}_{\text{EW}+f} = & \underbrace{\bar{\psi}_f(i\gamma^\mu \partial_\mu)\psi_f - \frac{y_f}{\sqrt{2}}v\bar{\psi}_f\psi_f}_{\text{massive fermions}} - \underbrace{\frac{y_f}{\sqrt{2}}h\bar{\psi}_f\psi_f}_{\text{Yukawa coupling}} - \underbrace{\bar{\psi}_f(\gamma^\mu g_Z \frac{1}{2}(c_{V,f} - c_{A,f}\gamma^5)Z_\mu)\psi_f}_{\text{weak neutral current}} \\ & - \underbrace{\bar{\rho}_{q'}\left(\gamma^\mu \frac{g_W}{2\sqrt{2}}(1 - \gamma^5)(\sigma_+ W_\mu^+ V_{q'q} + \sigma_- W_\mu^- V_{qq'}^*)\right)\rho_q}_{\text{weak charged current with quarks}} \\ & - \underbrace{\bar{\rho}_\ell\left(\gamma^\mu \frac{g_W}{2\sqrt{2}}(1 - \gamma^5)(\sigma_+ W_\mu^+ + \sigma_- W_\mu^-)\right)\rho_\ell}_{\text{weak charged current with leptons}} - \underbrace{\bar{\psi}_f(\gamma^\mu e Q_f A_\mu)\psi_f}_{\text{electromagnetic interaction}}, \end{aligned}$$

where the Lagrangian contains the fermion kinematic terms and their interactions with the Higgs and broken EW fields. The spinor fields are in the mass eigenstates, so that the CKM matrix elements $V_{qq'}$ are included in the quark weak charged current. Implicit summation over all fermions f , and three generations of quarks q and leptons ℓ is performed depending on the subscript. Similar to the weak bosons, the fermions now obtain their masses from the Brout-Englert-Higgs mechanism as

$$m_f = \frac{1}{\sqrt{2}}y_f v, \quad (1.9)$$

where y_f is the Yukawa coupling of the respective fermion. Unlike the manually added mass terms $m\bar{\psi}\psi$ of the free fermion in Equation (1.1), the masses originate from gauge-invariant Yukawa interaction terms with a pre-EWSB structure¹²

¹² Technically, the conjugate Higgs doublet $\phi_c = -i\sigma_2\phi^*$ transforming exactly like ϕ under $\text{SU}(2)$ is

$$\mathcal{L}_{\text{Yukawa}} = -y_f (\bar{\rho}_L \phi \psi_R + \bar{\psi}_R \phi^\dagger \rho_L). \quad (1.10)$$

The $V - A$ structure of the weak neutral current is dependent on the fermion considered via the coefficients $c_V = T_3 - 2Q \sin^2 \theta_W$ and $c_A = T_3$. Consequently, the weak neutral current is purely axial for right-handed particles, and maximally parity violating for left-handed neutrinos, while right-handed neutrinos cannot be generated in the SM at all. Because $e = g_W \sin \theta_W$ from the mixing of the gauge fields, the only free parameters of the electroweak sector are g_W and θ_W . Furthermore, the weak coupling constant and the W boson mass can be related to the Fermi constant G_F as

$$\frac{G_F}{\sqrt{2}} = \frac{g_W^2}{8m_W^2},$$

such that $v = (\sqrt{2}G_F)^{-1/2}$ is also known as the Fermi scale.

In the SM, the W bosons have universal interaction strength with all fermions. However, because of the three quark colour states, interactions producing hadrons in lepton collisions have their cross-sections inflated by a factor of three with respect to those expected on the grounds of weak universality. The weak interaction also gave rise to the prediction of the charm as a fourth quark after EW theory was found to be in disagreement with the heavy suppression of Flavour-Changing Neutral Current (FCNC) processes occurring in quantum loops, such as the $K_L \rightarrow \mu^+ \mu^-$ meson decay mode. This disagreement can be resolved via the Glashow-Iliopoulos-Maiani (GIM) mechanism [11], where destructive interference between the up and charm virtual exchange quarks leads to the observed suppression. In 1974, the existence of the charm quark was confirmed via its quarkonium state, the J/ψ meson [95, 96].

The weak neutral current predicted by the broken EW interaction was found in bubble chamber experiments at CERN in 1973 [97, 98]. Later, direct observations of the heavy W and Z bosons were made via their leptonic decay modes at the Sp̄p̄S in 1983 [99–102]. Almost another three decades passed until the discovery of the Higgs boson in 2012 by the ATLAS and CMS experiments [103, 104] provided an experimental signature for the last particle predicted by the SM.

1.1.4. Beyond the Standard Model

Despite the overwhelming success of the SM in explaining nature over the past decades, it falls short in answering some very fundamental questions of the universe. To explain the complete lack of anti-matter in the current universe [105, 106] after Big Bang baryogenesis, the existence of baryon number, C -, and CP -violating processes occurring out of thermal equilibrium is required according to the Sakharov condi-

additionally required to generate the up-type quark masses. These terms are not included in Equation (1.10).

tions [107]. The SM does contain C and CP violation in the weak interaction. It also allows for non-perturbative baryon-number-violating processes [108–110], while the early expanding universe offered out-of-equilibrium conditions. However, the magnitude of CP violation in the SM is not sufficiently large to explain the observed matter dominance over anti-matter. Moreover, by comparing the luminous mass of galaxies and clusters with mass estimates from cluster velocity dispersions [111], rotation curves [112, 113], and gravitational lensing [114] and from of the multi-pole structure of the Cosmic Microwave Background (CMB) [115, 116], it becomes evident that the particles described by the SM can only account for a small fraction of matter in the universe in the first place. Instead, the majority of it, dubbed *dark matter*, seems to be mostly – if not exclusively – interacting gravitationally. Using CMB data collected by the Planck satellite, the fraction of baryonic matter in the Λ -CDM cosmological model¹³ is currently estimated at around just 5 %, with dark matter accounting for a further 26.5 %, and the elusive dark energy for the remaining 68.5 % [116].

Additionally, the SM has some aesthetic issues that might hint at the existence of a more fundamental underlying theory¹⁴. The SM has a large number of 26 free parameters¹⁵ corresponding to the twelve fermion masses and the Higgs boson mass, the three coupling constants, the vacuum expectation value of the Higgs field, the eight parameters of the CKM matrix and the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [119, 120] required to explain neutrino oscillations [121], and a possible CP -violating phase θ_{QCD} in the strong interaction, which is found to be close to zero experimentally [122, 123]. These degrees of freedom make the model less predictive than a possible Grand Unified Theory (GUT) with fewer free parameters. The neutrino oscillations themselves are also problematic within the SM, since they rule out the possibility of massless neutrinos. At the same time, the small experimental value of θ_{QCD} gives rise to the *strong CP* problem. Assuming that the SM is valid up to the Planck scale $\Lambda_P \sim 10^{19}$ GeV, where gravitational effects become dominant, leads to the so-called *Higgs hierarchy* problem as an additional issue: quantum loop corrections to the Higgs boson mass result in a mass given by $m_H^2 = m_0^2 + O(\Lambda_P^2)$, thus requiring an incredible amount of fine-tuning in the *bare mass* m_0 for $m_H \sim O(10^2 \text{ GeV})$ [124, 125].

Since the SM cannot be the final description of nature, many theories incorporating physics beyond the SM (BSM) to solve some of its problems have been proposed over the last decades. A full taxonomy of them is well beyond the scope of this thesis. However, a feature of many BSM models, such as Supersymmetry (SUSY)¹⁶ [130], is an extended Higgs sector. One of the simplest such extensions is the so-called Two

¹³ Λ -CDM refers to a cold dark matter model in which dark energy is encoded in a non-vanishing cosmological constant Λ , which is commonly related to the vacuum energy density [117, 118].

¹⁴ Of course, nature could also simply be messy.

¹⁵ Assuming that neutrinos behave as Dirac fermions.

¹⁶ SUSY provides a solution to the Higgs Hierarchy problem via loop cancellations [126, 127], provided that the relevant SUSY energy scale is not too high, and a Weakly Interacting Massive Particle (WIMP) [128] candidate for dark matter [129].

Higgs Doublet Model (2HDM) [131, 132], which adds a single additional Higgs doublet resulting in four additional Higgs bosons after gauging away the Goldstone bosons: one additional electrically neutral CP -even scalar, one neutral CP -odd pseudoscalar, and a pair of oppositely charged Higgs bosons. The 2HDM can act as a standalone extension of the SM, but also features, for instance, as a part of the Minimally Supersymmetric Standard Model (MSSM) [127, 133], which gives every SM particle one superparticle partner. While additional Higgs bosons could be observed directly, extended Higgs models could also lead to measurable deviations in the couplings of the already discovered SM-like Higgs boson [134, 135].

1.2. The Interplay of Top and Higgs

1.2.1. The Top Quark

The top quark is the third-generation up-type weak isospin partner of the bottom quark. It was discovered in 1995 as the last missing quark of the SM by the CDF and DØ experiments at the TEVATRON [136, 137]. With a mass of about $m_t = 172.5$ GeV, similar to that of a gold nucleus, it is the heaviest particle in the SM¹⁷. First theoretical predictions of the top quark go back to the 1970s, when Kobayashi and Maskawa proposed the third quark generation to explain CP violation [86]. The top quark was needed in this scheme to extend the GIM mechanism to three generations, as no significant amount of FCNCs was observed in b -quark decays [138].

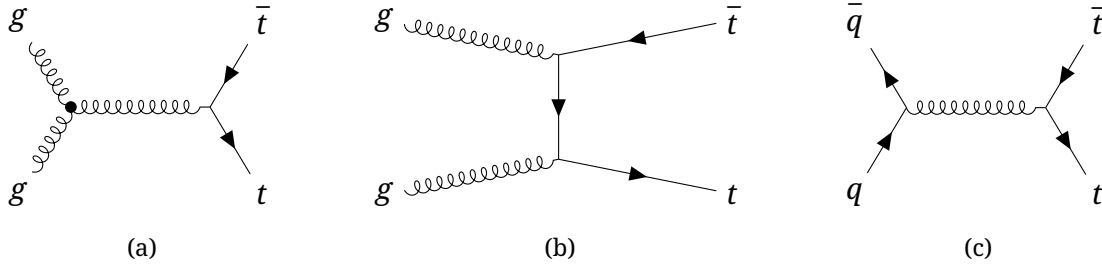


Figure 1.4: Tree-level Feynman diagrams of top quark pair production via (a) s-channel and (b) t-channel gluon–gluon fusion, and (c) quark–anti-quark annihilation.

Top quarks are produced predominantly in quark–anti-quark pairs. At the LHC, production via gluon–gluon fusion (ggF) is dominant. Production at the TEVATRON required higher parton momentum fractions x due to the lower centre-of-mass (CoM) energy there, resulting in top quark pair production mainly from quark–anti-quark annihilation with anti-proton valence quarks. Feynman diagrams of the main production modes are shown in Figure 1.4. The energy-dependent LHC cross-section is predicted as $\sigma_{t\bar{t}} = 834^{+20}_{-30}$ (QCD scales) ± 23 (PDF+ α_s) pb at a CoM energy of

¹⁷ $m_t = 172.5$ GeV is also the mass commonly assumed for the top quark in MC calculations and simulations.

$\sqrt{s} = 13$ TeV as calculated using Top++2.0 at next-to-next-to leading order (NNLO) in perturbative QCD with next-to-next-to-leading-log order (NNLL) soft-gluon resummation (c.f. Reference [139] and references therein). The QCD scale uncertainty comes from the independent variation of μ_F and μ_R , while the second uncertainty originates in variations in the PDFs and in α_s following the PDF4LHC prescription (c.f. References [140–142], and Reference [143] and references therein). Figure 1.5 shows the energy-dependent pp and $p\bar{p}$ cross-section predictions compared to the measurements of experiments at the TEVATRON and LHC at various CoM energies.

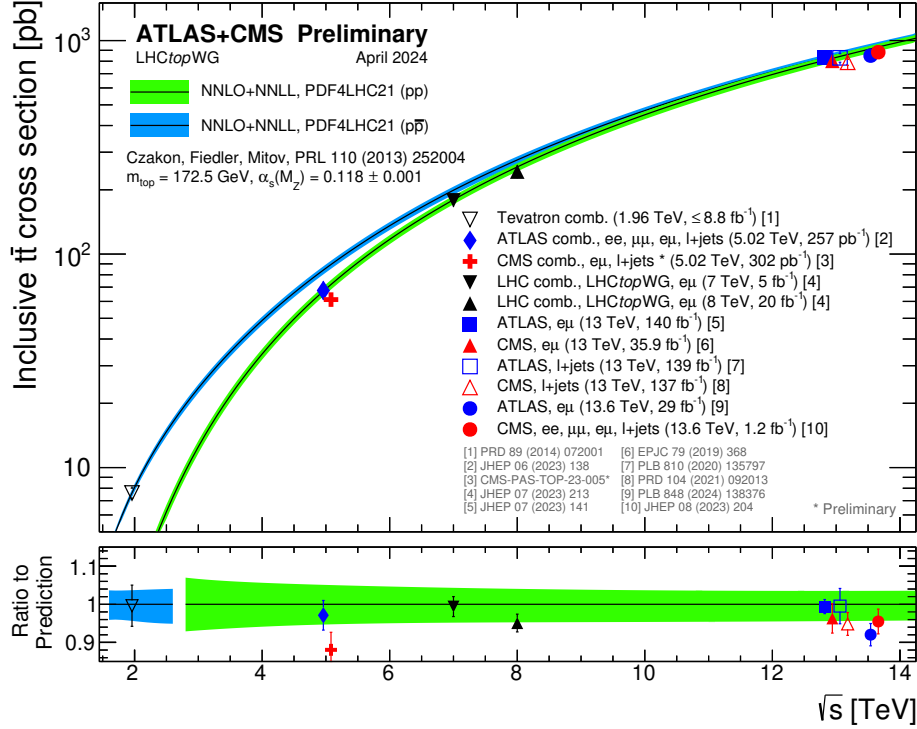


Figure 1.5: Semi-logarithmic plot of the energy-dependent top quark pair production cross-section in pp and $p\bar{p}$ collisions compared to measurements conducted at the LHC and the TEVATRON. The predictions and their uncertainty bands are calculated in the same way as described in the text. The plot is taken from Reference [144].

The production of single top quarks via electroweak processes constitutes the other primary top quark process encountered at hadron colliders. Single top quark production occurs via three main processes, whose tree-level Feynman diagrams are depicted in Figure 1.6: s -channel production, t -channel production, and the associated production of a top quark with a W boson. At 13 TeV they have pp cross-sections of 10.3 ± 0.4 pb, $214.2^{+4.1}_{-2.6}$ pb, and $79.3^{+2.9}_{-2.8}$ pb, respectively. The s -channel cross-section is calculated at NLO in QCD using HATHOR v2.1 [145, 146] with uncertainties consisting of QCD scale variations added in quadrature with PDF and α_s uncertainties in the same scheme as for $t\bar{t}$ production. Meanwhile, the t -channel and associated W boson production cross-sections constitute NNLO QCD calculations with additional third-order soft-gluon emission corrections via NNLL resummation in the

case of tW production [147, 148]. As for the s -channel cross-section, uncertainties include QCD scale variations, and PDF and α_s uncertainties based on the PDF4LHC21 PDF set [149] summed in quadrature.

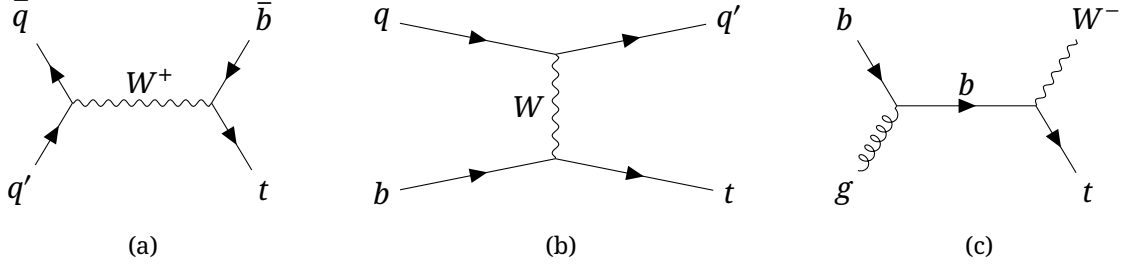


Figure 1.6: Representative tree-level Feynman diagrams of (a) s -channel, (b) t -channel, and (c) W boson associated single top quark production.

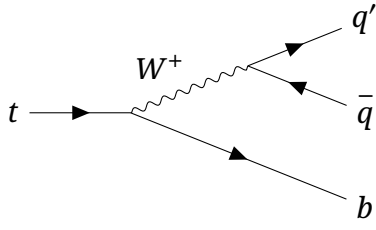
Because of the initial-state valence quarks entering the PDFs, top and anti-top quark production cross-sections for the s - and t -channel are asymmetric at the LHC and favour top quark production. Meanwhile, associated tW production interferes with LO $t\bar{t}$ production at NLO in QCD, for instance in $tW\bar{b}$ final states where $W\bar{b}$ originates from an intermediate anti-top quark. Two schemes, called *diagram removal* (DR) and *diagram subtraction* (Ds), have emerged to treat this interference in MC simulations when both processes are involved [150]. Theoretically, it is possible to distinguish single- and doubly-resonant processes as those with one or two on-shell top quarks, respectively, where the doubly-resonant processes can be associated with LO $t\bar{t}$ in the narrow-width approximation. The DR scheme then removes doubly-resonant Feynman diagrams from NLO tW calculations, while the Ds scheme introduces a subtraction cross-section term to cancel out LO $t\bar{t}$ contributions. In practice, ATLAS uses tW MC samples produced via both schemes to build a systematic uncertainty for the $t\bar{t}$ - tW interference treatment.

Top quarks can also be produced in association with other SM particles, such as the weak bosons, photons, the Higgs boson, or even other fermions. A detailed discussion is beyond the scope of this thesis. The associated production with a Higgs boson, which forms the main topic of this thesis, is outlined separately in Section 1.2.3.

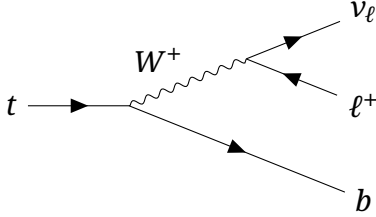
Because of $|V_{tb}| \sim 1$, the top quark decays almost exclusively into a bottom quark and a W boson. Top decays are classified into *leptonic* and *hadronic* based on the W boson decay modes shown in Figure 1.7. From lepton flavour universality and the colour factors involved, leptonic decays are expected with a Branching Ratio (BR) of about one third¹⁸. Therefore, processes involving a top quark pair are divided into fully hadronic, semileptonic, and dileptonic channels with BRs of 46 %, 45 %, and 9 %, respectively, as shown in Figure 1.8. Jets initiated by bottom quarks are present in every channel. However, only hadronic jets are present in the final state of fully hadronic decays¹⁹, so that the reconstruction of event topologies is challenging

¹⁸ Assuming $|V_{tb}| = 1$, the third quark generation is kinematically inaccessible to the W boson decay.

¹⁹ Hence the name.



(a) Hadronic top quark decay.



(b) Leptonic top quark decay.

Figure 1.7: Tree-level Feynman diagrams of the top quark decay modes. The decay modes of anti-top quarks are analogous.

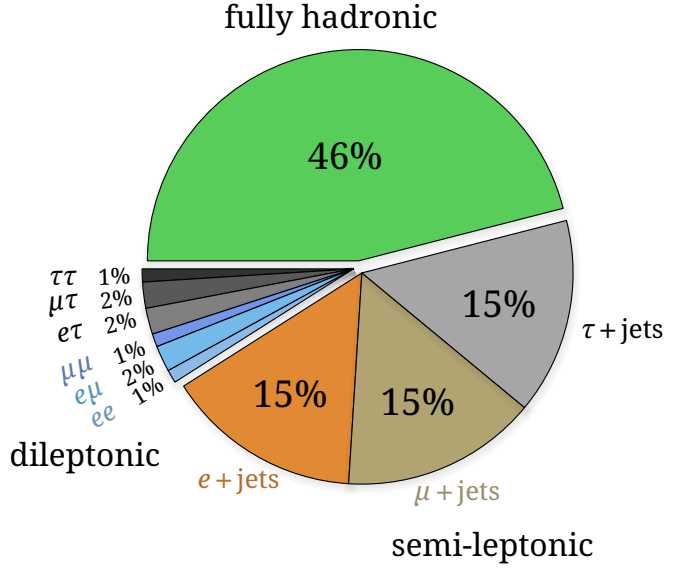


Figure 1.8: Pie chart containing the branching fractions of the various top quark pair decay modes grouped into fully hadronic, semi-, and dileptonic modes, where the tau lepton modes are not directly accessible because of successive tau decays.

Original graphic by Izaak Neutelings (CC-BY-4.0 license), modified.

experimentally. Additionally, QCD multi-jet background processes dominate the corresponding measurements. In contrast, the semi- and dileptonic decay channels contain high-energy leptons in combination with a sizeable energy imbalance in the plane transverse to the colliding beams, called the *missing transverse energy* E_T^{miss} , from associated undetected neutrinos. Therefore, they represent the preferred channels for many analyses of top quark processes.

The top quark width at NNLO in QCD and with NLO EW corrections is predicted to be $\Gamma_t = 1.33 \text{ GeV}$ [151], which is in agreement with direct ATLAS and indirect CMS measurements of $\Gamma_t = 1.9 \pm 0.5 \text{ GeV}$ and $\Gamma_t = 1.36 \pm 0.02 \text{ (stat.) }^{+0.14}_{-0.11} \text{ (syst.) GeV}$, respectively [152, 153]. Therefore, the top quark decays after around $\tau_t \sim 5 \times 10^{-25} \text{ s}$. The large top quark mass significantly contributes to the short decay timescale, because it lies over the threshold of the Wb system. Crucially, the top quark decays more rapidly than the hadronisation timescale of $\tau_{\text{Had}} \sim 10^{-23} \text{ s}$ [154] and before spin decorrelation can occur [155]. As a result, top quark processes provide a singular opportunity to study the properties of bare quarks. Additionally, top quark pair production provides a relativistic testbench for quantum information experiments involving entanglement and Bell-inequality [156] measurements [157, 158] with first observations of entanglement in the dileptonic channel by the ATLAS and CMS experiments [159, 160].

1.2.2. The Higgs Boson

After the discovery of the Higgs boson in 2012 [103, 104] as the last missing particle of the SM, the ATLAS and CMS collaborations have started to investigate its properties to test the consistency of the theory. Of central importance is the Higgs boson mass $m_H = \sqrt{2}\lambda v$, which has to be experimentally determined as a free parameter of the SM. Combining the measurements of the ATLAS and CMS experiments with data taken in 2011 and 2012, the so-called Run 1 of the LHC, yields a Higgs boson mass²⁰ of $m_H = 125.09 \pm 0.24$ GeV [161]. More recently, the ATLAS and CMS collaborations obtained $m_H = 125.11 \pm 0.11$ GeV [162] and $m_H = 125.38 \pm 0.14$ GeV [163], respectively, in combinations of their Run 1 and Run 2 measurements²¹. Both experiments have also confirmed that the particle discovered in 2012 behaves consistently with a scalar boson having CP -even couplings to the vector bosons [164–167] as postulated by the SM.

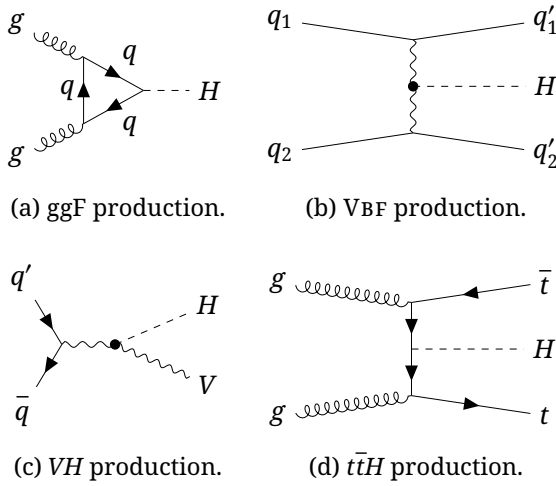


Figure 1.9: Representative tree-level Feynman diagrams of the four major Higgs boson production modes at the LHC.

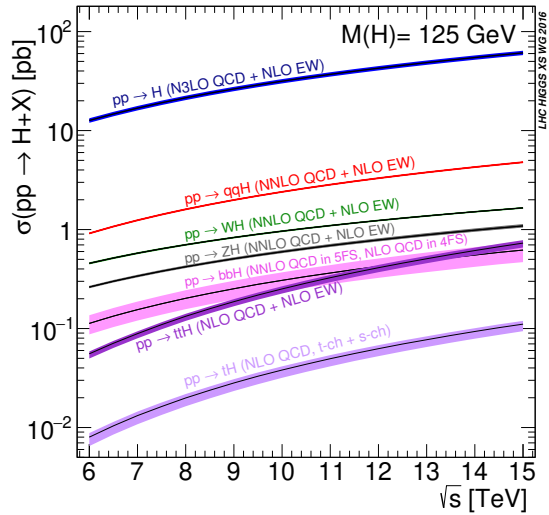


Figure 1.10: Semi-logarithmic plot showing the energy-dependent Higgs boson production cross-sections with their uncertainty bands. The plot is taken from Reference [168].

Higgs boson processes can be classified by their production and m_H -dependent decay modes. On the production side, there are four main processes as depicted in Figure 1.9. The dominant process, as for $t\bar{t}$ production, is ggF with a pp cross-section of 48.5 pb at 13 TeV and at N³LO in QCD [169]. Because the massless gluons do not directly interact with Higgs bosons, a virtual quark loop dominated by the heavy top quarks with a large coupling to the Higgs boson²² is mediating the interaction,

²⁰ Commonly used as the current reference mass value in analyses.

²¹ Run 2 corresponds to data taken between 2015 and 2018 at $\sqrt{s} = 13$ TeV. The CMS Higgs boson mass measurement uses the partial 2016 dataset.

²² $y_f^2 \propto m_f^2$ enters the matrix element at LO.

leading to an effective coupling. Vector Boson Fusion (VBF) represents the subleading contribution at 13 TeV with a cross-section about an order of magnitude lower than that of ggF [168]. Because of the t -channel-like scattering of the two initial-state quarks, VBF is characterised by high-energy jets with large momenta in the beam directions in addition to the Higgs boson decay products. VBF allows the direct measurement of the couplings between the Higgs boson and the weak bosons. The third Higgs boson production process is associated vector boson production (VH), also called Higgs-strahlung, where V is used to refer to both types of weak bosons. Overall, VH production has a cross-section in the same order of magnitude as VBF, but allows more precise studies of Higgs boson decay modes involving quarks, since leptonic vector boson decay modes provide non-hadronic final state particles to reduce QCD backgrounds in a similar fashion to the leptonic top decay channels. At the same time, the restriction to these decay modes decreases the ‘visible’ VH cross-section significantly.

The associated production of a Higgs boson with top quarks forms the fourth primary production mode, which is of special interest to this thesis and therefore covered in detail in Section 1.2.3. The $t\bar{t}H$ cross-section is almost two orders of magnitude smaller than ggF production. Analogously to top quark processes, tH production modes also exist with an even smaller cross-section than $t\bar{t}H$. Theoretically, associated production with bottom quarks is also possible with analogous Feynman diagrams to the associated production with top quarks, such as the t -channel production shown in Figure 1.9(d). However, $b\bar{b}H$ events are harder to identify in the collider data since the characteristic top quark decays are not present in the process. An overview of the energy-dependent Higgs boson production cross-sections is shown in Figure 1.10.

Higgs boson decays to fermion–anti-fermion and vector boson pairs are possible in principle, because of the Higgs boson interactions with all massive SM particles. From the kinematic restrictions set by the Higgs boson mass, $H \rightarrow WW^*$ and $H \rightarrow ZZ^*$ decays can only occur either via an off-shell Higgs boson or an off-shell vector boson, as signified by the star notation. The same also applies to the production of a $t\bar{t}$ pair from a Higgs boson, which nevertheless represents a significant contribution to $t\bar{t}t\bar{t}$ production. Additionally, Higgs boson decays to gluons or photons are possible via loop processes in a similar fashion to ggF Higgs boson production. Representative Feynman diagrams of the Higgs boson decay modes are shown in Figure 1.11, while the BRs for an SM Higgs boson are given in Table 1.1. To highlight the role of threshold effects, the BRs are also plotted as a function of the Higgs boson mass in Figure 1.12. As in the case of top quark decays, Higgs boson decays have a trade-off between channels with high BR but challenging signal extraction, such as $H \rightarrow b\bar{b}$, and low BR but clean signal extraction like $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$, which were the first individual decay channels to be observed [170–173].

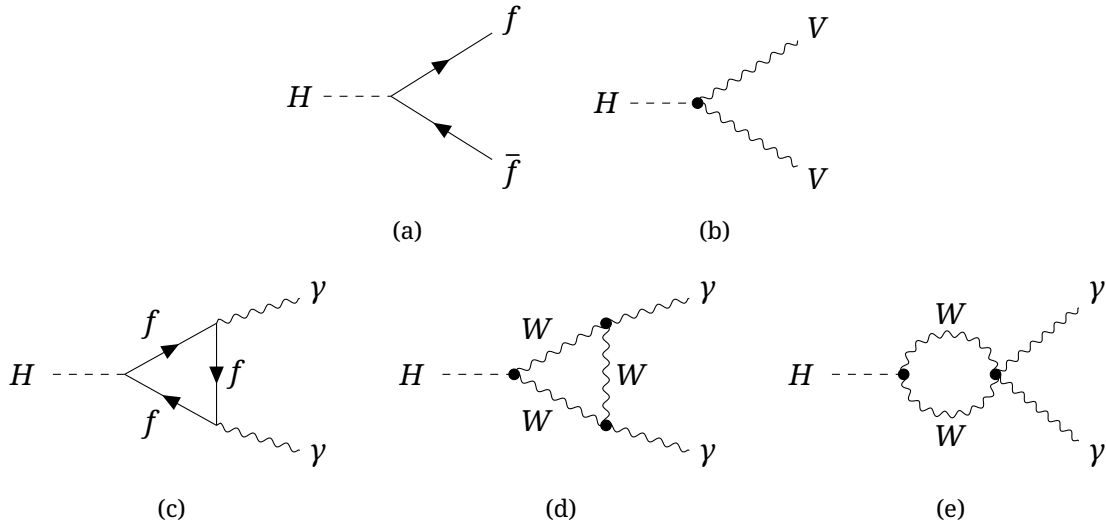


Figure 1.11: Representative Feynman diagrams of Higgs boson decays to (a) fermion anti-fermion pairs, (b) pairs of W or Z bosons, and photons via virtual (c) fermion and (d) + (e) W boson loops. Top quarks dominate the fermion loops. The decay into gluons occurs via virtual quark loops analogously to (c). Instead of the decay into two photons, a decay into a photon and a Z boson is also possible.

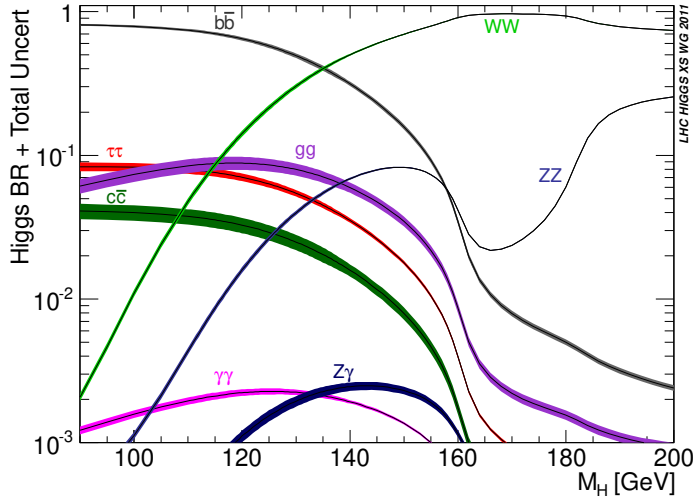


Figure 1.12: Semi-logarithmic plot of the Higgs boson branching ratios and uncertainty bands in relation to the Higgs boson mass. The threshold effects of the doubly-resonant $H \rightarrow WW^*$ and $H \rightarrow ZZ^*$ decay channels at around 160 GeV and 180 GeV, respectively, are visible. The plot is taken from Reference [174].

Table 1.1: Higgs boson branching ratios (BRs) for an SM Higgs boson of mass $m_H = 125.09$ GeV. Not all decay channels are listed. The BR values are taken from Reference [168].

Decay Mode	BR [%]
$H \rightarrow b\bar{b}$	58.09
$H \rightarrow WW^*$	21.52
$H \rightarrow gg$	8.18
$H \rightarrow \tau^+\tau^-$	6.26
$H \rightarrow c\bar{c}$	2.88
$H \rightarrow ZZ^*$	2.64
$H \rightarrow \gamma\gamma$	0.23
$H \rightarrow Z\gamma$	0.15
$H \rightarrow \mu^+\mu^-$	0.02

In principle, di-Higgs boson production processes are also attainable at the LHC. Their measurement allows a direct probe of the trilinear Higgs boson self-coupling,

$$g_{HHH} = \frac{m_H^2}{2v},$$

and of the vacuum expectation value v currently inferred to be $v \approx 246$ GeV [37]. Similar to single Higgs boson production, ggF-initiated di-Higgs boson production is the dominant process in the SM. However, the associated cross-section is about three orders of magnitude smaller than single Higgs boson production via ggF [168], which is in part due to destructive interference with the simultaneous production of two Higgs bosons from a virtual top quark loop [175, 176]. Nevertheless, the observation of di-Higgs boson production is one of the primary goals of the LHC physics programme going forward.

1.2.3. The Top Yukawa Coupling

A dozen years after the discovery of the Higgs boson, its couplings to the third generation fermions – except for ν_τ –, and the gauge bosons have been established. The ATLAS and CMS collaborations have also started tackling the second-generation couplings in dedicated analyses of $H \rightarrow \mu^+ \mu^-$ [177, 178] and $H \rightarrow c\bar{c}$ [179–181]. The current status of the ATLAS and CMS Higgs boson coupling measurements is depicted in Figure 1.13.

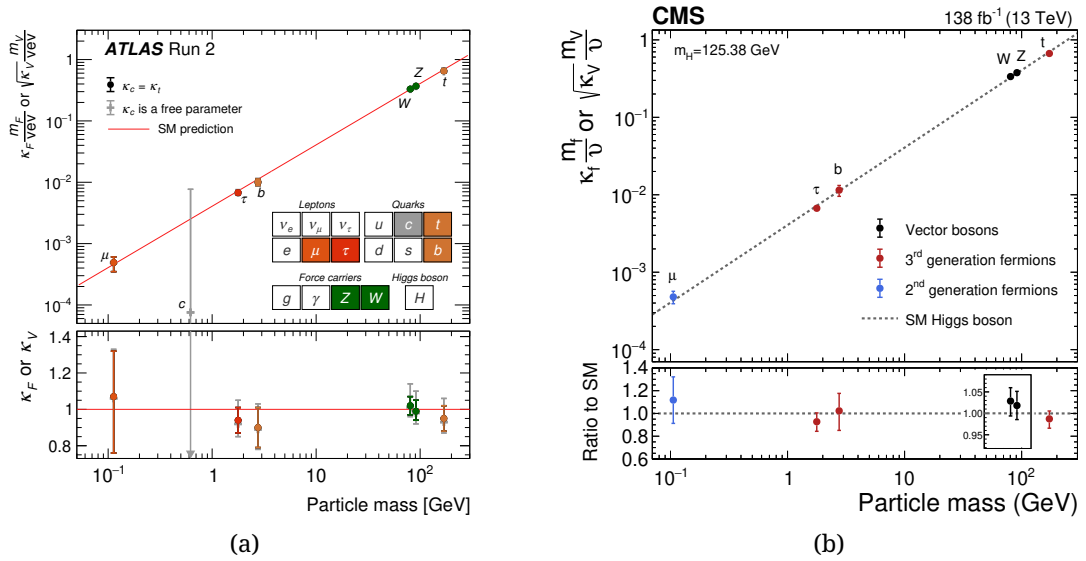


Figure 1.13: Overview over the Higgs boson couplings already measured by (a) the ATLAS and (b) CMS experiments in relation to the respective particle masses. The plots are taken from References [182, 183].

A special role among the Higgs boson couplings is played by the top Yukawa coupling y_t , defined from Equation (1.9) as

$$y_t = \sqrt{2} \frac{m_t}{v}.$$

It is the largest Yukawa coupling in the SM because of the large top quark mass and has a value of almost unity. While this might be a sheer coincidence, it might equally

hint at a special role of the top quark in EWSB.

The top Yukawa coupling is in principle experimentally accessible through the ggF production mode (Figure 1.9(a)) and the $H \rightarrow \gamma\gamma$ decay channel (Figure 1.11(c)), where top quarks play a dominant role in the associated fermion loops. However, to extract y_t from these measurements, an assumption of no BSM contributions to the loops has to be made, since modifications of the effective loop-induced couplings can be largely independent of y_t in BSM models (c.f. Reference [184], for instance). Therefore, no general statement about y_t can be made based on measurements of fermion loop processes. The top Yukawa coupling is also indirectly accessible via the top quark mass. However, problems in translating measured values based on the Mc top quark mass into the required top quark pole mass and uncertainties arising from QCD effects such as hadronisation following the coloured top quark decays render these interpretations difficult [185, 186]. Consequently, the only direct probes of the top Yukawa coupling in the SM via tree-level processes are $t\bar{t}H$ and tH production depicted in Figures 1.14 and 1.15, respectively. tH can be split further into associated tWH production, and s- and t-channel sub-processes in analogy to the production of a single top quark.

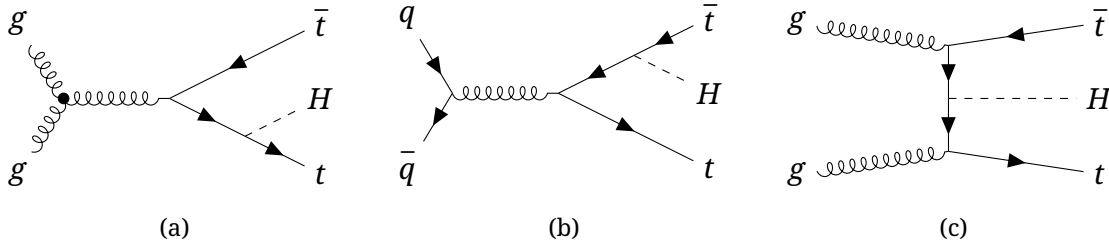


Figure 1.14: Representative tree-level Feynman diagrams of $t\bar{t}H$ production from (a) s-channel ggF, (b) s-channel $q\bar{q}$ annihilation, and (c) t-channel ggF.

$t\bar{t}H$ production has a cross-section of $\sigma_{t\bar{t}H} = 506^{+29}_{-47}$ (QCD scales) ± 18 (PDF+ α_S) fb at 13 TeV and using NLO in QCD with NLO EW corrections [168]. Meanwhile, tH production is predicted with cross-sections of 74^{+6}_{-11} fb and $16.3^{+0.9}_{-1.1}$ fb for t-channel production and tWH associated production²³, respectively, while s-channel production has a negligible contribution [168, 187]. In analogy to the top production processes discussed in Section 1.2.1, tWH also interferes with $t\bar{t}H$ production, so that similar interference removal schemes are used [187]. Additionally, there exists some negligible interference amongst the tH processes themselves [188]. t-channel tH production will hereafter be referred to as $tHjb$, where the final state b -quark is present at tree-level when the so-called Four-Flavour Scheme (4Fs) is used in the Mc simulations of the process. In the 4Fs, the b -quarks are assumed massive and do not contribute to the PDFs²⁴.

²³ Calculated with $m_H = 125$ GeV and the DR2 procedure described in [187].

²⁴ Calculations can also be performed in the Five-Flavour Scheme (5Fs) with massless b -quarks, which are then included in the PDFs. 4Fs and 5Fs deliver complementary information: The 4Fs provides more accurate exclusive and differential predictions, while the 5Fs yields accurate total cross-sections at higher perturbative order due to simpler setups [189].

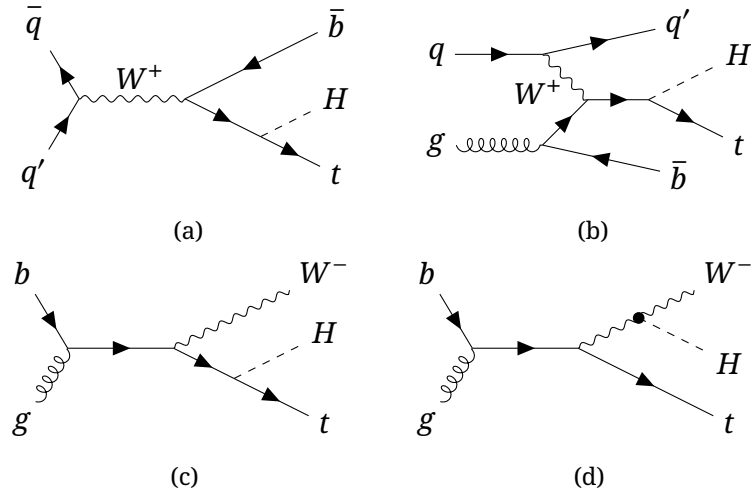


Figure 1.15: Representative tree-level Feynman diagrams of (a) s-channel tH production, (b) 4Fs t -channel tH production, and 5Fs tWH production with Higgs bosons radiated from (c) a top quark and (d) a W boson. The corresponding anti-top processes occur analogously.

The $tHjb$ and tWH processes are affected by destructive interferences between diagrams in which the Higgs boson is radiated from top quarks and W bosons, resulting in lowered SM cross-sections [187, 188]. The corresponding Feynman diagrams for tWH are shown in Figures 1.15(c) and 1.15(d), respectively. This interference renders the inclusive cross-section of tH production sensitive to the sign of the top Yukawa coupling. At the same time, the full CP -structure of y_t can be probed in $t\bar{t}H$ and tH measurements, for instance via angular correlations between final state objects or via cross-section comparisons, since $\sigma_{t\bar{t}H}$ decreases with increasing admixture of a CP -odd coupling component while σ_{tH} increases [190–192].

In addition to the BSM sensitivity already discussed in Section 1.1.4, the top Yukawa coupling also has important ramifications for the overall stability of the Higgs potential $V(\phi)$ of Equation (1.7): with the inclusion of quantum loop corrections at high energy scales, the shape of the potential is modified [193–195]. As a significant contribution to the corrections, large values of the top Yukawa coupling can cause a second minimum in the Higgs potential [196]. According to current measurements, $V(\phi)$ turns out to be metastable at the vacuum expectation value v , resulting in the possibility of *vacuum decay* of the universe unless new physics stabilising the Higgs potential emerges somewhere below the Planck scale [194, 196].

$t\bar{t}H$ searches have already been conducted in Run 1 by the ATLAS [197] and CMS [198, 199] collaborations. In 2018, the process was observed by both experiments [200, 201] by using a combination of the Run 1, and partial Run 2 datasets in the $H \rightarrow b\bar{b}$, $H \rightarrow \gamma\gamma$, resonant $H \rightarrow ZZ^* \rightarrow 4\ell$, and the so-called multi-lepton Higgs boson decay channel, $H \rightarrow ML$, which includes $H \rightarrow ZZ^*$, $H \rightarrow WW^*$, and $H \rightarrow \tau^+\tau^-$ decay modes with final state leptons²⁵ in addition to at least one additional lepton from the top

²⁵ Excluding $H \rightarrow ZZ^* \rightarrow 4\ell$ with an on-shell Higgs boson.

quark decays²⁶. Since then, $t\bar{t}(H \rightarrow \gamma\gamma)$ has been observed as the only single-channel $t\bar{t}H$ process so far [203, 204], although both collaborations have claimed strong evidence²⁷ for $t\bar{t}(H \rightarrow \text{ML})$ production in Run 2 data [202, 205, 206]. Additionally, analyses of the CP structure of y_t conducted in the various Higgs boson decay channels under inclusion of tH consistently find compatibility with the SM CP -even hypothesis [203, 207, 208].

In the $H \rightarrow b\bar{b}$ decay channel, which is targeted by the analysis presented in this thesis, no leptons are expected from Higgs boson decay products. As a result, in addition to the four b -jets expected from the top quark and Higgs boson decays, the tree-level final states either contain two leptons from the dileptonic top decay channel, a single lepton and two additional jets from the semileptonic channel, or four additional jets from the fully hadronic channel. So far, ATLAS analyses have been split in those targeting leptonic [1, 209–211] and hadronic [197] decay channels separately, while CMS analyses either proceeded with the same channel separations [198, 212, 213] or bundled all decay channels together [199, 208]. While the fully hadronic $t\bar{t}(H \rightarrow b\bar{b})$ channel has the largest BR and is fully reconstructable, its background contamination is also exceptionally high due to large QCD multi-jet backgrounds in addition to the $t\bar{t} + \text{jets}$ backgrounds dominating for the other $t\bar{t}(H \rightarrow b\bar{b})$ channels. Therefore, hadronic $t\bar{t}(H \rightarrow b\bar{b})$ analyses rely on data-driven multi-jet background estimates in combination with signal–background discrimination techniques. For the latter, *boosted* topologies [214, 215], where very high Higgs boson transverse momenta lead to collimated b -jets in a single fat jet, can be included in $t\bar{t}(H \rightarrow b\bar{b})$ analyses as well to profit from mass constraints and the substructure of the fat jet. The boosted dilepton $t\bar{t}(H \rightarrow b\bar{b})$ channel can additionally be used to probe the CP -structure of the top Yukawa coupling [216]. Boosted events have formed separate categories in the semileptonic channel of the Run 2 ATLAS analyses.

The first ATLAS full Run 2 leptonic $t\bar{t}(H \rightarrow b\bar{b})$ analysis [211] measured a $t\bar{t}H$ signal strength²⁸ of $\mu_{t\bar{t}H} = 0.35^{+0.36}_{-0.34}$, corresponding to an observed event excess with a significance of 1.0σ , while 2.7σ were expected. Recently, CMS also published a full Run 2 $t\bar{t}(H \rightarrow b\bar{b})$ cross-section analysis encompassing all top decay modes [208], which resulted in a measured signal strength of $\mu_{t\bar{t}H} = 0.33 \pm 0.26$ (observed significance of 1.3σ compared to 4.1σ expected). Meanwhile, the ATLAS collaboration conducted a re-analysis of the Run 2 data in the leptonic top decay channels to supersede Reference [211] – the Run 2 legacy analysis [1] presented in Part III of this thesis. A discussion of the improvements incorporated into the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis is deferred to Chapter 6.

²⁶ CMS used one category targeting $H \rightarrow \tau^+\tau^-$ in combination with fully hadronic top decays in the later full Run 2 $t\bar{t}(H \rightarrow \text{ML})$ analysis as well [202].

²⁷ While an event excess with a significance of more than 5 standard deviations (denoted as 5σ) is required to claim an observation, significances above 4σ are considered strong evidence.

²⁸ The signal strength μ is the ratio of measured cross-section σ_{obs} to SM cross-section σ_{SM} .

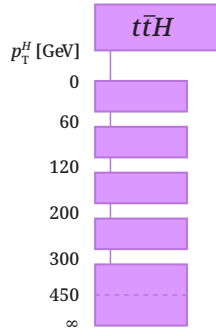


Figure 1.16: Schematic of the $t\bar{t}H$ stage 1.2 Simplified Template Cross-Section (STXS) formalism using p_T^H , the Higgs boson transverse momentum.

Individual $t\bar{t}H$ analyses do not extract y_t directly, instead measuring cross-sections. Afterwards, global Higgs combinations are performed in the so-called κ -framework [217], which gives every (effective) Higgs boson coupling a scale factor κ to disentangle effects of the different Higgs boson production and decay modes in the narrow-width approximation. Additionally, the Simplified Template Cross-Section (STXS) formalism [168, 218] is used as a guideline for combined Higgs measurements. While the original motivation for the STXS formalism was the disentanglement of QCD scale effects during combinations of processes such as ggF and VBF, it has evolved to provide a standardised binning for differential cross-section measurements in the Higgs sector. The STXS Higgs boson transverse momentum (p_T^H) bins in stage 1.2, as used for the differential cross-section measurement of the ATLAS $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis, are shown in Figure 1.16. In global STXS combinations, $t\bar{t}(H \rightarrow b\bar{b})$ analyses contribute especially in the high- p_T^H regime, where other Higgs boson decay channels suffer from high uncertainties. This strength of $t\bar{t}(H \rightarrow b\bar{b})$ for combinations is further compounded when including boosted topologies.

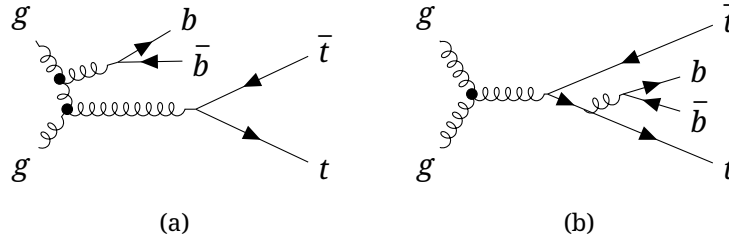


Figure 1.17: Representative tree-level Feynman diagrams of $t\bar{t} + b\bar{b}$ production via gluon splitting from (a) an initial state gluon and (b) a top quark. The final state in either case is identical with $t\bar{t}(H \rightarrow b\bar{b})$.

The main challenge for $t\bar{t}(H \rightarrow b\bar{b})$ measurements lies in the dominant $t\bar{t} + b\bar{b}$ background, which results in an identical final state as depicted in Figure 1.17. For this reason, $t\bar{t} + b\bar{b}$ is called an *irreducible* background, meaning that simple event cuts cannot remove it. Additionally, it dominates the analysis phase space with a cross-section that is about two orders of magnitude above that of $t\bar{t}(H \rightarrow b\bar{b})$ [168]. Consequently, analyses rely on Multivariate Analysis (MVA) techniques to discriminate between signal and background events. The development of event classification and reconstruction techniques for $t\bar{t}(H \rightarrow b\bar{b})$ forms one of the main topics of this thesis and is discussed in detail in Chapter 7. Because of the significant impact on MC simulations when considering the non-negligible b -quark mass, modelling of $t\bar{t} + b\bar{b}$ exists both in 4Fs and 5Fs and has been extensively studied recently [219, 220] to inform the setups for future LHC analyses and to reduce modelling uncertainties associated with $t\bar{t} + b\bar{b}$. Additionally, schemes *fusing* 4Fs and 5Fs [221, 222] have

been developed recently with the promise of more precise background modelling for future $t\bar{t}(H \rightarrow b\bar{b})$ analyses. To complement the development of Mc simulations, dedicated measurements of inclusive and differential $t\bar{t} + b\bar{b}$ cross-sections are also being performed by the ATLAS and CMS experiments [223, 224]. So far, they have found the Mc predictions to underpredict data yields consistently. In the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis, $t\bar{t} + b\bar{b}$ is further subdivided into three categories (denoted as $t\bar{t} + \geq 1b$ in combination), as illustrated in Figure 1.18:

- $t\bar{t} + \geq 2b$, containing at least two b -jets in addition to those produced in the top quark decays.
- $t\bar{t} + 1B$, where the b -quarks are produced collinearly, such that their hadron showers overlap and combine into a single jet.
- $t\bar{t} + 1b$, where one of the additional b -jets cannot be reconstructed in the detector because it falls out of acceptance. As a result, $t\bar{t} + 1b$ is the most difficult subprocess to be modelled by Mc generators.

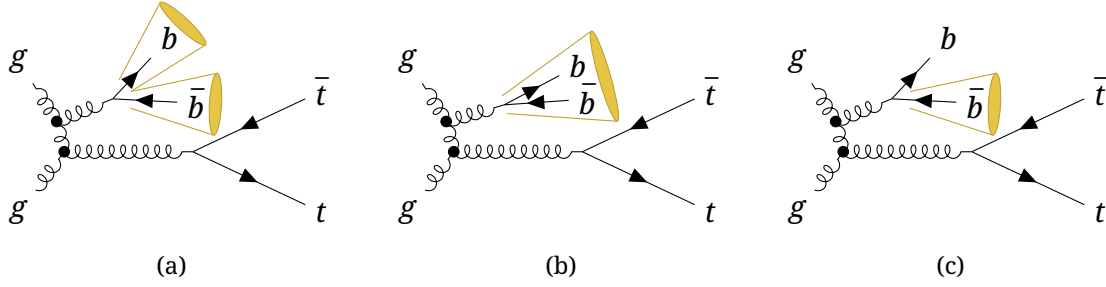


Figure 1.18: Schematic depictions of the $t\bar{t} + b\bar{b}$ categories used in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis to split the process into (a) $t\bar{t} + \geq 2b$, (b) collimated $t\bar{t} + 1B$, and (c) $t\bar{t} + 1b$, which is dominated by acceptance effects in 4fs simulations. Jets not originating from the top quark decays are represented by cones encircling the respective final-state particles.

The CERN Accelerator Complex and the ATLAS Experiment

In the early days of particle physics research, experimental apparatus was small enough to fit onto a workbench – or at least into a single room. Nowadays, the experimental tools of the trade have increased in size slightly: currently, the Large Hadron Collider (LHC) [225] at the *Conseil Européen pour la Recherche Nucléaire* (CERN) is the world’s most powerful particle accelerator at the high-energy frontier. Situated at the Swiss-French border near Geneva, it is a near-circular¹ proton–proton (pp) and heavy-ion synchrotron with a circumference of approximately 26.7 km.

Around the LHC, several collider experiments make use of the TeV-scale particle collisions generated at a total of four interaction points: ATLAS [226] and CMS (Compact Muon Solenoid) [227] situated on opposite sides of the LHC are general-purpose detectors designed to investigate a large variety of high-energy physics phenomena. LHCb (Large Hadron Collider beauty) [228], built in a forward spectrometer configuration, is specialised on flavour physics to search for the origin of CP -violation, while ALICE (A Large Ion Collider Experiment) [229] is designed to investigate the quark-gluon plasma in heavy-ion collisions for a better understanding of conditions in the very early universe. In addition to the four large experiments, the LHC additionally houses the smaller LHCf [230], TOTEM [231], MoEDAL [232], and – as more recent additions – the FASER [233] and SND@LHC [234] experiments.

In the following, the LHC and its injector chain will be described in Section 2.1. The concepts of luminosity and pile-up are additionally outlined in Section 2.1.2. As the work described in this thesis was conducted with the ATLAS detector, its detector design will be presented in detail in Section 2.2.

¹ The LHC is an octagon with rounded corners consisting of eight approximately 528 m long straight sections connected by eight arcs.

2.1. The Large Hadron Collider

The LHC is located between 45 m and 170 m underground in the preexisting tunnel of the Large Electron-Positron Collider (LEP) [235] running at CERN from 1989 until 2000. It consists of 1232 superconducting NbTi dipole magnets cooled to 1.9 K by superfluid helium to generate magnetic fields of up to 8.33 T for bending the proton beams circulating in two separate beam pipes around the curved sections [225]. Additionally, a large number of higher-order multipole magnets is used for beam correction and insertion into the eight straight sections housing the four main experiments, the 16 superconducting radiofrequency cavities for beam acceleration, two beam cleaning systems, and a beam dump system. For instance, to focus the beams at the interaction points at the centre of the four experiments, triplet quadrupole magnet assemblies are used [236].

Instead of continuous beams, up to 2808 bunches of around 1.15×10^{11} protons each are rotating in each of the two rings of the LHC. During operation, collisions of the around 7.5 cm long bunches occur every 25 ns at the interaction points, corresponding to 40 million so-called bunch crossings per second.

The LHC is designed to accelerate protons from beam energies of 450 GeV at injection to 7 TeV, resulting in a CoM energy of $\sqrt{s} = 14$ TeV – the highest currently achievable in collider experiments. So far, the design energy has not yet been reached because of the associated strain on the magnets. Instead, pp collisions were attained at energies of up to 13.6 TeV so far.

Operation of the LHC is split into distinct multi-year periods called *Runs*. These are interspersed with *Long Shutdown* periods during which the accelerator and the experiments conduct major upgrade and replacement works. During the Runs, physics data taking is performed for about 150 days per year [237, 238] in so-called *fills*, while commissioning, LHC machine development, special beam setups for calibration or specific physics objectives, and a regular Year-End Technical Stop (YETS) make up the remaining time. In each fill, the beam energy is first ramped up to the target value after all bunches have been injected into the LHC. Afterwards, the beams are *squeezed* (i.e. focused) before data taking can commence during the so-called *stable beams* period. At the end of each fill or in case of issues such as beam degradation from unidentified falling objects [239, 240], the beams are dumped into graphite cylinders weighing about 4.5 tonnes by two assemblies of 15 kicker magnets each. A summary of LHC operations to date is given in Table 2.1.

2.1.1. The LHC Accelerator Chain

In order to accelerate the proton beams to their final energy, several pre-stages are used in addition to the LHC itself. This injection chain [241] starts by stripping hydrogen atoms off of their electrons by use of an electric field. The protons created

2.1. The Large Hadron Collider

in this way are then accelerated from rest to 50 MeV in the CERN Linear Accelerator 2 (LINAC2) [242] operating since 1978. Afterwards, they enter the Proton Synchrotron Booster (PSB) consisting of four superimposed rings, where they are accelerated further to an energy of 1.4 GeV before transferring into the Proton Synchrotron (PS). The energy bump from the PSB hereby helps to increase the number of protons accepted by the PS at a time. In the PS, protons destined for the LHC reach an energy of 26 GeV. At this point, between two and four PS batches² sequentially fill the Super Proton Synchrotron (SPS) for a further ramp-up to an energy of 450 GeV. Finally, the protons in the SPS are used to fill both rings of the LHC via two transfer lines running in opposite directions. Each SPS cycle lasts 21.6 s, such that the LHC can be filled by 24 SPS cycles in about 9 minutes. During Long Shutdown 2 following Run 2, LINAC2 was replaced by LINAC4³ [243] as part of the upgrades towards the high-luminosity LHC slated to start operation after Long Shutdown 3.

The injection of heavy ions makes use of LINAC3 and the Low Energy Ion Ring (LEIR) before the PS. The layout of the CERN accelerator complex is shown in Figure 2.1.

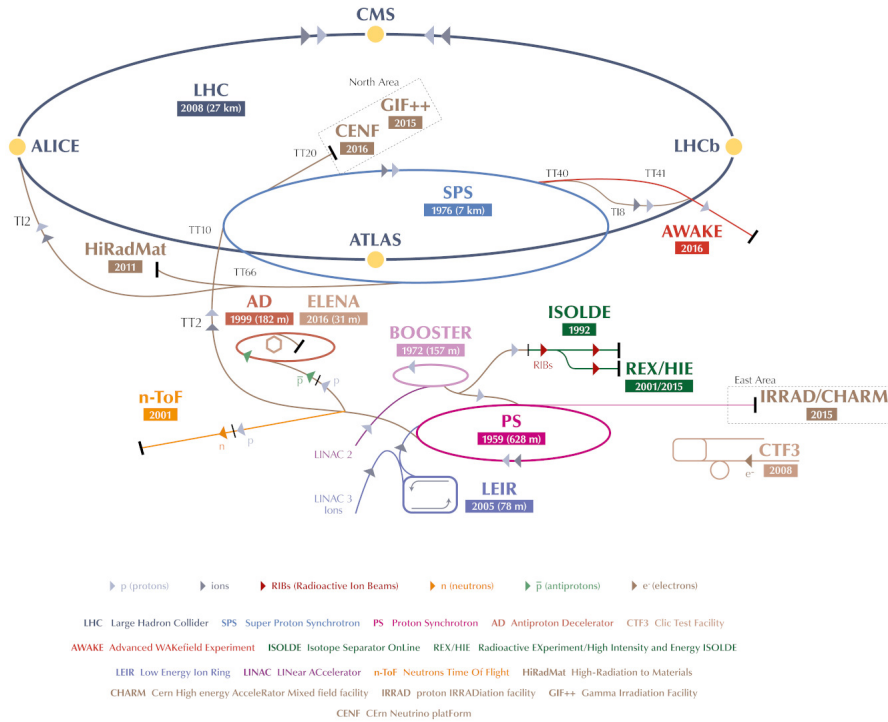


Figure 2.1: Layout of the CERN accelerator complex during Run 2 in 2016. Protons destined for the LHC come from LINAC2, travel through the PSB, the PS, and the SPS before they are finally injected into the LHC. During Long Shutdown 2 following Run 2, LINAC2 was replaced by LINAC4.

© 2016 CERN.

² This depends on the bunching scheme used, i.e. how the individual proton packets are arranged in the LHC.

³ LINAC4 accelerates negatively charged hydrogen ions, which have their electrons stripped during the transfer to the PSB.

2.1.2. Luminosity and Particle Collision Conditions

In particle physics, detectors are primarily able to measure the number of collision events N instead of the physically more meaningful cross sections σ . Therefore, the *instantaneous luminosity* $\mathcal{L}(t)$ integrated over the collider runtime is used as a normalisation factor between the two:

$$\begin{aligned} N &= \int \sigma \cdot \mathcal{L}(t) dt \\ &= \sigma \cdot \mathcal{L}_{\text{int}}, \end{aligned} \quad (2.1)$$

where it is assumed for now that all collision events are counted without inefficiencies. The second line instead makes use of the *integrated luminosity* $\mathcal{L}_{\text{int}}$.

At the LHC, the instantaneous luminosity can be directly calculated from the beam parameters under the assumption of a Gaussian beam profile as

$$\mathcal{L}(t) = \frac{N_b^2 n_b f_{\text{rev}}}{4\pi\sigma_x\sigma_y} F(\theta_c). \quad (2.2)$$

Hereby, N_b is the number of protons per bunch, n_b the number of colliding bunches per beam, f_{rev} denotes the revolution frequency of the beams in the LHC (about 11.2 kHz), σ_x and σ_y are the horizontal and vertical root-mean-square sizes of the beam at the interaction point, respectively, and the geometric factor $F(\theta_c)$ corrects for the fact that the two beams do not collide head-on but with a crossing angle θ_c . The LHC is designed for an instantaneous luminosity $\mathcal{L}(t) = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ at the ATLAS and CMS interaction points. In practice, the experiments do not rely on the LHC luminosity derived from Equation (2.2). Instead, they employ dedicated subdetectors (c.f. Section 2.2.5 in the case of ATLAS) for its measurement via well-understood reference processes, essentially turning Equation (2.1) around. The accumulated pp luminosity delivered to and recorded by the ATLAS detector in Run 2 and Run 3 is shown in Figure 2.2.

Because of the large number of protons per bunch, multiple proton–proton interactions in the same bunch crossing are possible. This process is called *pile-up*⁴. It gives rise to complicated detector signals due to the superposition of collision products from multiple sources. Disentangling the original collision events from each other is therefore essential for later data analysis. Additionally, so-called out-of-time pile-up from particle signature overlaps of different bunch crossings in the detector is equally possible. Pile-up distributions in ATLAS for the three Runs to date are shown in Figure 2.3. Meanwhile, a table summarising LHC conditions in the different Runs is provided in Table 2.1.

⁴ Denoted by μ for the number of pp collisions occurring in a bunch crossing.

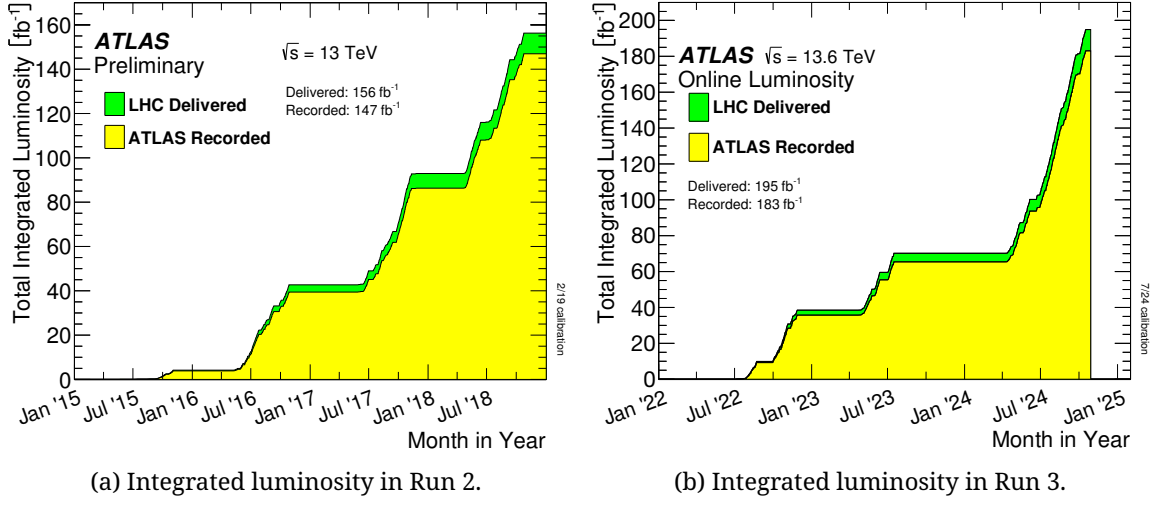


Figure 2.2: Integrated luminosity delivered by the LHC and collected by the ATLAS detector in pp collisions during (a) Run 2 at 13 TeV and (b) Run 3 at 13.6 TeV. The YETs are clearly discernible from the luminosity plateaus. The lower amount of collected data originates from both inefficiencies in the data acquisition and from the detector start-up period following the declaration of stable beams by the LHC. The plots are taken from References [244, 245].

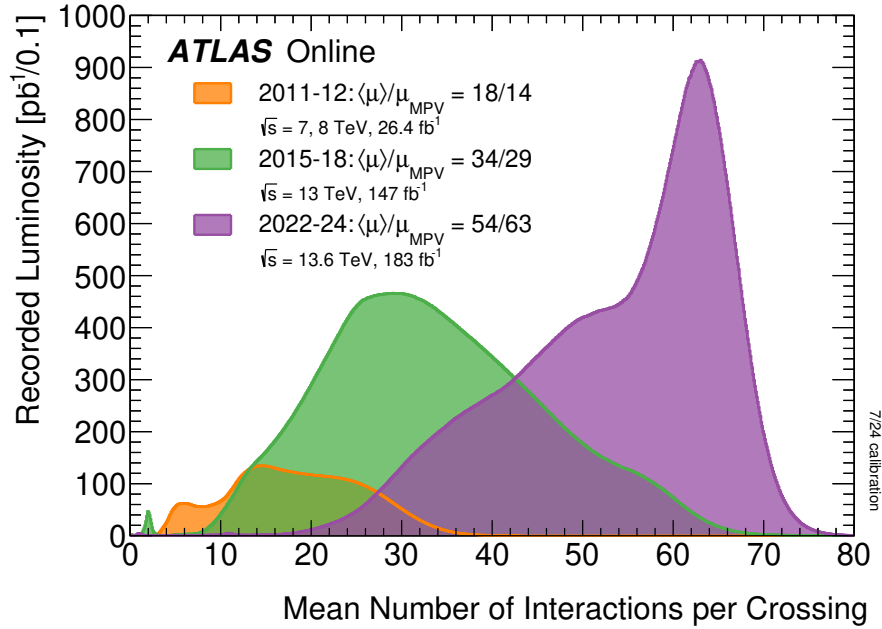


Figure 2.3: Distributions of pile-up conditions encountered by the ATLAS detector in pp collisions during the full LHC operations to date. Both the mean pile-up $\langle\mu\rangle$ and the most probable pile-up value μ_{MPV} are shown for each Run. The plot is taken from Reference [245].

Table 2.1: LHC operation conditions for pp collisions in each Run to date. For the number of bunches and the collision interval, the maximum values are taken. The integrated luminosity as recorded by the LHC is used. For Run 3, the design target values as of the Chamonix 2024 Workshop for the peak instantaneous luminosity and the bunch population are given. The remaining values for Run 3 are as of the end of data-taking for 2024. The parameter values are taken from References [237, 238, 244–247].

Parameter		Run 1 (7 TeV)	Run 1 (8 TeV)	Run 2	Run 3
CoM energy	[TeV]	7	8	13	13.6
Peak inst. lumi.	$[10^{34} \text{ cm}^{-2} \text{ s}^{-1}]$	0.35	0.77	2.1	2.3
Integrated lumi.	$[\text{fb}^{-1}]$	5.6	23.3	160.5	195
Number of bunches		1380	1380	2556	2464
Protons per bunch	$[10^{11}]$	1.45	1.6	1.25	1.8
Collision Interval	[ns]	50	50	25	25
Mean pile-up		9.1	20.7	33.7	54

2.2. The ATLAS Detector

The ATLAS detector [226] is a cylindrical forward–backward symmetric multi-purpose detector with a near-complete coverage of the 4π solid angle around the LHC interaction point closest to the CERN Meyrin site. With a diameter of 25 m, a length of 44 m, and a weight of around 7000 tonnes – comparable to that of the Eiffel tower –, ATLAS is the largest particle detector at the LHC.

To extract the maximum physics information, ATLAS consists of multiple layers of subdetectors arranged cylindrically around the beam pipe in an onion-like⁵ fashion. Complementing the cylindrical layers in each subdetector are the so-called end-caps for additional coverage in the forward regions. A schematic is shown in Figure 2.4.

Each layer is specialised in detecting different physics signatures: closest to the beam pipe sits the Inner Detector described in Section 2.2.1. It is designed to provide precise tracking and momentum information. Positioned around the Inner Detector are electromagnetic and hadronic sampling calorimeters for particle energy determination, as detailed in Section 2.2.2. The outermost layer consists of the Muon Spectrometer outlined in Section 2.2.3. It provides improved tracking, momentum, and energy resolution for the muons, which are not stopped in the calorimeters.

A two-level trigger system is used in the ATLAS detector to reduce the immense amount of data collected to manageable levels. Additionally, ATLAS houses several forward detectors for luminosity measurements, event topology determination in heavy-ion collisions, and forward region physics. The ATLAS trigger system and luminosity determination are described in Sections 2.2.4 and 2.2.5, respectively.

⁵ Or possibly more like a leek, given the cylindrical structure.

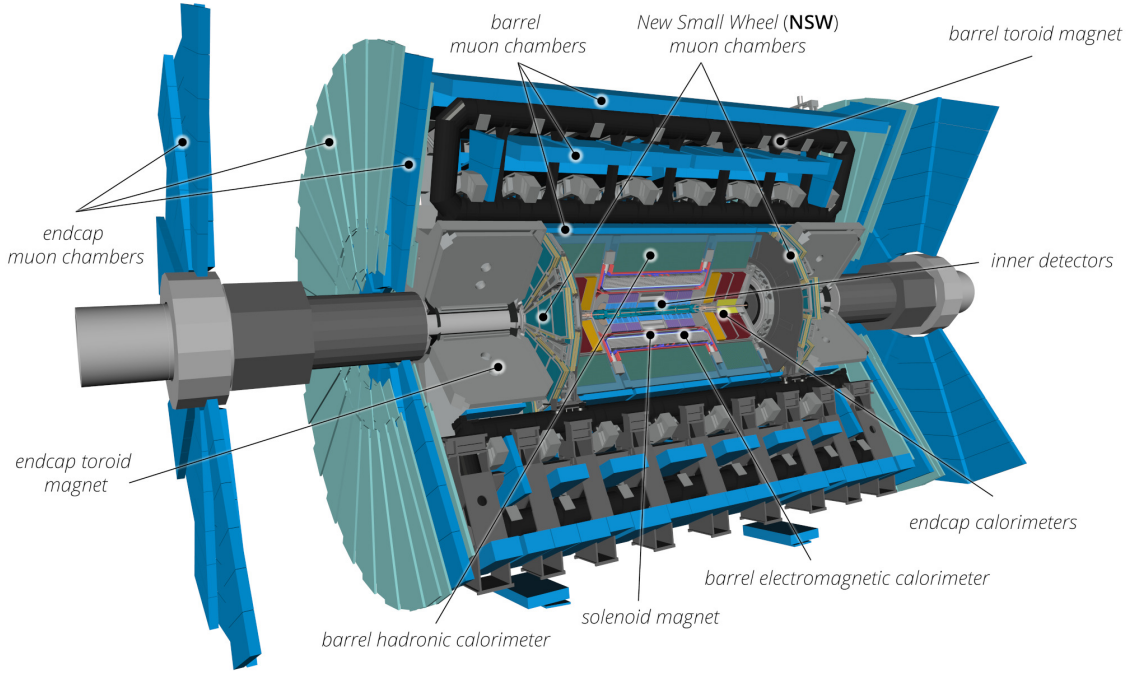


Figure 2.4: Schematic of the ATLAS detector in 2022. The overall detector structure in Run 1 and Run 2 was very similar, except for the exchange of the Small Wheel with the New Small Wheel [248] in Long Shutdown 2.

ATLAS Experiment © 2024 CERN (CC-BY-4.0 license).

The ATLAS detector employs a right-handed coordinate system with the nominal interaction point at its origin. The positive x -axis points towards the centre of the LHC, the positive y -axis upwards, and the z -axis along the beam pipe with the positive direction towards *Lac Léman*. A polar system is normally used for particle directions: the azimuthal angle ϕ is measured from the x -axis in the x - y plane, while the polar angle θ is measured from the beam axis. Additionally, the pseudorapidity η with

$$\eta = -\ln(\tan \theta/2)$$

is often used instead of θ since for highly relativistic particles it approaches the rapidity y with

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

where E denotes the particle energy, and p_z the momentum in the z -direction. As a result, differences $\Delta\eta$ in pseudorapidity can be taken as invariant under Lorentz-boosts along the z -axis, which is useful since the CoM system of the interacting partons does not necessarily correspond to that of the beams. Angular differences ΔR between objects are usually measured via the Euclidean distance in η - ϕ space as $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$. Lastly, transverse quantities such as the transverse momentum p_T or missing transverse energy E_T^{miss} are defined in the x - y plane, since only momentum components in this plane can be measured directly by the Inner Detector.

2.2.1. The Inner Detector

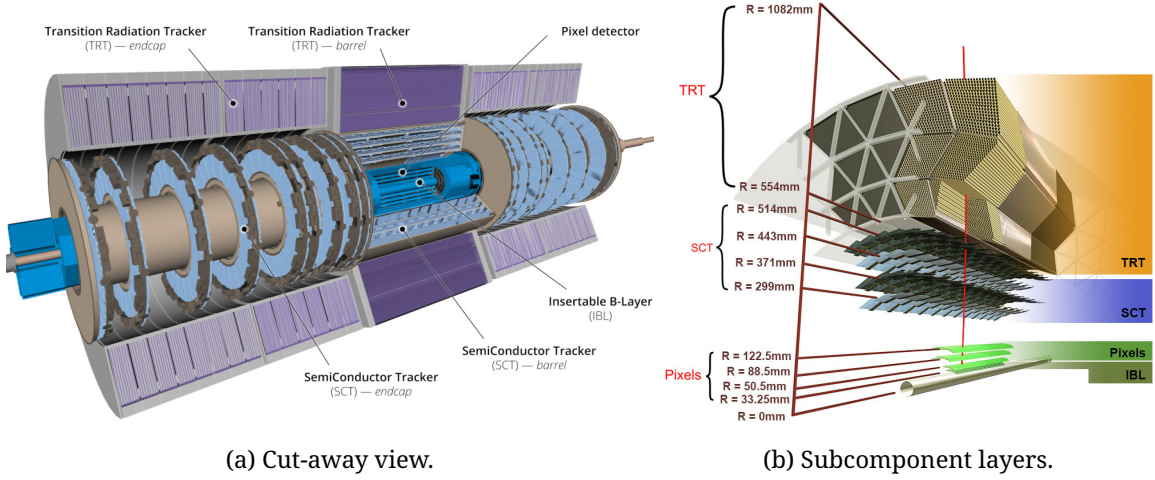


Figure 2.5: Illustrations of (a) a cut-away view of the ATLAS Inner Detector with barrel and end-caps and (b) the subsystems Pixel (with IBL), SCT, and TRT in the barrel region with their respective distances from the interaction point.

ATLAS Experiment © 2024 CERN (CC-BY-4.0 license), modified.

The ATLAS Inner Detector (ID) [249, 250] provides particle hit information required for particle tracking, vertexing, and transverse momentum determination in a pseudorapidity range of $|\eta| \leq 2.5$. Based on physics considerations, the ID was originally designed among other criteria to reconstruct tracks with $p_T > 5$ GeV at better than 95 % efficiency and have a momentum resolution better than 30 % at $p_T = 500$ GeV and $|\eta| < 2.0$ [251]. At the start of Run 1, the ID had a transverse momentum resolution σ_{p_T} / p_T of [226]

$$\sigma_{p_T} / p_T = 0.05 \% \cdot p_T [\text{GeV}] \oplus 1 \% .$$

The ID, shown schematically in Figure 2.5, encompasses three subsystems based on different tracking technologies: the Pixel detector consists of silicon pixels, the Semiconductor Tracker of silicon semiconductor strips, and the Transition Radiation Tracker of many small straw tubes filled with a gas mixture. All three are described in more detail in the following.

To afford momentum measurements via particle deflections, the ID is immersed in a homogeneous 2 T magnetic field aligned with the z-axis generated by a thin superconducting solenoid of 5.3 m length and 2.5 m diameter.

The Pixel Detector

The Pixel detector [252, 253] forms the innermost part of the ATLAS tracking system, providing true three-dimensional particle hit information in the radiation-harsh environment right around the beam pipe. Originally, it was designed with three concentric barrel layers totalling 67.1 million read-out channels [254]: the *B*-Layer, Layer 1, and Layer 2 sit at radial distances from the interaction point of 50.5 mm, 88.5 mm, and 122.5 mm, respectively. They are complemented on either side by three end-cap disks at axial distances $|z|$ of 495 mm, 580 mm, and 650 mm from the nominal interaction point, which cover the forward region up to $|\eta| = 2.5$ and add a further 13.3 million read-out channels [254].

While the *B*-Layer had initially been conceived as an insertable layer for the initial low-luminosity phase of the LHC to be replaced thereafter [252], this proposal was already scrapped during later detector planning stages. Nevertheless, the *B*-Layer is not designed for the high Run 2 radiation doses, which reduce charge collection efficiency, and for the instantaneous luminosity ramp-up past the design value of the LHC listed in Table 2.1, challenging the *B*-Layer read-out bandwidth [253]. Consequently, the Insertable *B*-Layer (IBL) [253, 255] containing an additional 12 million read-out channels was added during Long Shutdown 1 as a new innermost layer at a radius of 33.25 mm. The *B*-Layer was left in place instead of attempting its removal from the detector. In addition to mitigating the projected rapid performance drop for the *B*-Layer, the IBL was conceived to improve vertexing and the identification of jets originating from *b*-hadrons.

The sensors of the original Pixel detector (referred to as PIXEL from now on) and of the IBL consist of pixels with sizes $(50 \times 400) \mu\text{m}^2$ and $(50 \times 250) \mu\text{m}^2$, respectively⁶. The sensors follow a hybrid design, where the read-out electronics are placed on a separate chip bump-bonded to the active silicon sensor volume, which consists of deep-depleted reverse-biased p–n junctions to allow for full depletion even after type inversion of the n-bulk due to radiation damage and subsequent annealing [254, 255]. Passing charged particles create electron-hole pairs, whose charge build-up is processed by the read-out chip to generate hit signals. The IBL sensors are designed for higher radiation tolerance and use an updated read-out chip design for an improved bandwidth compared to the PIXEL sensors. The Pixel read-out system, including off-detector components, is described in more detail in Chapter 3.

The segmentation into pixels results in intrinsic hit position accuracies of 10 μm (azimuthal) and 60 μm (axial) for hits in the IBL [253], 10 μm (azimuthal) in PIXEL and 115 μm (axial/radial) in barrel/end-cap PIXEL, respectively [226]. The high spatial resolution is vital in the reconstruction of vertices and the estimation of track impact parameters for pile-up suppression and *b*-jet identification. With the addition of the IBL, impact parameter estimation was further improved, especially for low- p_T tracks where multiple scattering in the ID dominates [253].

⁶ The shorter side extends in the azimuthal direction, the longer in the axial one.

The Semiconductor Tracker

Situated around the Pixel detector are four barrel layers of p-in-n silicone semiconductor strips [256] with a length of 64 mm and 80 μm pitch at radii of 29.9 cm, 37.1 cm, 44.3 cm, and 51.4 cm [226]. Nine additional end-cap disks on either side reach up to $|z| = 2.72$ m. Together, they form the Semiconductor Tracker (SCT) with 6.3 million read-out channels and coverage of four particle hits for tracks up to $|\eta| = 2.5$.

The SCT modules consist of two back-to-back sensors, which have their strips rotated against each other by a stereo-angle of 40 mrad in order to increase the hit position accuracy in the axial and radial directions for the barrel layers and end-cap discs, respectively. As a result, the strips have an intrinsic accuracy of 17 μm in azimuthal and 580 μm in axial (radial) direction in the barrel (end-cap) region. The PIXEL and SCT detectors are bi-phase cooled with C_3F_8 to an operating temperature below 0 °C, while CO_2 is used in the IBL for the same purpose.

The Transition Radiation Tracker

The third subcomponent of the ID is the Transition Radiation Tracker (TRT). It employs a total of around 298 000 [257, 258] polyimide drift (straw) tubes of 4 mm diameter filled with a noble gas mixture⁷.

Similar to the Pixel detector and the SCT, the TRT consists of a barrel and two end-cap regions. In the central barrel region of $|\eta| < 1.0$ [257], about 53 000 straw tubes are aligned axially in up to 73 layers arranged in three module rings at radii of 55.4 cm to 108.2 cm. Meanwhile, the around 123 000 Drift tubes in each end-cap are aligned radially in 160 layers. Overall, passage through at least 36 drift tubes is ensured for tracks at $|\eta| < 2.0$, except in the transition region between barrel and end-caps [226].

At their centre, the straw tubes contain a gold-plated tungsten wire for collecting electrical signals generated during gas ionisation from transition X-ray photons. The transition radiation is generated via a fibre matrix (foils) of polypropylene between the straws in the barrel (end-cap) region [226]. Because of the dependence of the amount of emitted forward transition radiation on the Lorentz factor γ of a particle [261], the TRT is used for distinguishing electrons from pions [249]. Additionally, it assists in track reconstruction with an intrinsic azimuthal resolution of 130 μm for the individual straws due to their drift-time accuracy [259].

⁷ In Run 1, a Xe (70 %)/ CO_2 (27 %)/ O_2 (3 %) mixture was used [226]. Because of leaks and the high cost of Xe, from Run 2 onwards parts of the TRT were filled with a cheaper Ar (70 %)/ CO_2 (30 %) mixture already used for quality assurance during production [259, 260].

2.2.2. The Calorimeters

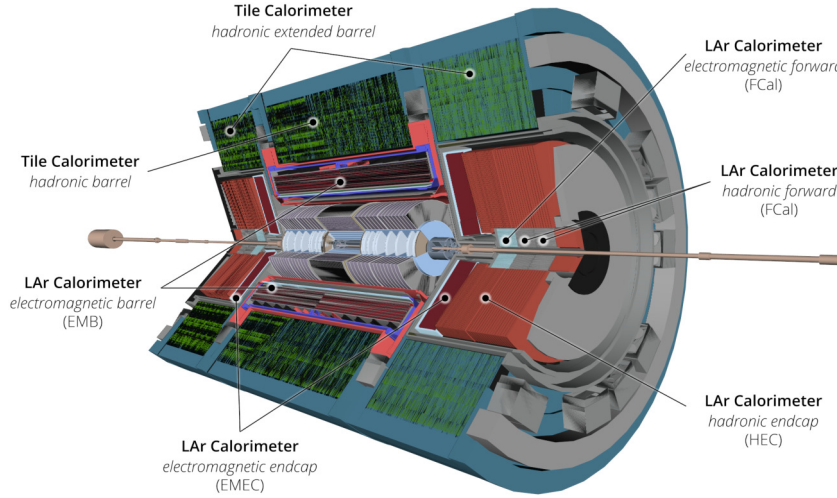


Figure 2.6: Schematic of the ATLAS electromagnetic and hadronic calorimeters.

ATLAS Experiment © 2024 CERN (CC-BY-4.0 license).

The ATLAS detector uses several sampling calorimeters [262, 263] with coverage up to $|\eta| = 4.9$, depicted in Figure 2.6, for precise energy measurements of electrons, photons, and hadronic jets produced in the particle collisions. To this purpose, the initial high-energy particles interact with absorber material to create showers of low-energy particles, which generate electrical signals in the active detector material while being fully absorbed. Material-dependent characteristic lengths leading to a particle energy reduction by a factor e are often used for shower parameterisation: for electromagnetic showers, this is the radiation length X_0 , for hadronic showers the interaction length λ .

Like the ID systems, the calorimeters consist of central barrel systems complemented by end-cap detectors in the forward regions. The inner subsystem is the Electromagnetic Calorimeter (ECAL) designed to measure the energy of electrons and photons. It consists of lead absorber plates interspersed with electrodes in liquid argon (LAr) cooled to about -184°C [262]. For homogeneous coverage and straightforward signal extraction, absorber and active detector volume follow an accordion geometry.

In the central barrel region at $|\eta| < 1.475$, the ECAL is split into up to three coaxial sections of decreasing η -granularity with increasing radii: the inner layer thus allows precise shower direction measurements and π^0 vetoing [264] using a granularity of about $\Delta\phi \times \Delta\eta = 0.1 \times 0.003$, the middle layer absorbs the bulk of the shower at a granularity informed by the Molière radius⁸, and the coarse outer layer contains the remaining energy. The end-caps ($1.375 < |\eta| < 3.2$) are structured similarly but coarser. An additional LAr presampling layer is included for $|\eta| < 1.8$ to correct for energy lost between the interaction point and the ECAL. Containment of electromag-

⁸ The characteristic lateral extent of an electromagnetic shower.

netic showers in the ECAL is ensured by its depth, which ranges from $22 X_0$ in the central region to $38 X_0$ in parts of the end-caps.

Around the ECAL, the Hadronic Calorimeter (HCAL) is used to measure the energy of hadronic jets. The Tile calorimeter with passive steel absorbers and polystyrene scintillator tiles as the active material is used in the low-radiation barrel region at $|\eta| < 1.7$. Low-energy particles generate UV photons in the tiles, which are grouped into three layers with a minimum granularity of $\Delta\phi \times \Delta\eta = 0.1 \times 0.1$ [226]. The photons are then guided by wavelength-shifting fibres to Photomultiplier Tubes (PMTs) for signal read-out. Overall, the Tile calorimeter consists of one central and two extended barrel segments at radii between 2.28 m to 4.25 m for a depth of 7.4λ ⁹.

In the forward regions of $1.5 < |\eta| < 3.2$, hadronic shower energy is measured by LAr sampling calorimeters with copper absorbers, and with copper/tungsten absorbers at $3.1 < |\eta| < 4.9$. The requirement of momentum balance in the transverse plane, combined with the precise accounting of particle energies, allows the determination of missing transverse energy E_T^{miss} from neutrinos and possible BSM signatures. In total, the calorimeters add a further 192 000 read-out channels [226].

2.2.3. The Muon Spectrometer

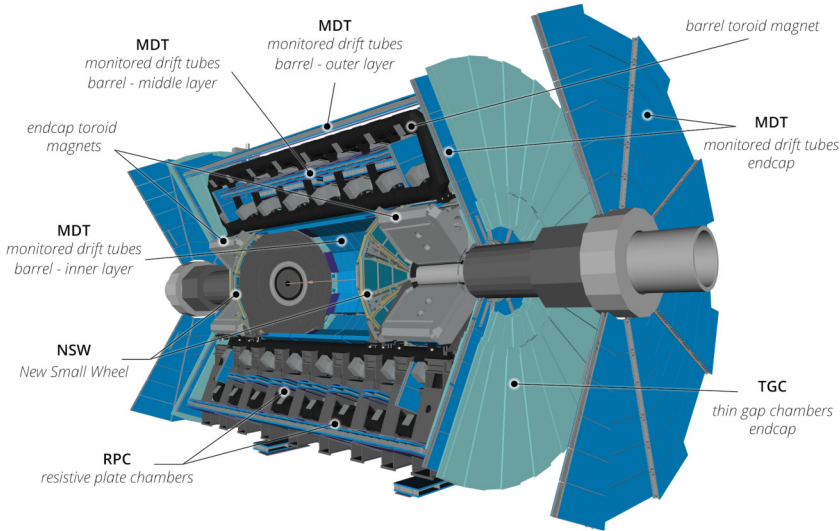


Figure 2.7: Schematic of the ATLAS Muon Spectrometer with an overview of the various detector technologies used.

ATLAS Experiment © 2024 CERN (CC-BY-4.0 license).

As high-energy muons are not stopped in the ATLAS calorimeters, the outermost detector layer consists of the Muon Spectrometer (MS) [265], shown in Figure 2.7, for improved muon momentum measurements and tracking. To this purpose, the muon system is embedded in three large toroidal magnets, each containing eight

⁹ Detector components in front of the HCAL contribute a further 2.3λ .

superconducting coils [266]: a non-uniform toroidal field of typically 0.5–1.0 T is generated by the air-core barrel toroid and the two end-cap toroids at $|\eta| < 1.4$ and $1.6 < |\eta| < 2.7$, respectively [265]. In the intermediate transition region, the fields of both magnet systems overlap.

The muon system itself consists of dedicated tracking layers (so-called *stations*) for precise muon hit detection. Additional low-latency sensors are used as inputs to the hardware triggers. In the barrel region at $|\eta| < 1.05$, three Monitored Drift Tube (MDT) stations are placed near the inner boundary, the maximum, and the outer boundary of the magnetic field for a measurement of the curved track sagitta in the bending direction. The MDTs consist of aluminium tubes of 30 mm diameter filled with an Ar (97 %)/CO₂ (3 %) gas mixture and a central tungsten-rhenium wire electrode [265] with a resolution of 60 μm for a single MDT [226].

In the end-cap region, no station placement is possible near the field maximum in the closed end-cap toroids. Instead, the outer stations arranged as Large Wheels are pushed axially outwards to provide measurements with a long lever arm after muon deflection. Because of the high particle fluxes at $2.0 < |\eta| < 2.7$ and the long MDT dead-time of around 750 ns, Cathode Strip Chambers (CSCs) – a type of multiwire proportional chamber – were used for the innermost station in this high-pseudorapidity region until the end of Run 2 [265]. The remaining end-cap stations consist of MDTs.

The trigger system covers the range of $|\eta| < 2.4$ via Resistive Plate Chambers (RPCs) with an intrinsic time resolution of 1.5 ns in the barrel, and Thin Gap Chambers (TGCs) with a resolution of 4 ns in the end-cap regions [226]. The trigger components additionally provide two-dimensional position measurements to assist in tracking.

During Long Shutdown 2, the Small Wheels containing the CSCs were replaced by the New Small Wheels (NSWs) [248] covering $1.3 < |\eta| < 2.7$. The new system employs small-strip TGCs for the hardware triggers, while Micromegas detectors are used for the precision momentum measurement.

2.2.4. The Trigger System

At a rate of 40 MHz, the LHC pp collisions would generate an immense amount of $\mathcal{O}(100 \text{ TB s}^{-1})$ [226] of ATLAS detector data, which is impossible to handle with current computing technology. At the same time, the vast majority of collision events are uninteresting for ATLAS physics analyses. For instance, ggF-initiated Higgs boson production with an inclusive cross-section of 48.5 pb at $\sqrt{s} = 13 \text{ TeV}$ [169] and $\mathcal{L}(t) = 2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ only occurs about once per second. Therefore, a two-level trigger system [268, 269] is employed to filter out interesting collision events¹⁰ and reduce the amount of data to be processed and stored to manageable levels.

The first stage hardware-based Level-1 (L1) trigger uses low-latency signals from the

¹⁰ There are additional minimum-bias event selections.

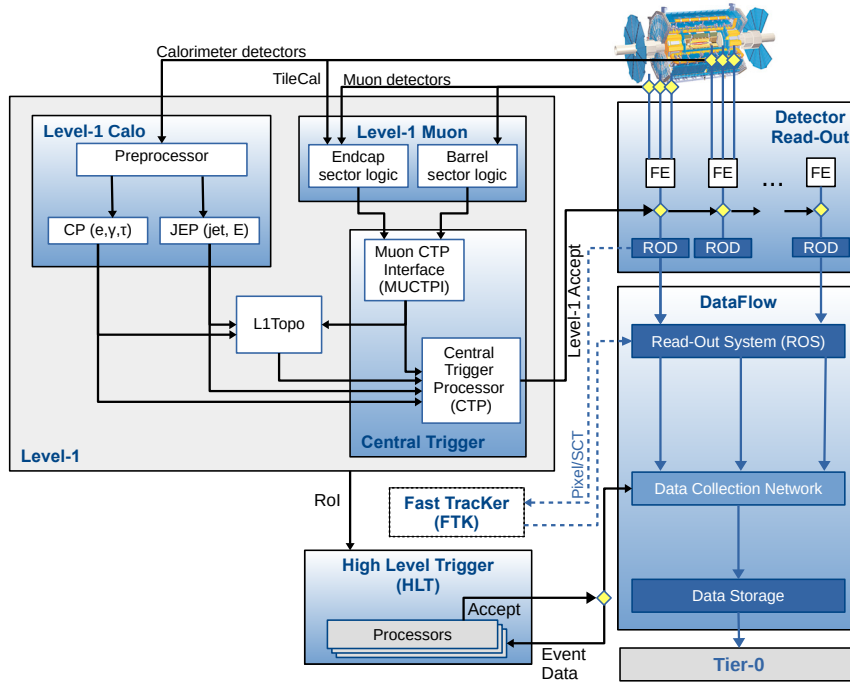


Figure 2.8: Schematic of the ATLAS two-level trigger system used during Run 2. The hardware-based Level-1 trigger relies on low-latency sensors in the calorimeters and the Muon Spectrometer. Meanwhile, the software-based High-Level Trigger makes use of the full detector read-out. The image is taken from Reference [267].

calorimeters and the MS. Afterwards, the software-based High-Level Trigger (HLT) consisting of about 40 000 compute nodes processes the events selected by the L1 trigger more thoroughly. The entire Run 2 system is shown in Figure 2.8.

Application-Specific Integrated Circuits (ASICs) and Field-Programmable Gate Arrays (FPGAs) are used in the L1 trigger for fast processing of the MS and calorimeter signals. The L1 calorimeter triggers react to photons, electrons, and tau leptons depositing energy in the ECAL and hadronic jets depositing energy in the HCAL on a granularity of 0.1×0.1 in $\eta \times \phi$. Meanwhile, the L1 muon triggers are designed to identify high- p_T muons via hit patterns in the MS stations. Triggers based on the jet scalar sum p_T , E_T^{miss} , and (from Run 2 onwards [270]) topological association between trigger objects are also used. In Long Shutdown 1, the L1 topological trigger subsystem was overhauled, while the L1 calorimeter system was upgraded in Long Shutdown 2 [270]. Decisions of the L1 subsystems are combined in the Central Trigger Processor (CTP) for a final trigger decision under a set of conditions defined in the so-called *trigger menus*. As a result, events are prefiltered at a rate of 100 kHz (75 kHz in Run 1) with a maximal latency of $2.5 \mu\text{s}$ [226, 267].

For events passing the L1 selection, the full event information is queried from detector module buffers and sent to the Read-Out System (ROS). The HLT then makes a final event selection based on both regions of interest defined by the L1 trigger for time-critical algorithms and the entire detector for the most time-consuming trigger

decisions. At the start of Run 2, the FPGA-based Fast Tracker (FTK) [271] was added as a pre-HLT component speeding up online track reconstruction for b -jet and tau lepton trigger decisions. In 2018, the HLT event processing time averaged 400 ms at peak luminosity [272]. After the HLT, the event rate is reduced down to 1200 Hz (400 Hz in Run 1) with a data throughput of 1.2 GB s^{-1} [267, 273].

2.2.5. Luminosity Determination

The ATLAS detector uses dedicated subsystems for instantaneous luminosity determination. In Run 1, this task was mainly fulfilled [274] by the Beam Conditions Monitor (BCM) [275], a set of diamond sensors primarily monitoring beam conditions to safeguard the ID. Meanwhile, the Luminosity Cherenkov Integrating Detector (LUCID), updated to LUCID-2 [276] during Long Shutdown 1, took over as the primary luminosity detector for Run 2 and Run 3. LUCID-2 consists (on either side of the interaction point) of 16 aluminium Cherenkov tubes located around the beam pipe at around $|\eta| = 5.6$ and at an axial distance of around 17 m from the detector centre. The tubes contain a quartz window as the Cherenkov medium coated with a ^{207}Bi source used for calibrating the PMTs, which amplify the Cherenkov signal.

LUCID-2 provides per-bunch luminosity estimates, which are also transmitted to the LHC as feedback on the beam conditions. Additionally, the instantaneous luminosity is averaged over so-called *luminosity blocks* (LBs) – periods about one minute in length, during which data-taking conditions are assumed to be near-constant. The LBs also form the smallest unit at which collision data is rejected in offline reconstruction if data quality issues are discovered.

The measurements of LUCID-2 are cross-checked with several other independent luminosity determinations from the BCM and track and cluster counting methods using the silicon detectors of the ID and the calorimeters [277]. The absolute luminosity measurement is calibrated via *van der Meer* scans [278] carried out yearly in special LHC fills. During these fills, the beam separation at the interaction point is slowly varied to obtain separation-dependent interaction rates μ_{vis} , which can then be used to determine a reference visible inelastic pp cross-section σ_{vis} . This allows for the measurement of the per-bunch luminosity $\mathcal{L}_b$ during physics data-taking via

$$\mathcal{L}_b = \frac{\mu_{\text{vis}} f_{\text{rev}}}{\sigma_{\text{vis}}},$$

where f_{rev} is the revolution frequency of the LHC and $\mathcal{L}(t) = \mathcal{L}_b \cdot n_b f_{\text{rev}}$ with the number of colliding bunches n_b per beam.

Part II.

Data Acquisition with the ATLAS Pixel Detector

The Pixel Detector Read-out Chain

While the sensors of the Pixel detector introduced in Section 2.2.1 are responsible for generating electrical hit signals from passing charged particles, these data subsequently have to be processed, transferred to the off-detector electronics located in the counting room, and finally routed to the central ROS as part of the Data Acquisition (DAQ) chain of the entire ATLAS detector. Additionally, the detector has to be calibrated periodically to ensure continued peak data-taking efficiency in the challenging environment close to the beam pipe. Both of these tasks rely on the Pixel detector read-out chain described in Section 3.1. Afterwards, detector calibration procedures are introduced in Section 3.2. Finally, Section 3.3 summarises Pixel detector operations in Run 3 with a focus on challenges and mitigation strategies. The development of the automated system tests for the Pixel read-out chain is instead deferred to Chapter 4.

3.1. Components of the Read-out Chain

3.1.1. From the Sensor to the Counting Room

The read-out chain of the original PIXEL detector [254], shown in Figure 3.1, and that of the IBL [255, 279] starts with the Front-End (FE) ASICs, which are connected to the sensors via bump bonds. The PIXEL and IBL FE chips are based on the FE-I3 [280] and FE-I4B [281, 282] designs, respectively. They amplify the electrical signal induced in the pixels by traversing high-energy charged particles and generate a digital hit signal if the amplified analogue pulse exceeds an adjustable threshold. As a result, the read-out circuits consist of an analogue part prior to the signal discriminator and a

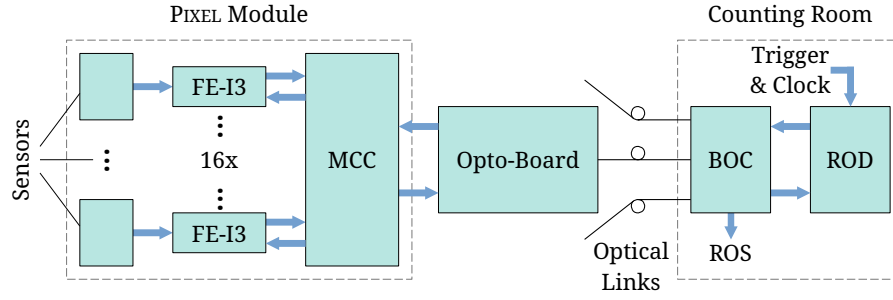


Figure 3.1: Simplified block diagram showing the read-out chain for a single PIXEL module. The IBL read-out chain is structured similarly, except for the module setup.

digital part following it. Both can be individually tested via dedicated circuits injecting a charge or a digital signal, respectively. Additionally, the Time-over-Threshold (ToT) of the discriminator in units of the 25 ns bunch crossing period can be used for an estimate of the charge on the sensor, since the discharge current is largely independent of the sensor charge and can be adjusted via dedicated registers in a similar way to the discriminator threshold. The timestamps of the leading and trailing edges of the amplified signals, which constitute the digital data, are then stored in buffers until an L1 trigger signal requests hits for a specific bunch crossing.

In PIXEL, 16 FEs¹ form a module controlled by a Module Controller Chip (MCC) ASIC [283], which coordinates the communication between the FEs on the module and the downstream components: it distributes configurations, trigger signals, and reset commands among the FEs, and aggregates the FE hit data before passing them downstream. Overall, 13 modules are arranged along z at constant ϕ on so-called *staves* in the barrel region. The FEs in the IBL are instead arranged in six double-chip modules at the stave centres, while four additional FEs connected to 3D² IBL sensors at the outer end of each IBL half-stave form single-chip modules. The IBL read-out chain also does not require an MCC. Instead, every IBL half-stave is connected directly to an opto-box via electrical lines of approximately 6 m length. The opto-boxes house opto-boards to interconvert the electrical signals running to and from the modules to optical signals running to and from the off-detector electronics in optical links. The MCC of the PIXEL modules are connected to the counting room via opto-boards as well.

¹ The sensors and some other components are technically also part of each module. They are not explicitly mentioned in the read-out chain for simplicity.

² The IBL pioneered the usage of 3D semiconductor sensors [284], in which the electrodes penetrate the wafer substrate, thus allowing independent control over sensor thickness and required depletion voltage.

3.1.2. Data Processing in the Counting Room

The ATLAS counting room is located in service cavern USA15 next to the detector cavern. Among the read-out systems for other subdetectors, it houses the Pixel off-detector electronics. These are primarily comprised of a number of Read-Out Driver Cards (RODs) [285] to process the detector data and control the FEs. Additionally, every ROD is connected to a Back-of-Crate Card (BOC) [286], which provides the optical interfaces to the detector modules, but also to the ROS and directly to the FTK. While a common BOC/ROD flavour had been developed for both the PIXEL and SCT subsystems, this original Hardware (HW) was gradually replaced with IBL-style BOC/RODs [255] in the YETS of Run 2 to benefit from bandwidth improvements in the optical links of Layer 1 and Layer 2, and to harmonise the off-detector part of the read-out chain for Firmware (FW) and Software (SW) development [287, 288]. As a result, the Pixel detector employs unified off-detector HW with different FW versions for the BOC/RODs connected to IBL and PIXEL FEs.

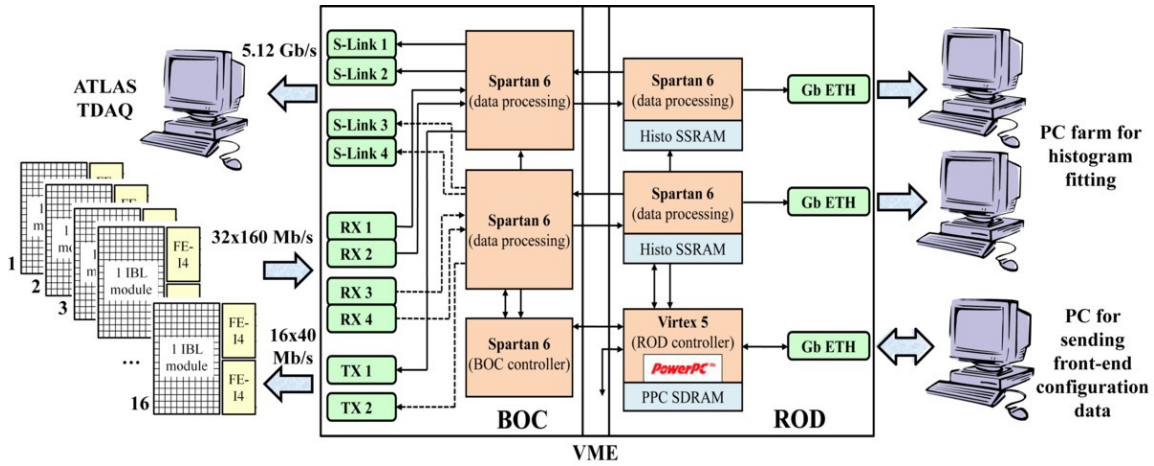


Figure 3.2: Schematic of the BOC/RODs used in the Pixel detector. While the BOC provides optical interfaces to the detector components and the ROS for the ROD, the ROD itself consists of three FPGAs to process data received from the detector modules during data-taking or calibration. Each Tx and Rx plugin can house twelve optical channels.

JINST **10** (2015) C03037 (CC-BY-3.0 license).

The overall BOC/ROD structure is shown in Figure 3.2. The BOCs are responsible for transmitting configurations, L1 trigger requests, and the LHC clock signal to the modules while receiving data in return. Additionally, they transmit the event fragments prepared by their connected ROD downstream to the ROS via S-LINK connections³, where the data are stored in Read-Out Buffers (ROBs) to await an HLT request. Internally, every BOC consists of one BOC Control FPGA (BCF) and two BOC Main FPGAs (BMFs) [286]. Each BMF in turn is connected to two S-LINK connections as well as two RX-plugins and one Tx-plugin (two in Layer 2). Each of the twelve

³ S-LINK is a data transmission format developed at CERN.

channels of an Rx-plugin (Tx-plugin) in turn operates at a bandwidth of 160 Mbit s^{-1} (40 Mbit s^{-1}) [255].

The RODs instead are responsible for building event fragments during data-taking and accumulating histogramming information during detector calibration scans, which is then transmitted via Gigabit Ethernet to a dedicated fit farm for further processing [279]. The ROD also receives and distributes trigger and timing information from dedicated trigger cards, and sends commands and configuration data to the FEs via the BOC. To fulfil these tasks, each ROD consists of one control FPGA with an embedded PowerPC (PPC) [289] as well as two further FPGAs for fragment building and histogramming. Lastly, the Program Reset Manager (PRM) FPGA is contained within the RODs as a low-level controller.

3.1.3. Emulating Detector Data in the Pixel Testbed

In general, each type of FPGA in use in the off-detector electronics uses its own dedicated Fw. As a result of the heterogeneous interconnected system, rigorous testing of system capabilities is required after changes in Fw or Sw. A dedicated testbed is available for this purpose at the CERN SR1 facility near the ATLAS detector. It consists of a full detector slice, meaning that the entire read-out chain described above is implemented on a smaller scale in a non-critical system to allow validation of Sw and Fw developments.

To complement the Pixel Hw available in the SR1 testbed, an emulator is available on the BOC to mimic real FE data in data-taking conditions and some calibration scans [286]. The off-detector part of the read-out chain can therefore be validated using the emulator without having to enable the upstream on-detector parts of the chain. Since high bias voltages, module cooling, and other complications are involved in operations with the real detector modules even in the testbed, emulated data present a low-risk alternative for many common read-out scenarios. Additionally, the emulator allows the generation of repeatable data patterns for some of the calibration scans outlined in the next section.

3.2. Calibrating the Pixel Detector

Frequent assessment of the Pixel detector operating characteristics has to be carried out to ensure optimal detector performance during data-taking. Therefore, the Pixel read-out chain has to support various scanning and tuning procedures [290, 291] in addition to the processing of particle hit data during physics data-taking runs. While a complete outline of all calibration procedures performed with the detector is beyond the scope of this thesis, some standard scans and tuning procedures will be introduced in the following.

3.2.1. Module Scans

Digital scans represent the simplest monitoring procedure in the Pixel detector: they consist of the injection of digital pulses into the FE read-out electronics located downstream of the threshold discriminator to validate the digital part of the FE read-out. In order not to overload the circuits accumulating the single-pixel hit signals, the digital pulses are strobed across the individual pixels so that only a fraction of them registers a digital hit at a time [292]. Additionally, the digital pulse is repeated several times to histogram the number of digital hits registered for each pixel with the RODs. Digital scans represent the first quick check to be performed during detector monitoring or tuning to ensure correct operation of the system.

Analogue scans are the next more complicated monitoring method. Instead of injecting a digital pulse into the digital pixel read-out circuits, they inject a known charge into the analogue read-out circuits before the amplification stage by means of a capacitor. Similar to the digital scans, these injections are strobed and repeated.

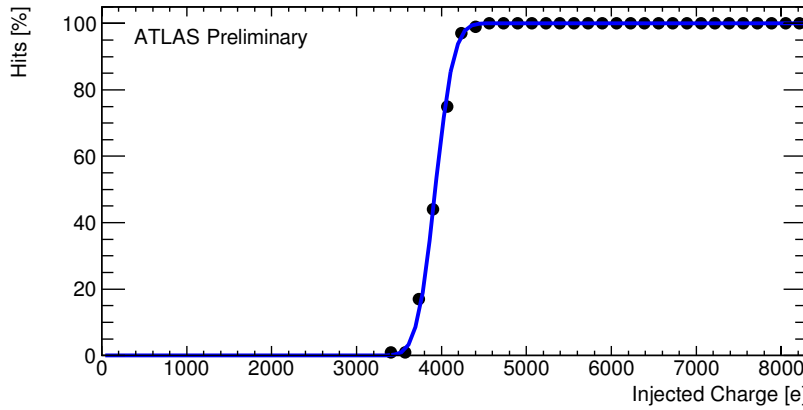


Figure 3.3: Example of an S-curve generated in a threshold scan. The average threshold and its root-mean-square deviation can be extracted in fits and compared to target values. ATL-INDET-PUB-2010-001 (CC-BY-4.0 license).

Threshold and ToT scans are the two other basic detector monitoring techniques and essentially consist of the repeated application of analogue scans with different injected charges. In threshold scans, the fraction of registered hits, also called the occupancy in this context, is scanned in dependence of the injected charge to arrive at so-called *S-curve* plots shown in Figure 3.3. Ideally, these plots have a step from no hits to a hit fraction of 100 % at the desired threshold. In practice, the average threshold and its uncertainty resulting from electronic noise and threshold spread in the case of multiple pixels can be extracted using a fit to a Gaussian error function. Meanwhile, ToT scans use the timestamps of the leading and trailing edges generated by pulses from analogue injections. Since the ToT is used for improving the hit resolution in clusters, where multiple pixels register a hit for the same traversing particle, a precise calibration of the ToT via adjustments of the pulse slope during the return to baseline, shown in Figure 3.4, is critical.

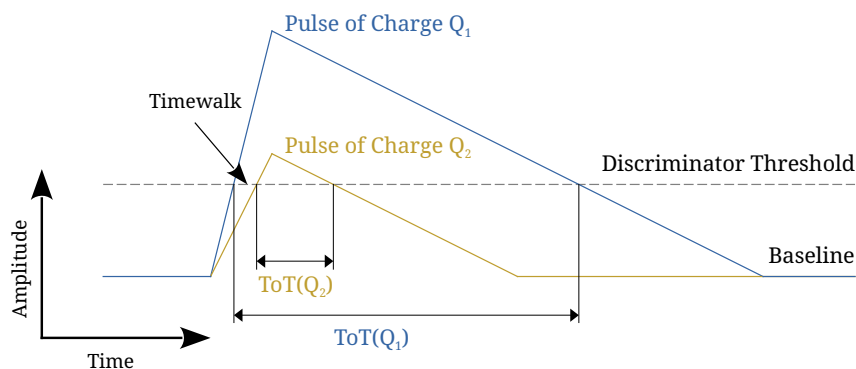


Figure 3.4: Schematic of the signals generated by two analogue signals of different charges $Q_1 > Q_2$. Due to the constant discharge rate, the pulses return to baseline with the same slope, resulting in different ToTs. Timewalk is visible between the leading edges.

3.2.2. Detector Tuning

Detector tuning is carried out if the threshold and ToT scans reveal deviations from the calibration target values or a large spread of values. In a similar fashion to the step from analogue scans to threshold and ToT scans, tunings consist of repeated threshold or ToT scans with intermittent tuning iterations: in threshold tunings, the FE-level GDAC and pixel-level TDAC registers⁴ [280, 293], which control the discriminator threshold, are iteratively adjusted to ensure a uniform response of the detector to injected charges. Meanwhile, ToT tunings aim at obtaining a uniform ToT response to injected charges of a certain magnitude by adjusting the FE-level IFDAC and pixel-level FDAC registers⁵, which modify the pulse slope during the return to baseline, resulting in a change of ToT.

3.2.3. Timing Scan

Since charge pulses of different magnitudes build up in approximately constant time in the analogue read-out circuits, the hit discrimination threshold is crossed at different times after particles traverse the sensor and generate charges – a phenomenon known as *timewalk* visible in Figure 3.4. As a result, signal pulses originating from small charges may not be registered in the bunch crossing bin corresponding to the collision, thus getting lost for the reconstruction, as only hits from a single bunch crossing are read out for each L1 trigger during data-taking⁶. Additionally, the overall trigger delay has to be determined to ensure detector synchronisation. Therefore, the initial coarse detector timing adjustment is performed in special ATLAS runs

⁴ These get their names from the fact that they are connected to ‘global’ and pixel-level Digital-to-Analogue Converters (DACs) to set the discriminator threshold.

⁵ Their names originate from the fact that these DACs control the feedback current I_f in the amplification stage before the hit discriminator circuits.

⁶ Reading out the surrounding bunch crossings would consume too much read-out bandwidth during physics data-taking conditions.

without LHC beam, in which the detector triggers on cosmic muons instead [294]. Afterwards, the detector timing is finely adjusted in special LHC fills containing only single isolated bunches by scanning clock delays in steps of 1 ns. Both of these adjustments use a special detector configuration in which multiple bunch crossings are read out for each L1 trigger. The optimal timing delay is determined as the point at which high-ToT⁷ hits begin migrating from the target bunch crossing to the one before it to minimise timewalk losses⁸.

3.3. Pixel Detector Operations in Run 3

During Run 3 *pp* physics fills, typical mean pile-up conditions of $\langle\mu\rangle = 52$ in 2022 and $\langle\mu\rangle = 62$ in 2023 [295] were about a factor of three higher than the design values of the PIXEL detector. The number of particles produced in collisions scales with the pile-up, thus leading to an increase in the pixel hit occupancies. The greater number of hits puts a strain on the read-out chain of the Pixel detector, where the read-out bandwidth of the *B*-Layer is the critical bottleneck: as the bandwidth usage exceeds the read-out limit, not all data can be transmitted off the modules in time. As a result, modules can become desynchronised from the L1 triggers until a subsequent Event Counter Reset (ECR) – a special signal issued by the CTP every five seconds to synchronise the entire ATLAS detector.

3.3.1. Smart Level-1 Trigger Forwarding

Module desynchronisation can also occur at a bandwidth usage below the critical data rate, for instance, if many L1 trigger signals are sent to a module registering many hits in a short time. Technically, the MCC is supposed to skip L1 triggers in this case to allow the modules sufficient time for building event fragments, while reporting the number of skipped triggers back to the ROD to compensate. However, discrepancies likely stemming from the MCC were observed in the expected and reported number of skipped triggers during Run 2 operations, resulting in ROD-level desynchronisation until the next ECR.

The on-detector components are inaccessible during LHC Runs and consist of ASICs without any programmable FW or SW components. Therefore, all mitigation strategies have to be implemented in the off-detector electronics via the FW of the BOC/ROD components. Consequently, the PIXEL-flavour ROD FW was modified during 2022 to carry out the L1 trigger skipping logic instead of relying on the MCC [297, 298]. In the new procedure, the ROD sends L1 triggers to the modules until the difference between sent triggers and module responses exceeds a certain threshold. At this

⁷ To avoid timewalk effects.

⁸ A safety margin of 5 ns is added on top.

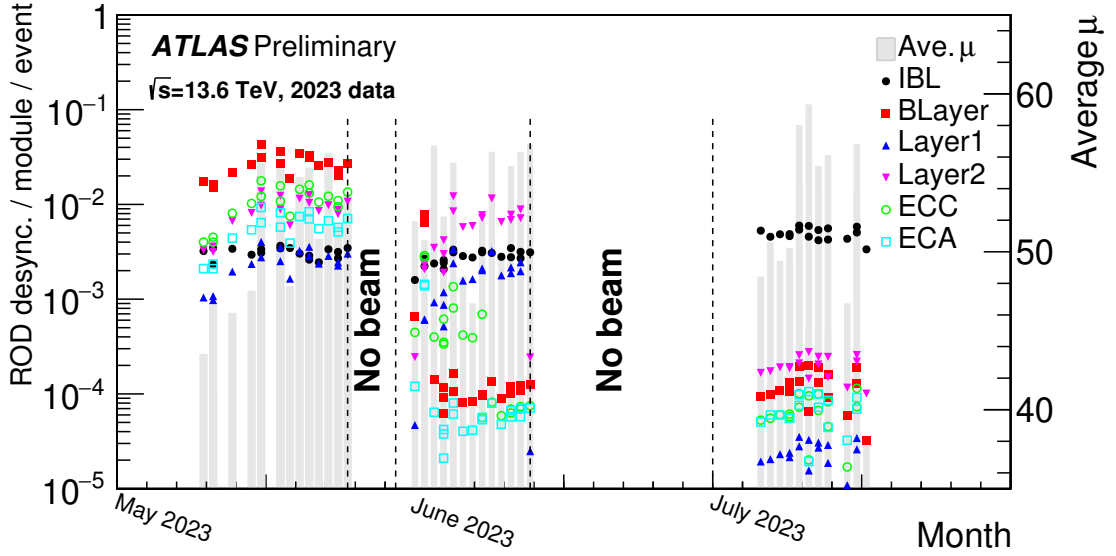


Figure 3.5: Semi-logarithmic plot showing the gradual reduction of the average per-event fraction of PIXEL modules with synchronisation errors at the ROD level during 2023 as a result of the rollout of smart L1 forwarding. The procedure was not implemented in IBL during 2023. Additionally, the pile-up averaged for each ATLAS run is shown. The plot is taken from Reference [296].

point, the ROD stops forwarding triggers and instead inserts empty dummy events flagged for later removal within the data quality processing into the data stream to keep synchronisation until further module responses are received and triggers can be forwarded again. This smart L1 forwarding was subsequently rolled out to all PIXEL layers during 2023. The resulting reduction in the module desynchronisation fraction is shown in Figure 3.5. Currently, work is ongoing to port the functionality to the IBL-flavour FW as well, where similar issues with skipped triggers have been observed.

3.3.2. Periodic Module Register Reconfiguration

Apart from large module hit occupancies, the high pile-up also presents challenges to the Pixel modules via a second phenomenon: secondary particles from hadronic interactions of the high-energy collision products also deposit charge in module registers and transmission lines, which can lead to bit-flips via so-called Single Event Upsets (SEUs) and Single Event Transients (SETs), respectively. The global registers of both the FE-I3 and FE-I4B chips are therefore implemented in radiation-hardened configurations [280, 282, 299]. However, some level of global deconfiguration from SEUs/SETs is still present and can lead to noisy FEs and sudden changes in module current consumption during data-taking [300]. Similarly, analogous effects can occur in single-pixel registers, leading to individual noisy and quiet pixels.

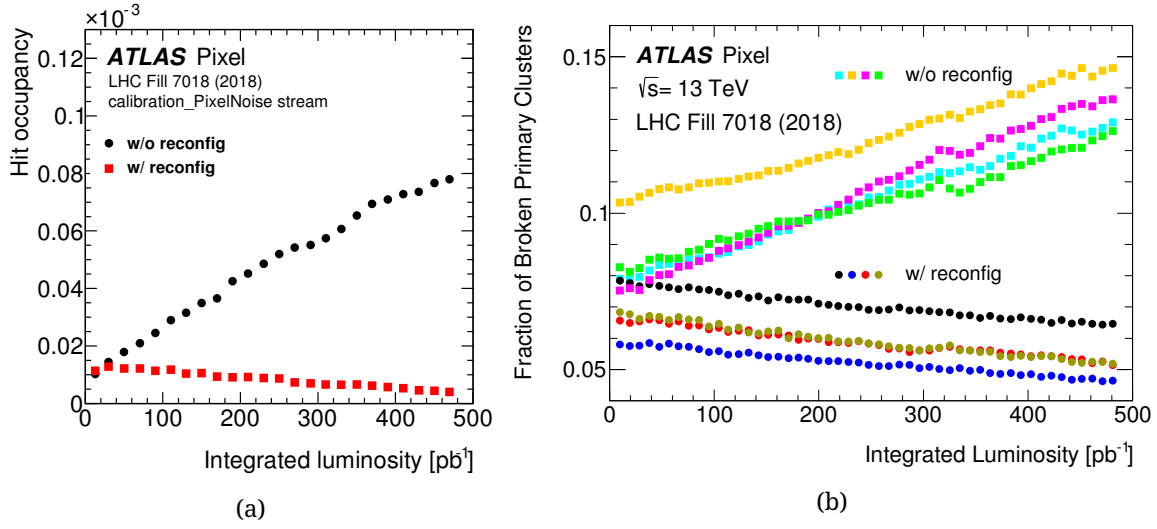


Figure 3.6: Plots of the development of (a) the hit occupancy and (b) the fraction of broken clusters in the IBL 3D modules during a fill towards the end of Run 2, in which the single-pixel register reconfiguration was tested. While the hit occupancy correlates with the amount of noisy pixels present, the fraction of broken clusters instead correlates with the amount of quiet pixels. The plots are taken from Reference [300].

In order to mitigate these module deconfiguration effects, a procedure to periodically reconfigure FE-level registers was rolled out to the IBL during Run 2 [300]. To avoid introducing additional dead-time during the reconfiguration step, the procedure makes use of a 2 ms period without triggers accompanying the ECR signal. During Run 3, the procedure was expanded to the single-pixel IBL registers, and FE- and MCC-level registers in the other Pixel layers [295, 301]. Due to the short ECR period, only a fraction of IBL pixels are reconfigured with every ECR on a rolling basis. A complete reconfiguration occurs every eleven minutes [300]. The effect of the single-pixel register reconfiguration on the development of noisy and quiet pixels was tested towards the end of Run 2 [300] via the hit occupancy in the IBL 3D modules and the observed fraction of broken clusters, respectively, as shown in Figure 3.6.

3.3.3. Thresholds & Bandwidth Limitations

Since the Pixel detector sits right next to the beam pipe, it experiences a constant high level of radiation from particle collisions during LHC operation. Over time, this leads to a decrease in the charge collection efficiency of the Pixel sensors because of charge trapping on radiation-induced defects in the silicon sensors. To retain tracking efficiency, the module hit thresholds can be lowered as a mitigation strategy. However, because of the simultaneous pile-up dependence of the bandwidth usage and the associated risk of module desynchronisation outlined previously, a balance has to be struck in the thresholds. Simultaneously, the constant radiation also leads to higher sensor bias voltage requirements and to drifts in detector thresholds and

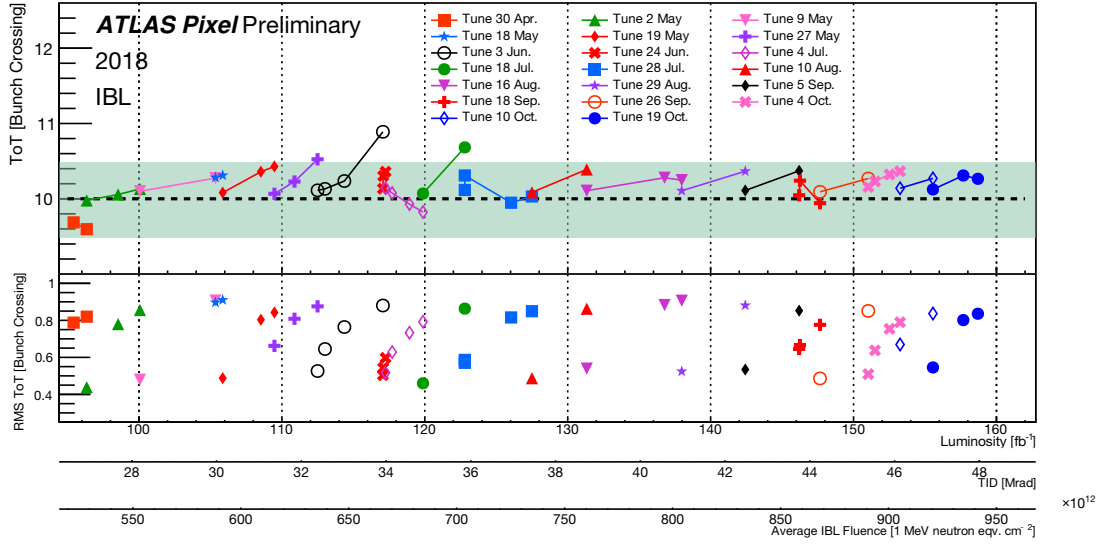


Figure 3.7: Evolution of the average IBL ToT values and their root-mean-square errors during 2018, where every retuning is denoted with a new marker style and colour. The plot is taken from Reference [302].

ToT as shown for the IBL during 2018 in Figure 3.7. Because of these shifts, frequent monitoring and retuning of the detector as outlined in Section 3.2 is required.

Based on data-taking experience collected in 2022, the *B*-Layer thresholds were raised at the beginning of 2023 [303]. Additionally, dedicated LHC fills in 2022 and 2023 were used to probe detector capabilities in various pile-up and trigger rate conditions for further Run 3 planning. Figure 3.8 shows the result of these condition scans on the IBL and *B*-Layer on-track hit efficiency⁹, suggesting that the smart L1 forwarding and the threshold optimisation introduced in between the two tests have led to an improved *B*-Layer performance.

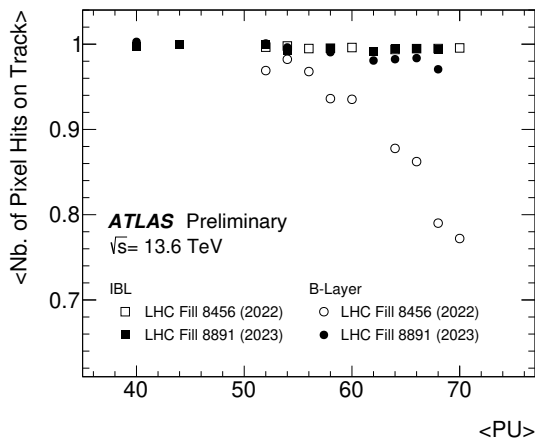


Figure 3.8: Hit-on-track efficiencies in IBL and *B*-Layer during 2022 and 2023 data-taking plotted in dependence on the average pile-up. The plot is taken from Reference [304].

⁹ The fraction of tracks reconstructed with $p_T > 0.7$ GeV with a hit registered in the respective layer.

Automated Testing Framework Development

The Run 3 improvements made to the Pixel detector outlined in Chapter 3 are primarily based on developments in the FW of the off-detector electronics. Additionally, the SW framework of the detector is continuously evolving to face emerging challenges and retain optimal detector performance. These FW and SW changes require a rigorous testing regime to ensure that no bugs are introduced to the read-out functionalities by new features.

While testing of the new functionalities is highly non-trivial and requires manual supervision, many tests required to validate changes in the read-out architecture are related to core detector operations, which are usually not modified by the new features. It was therefore decided during Long Shutdown 2 to develop automated tests covering basic data-taking and calibration procedures. The usage of the BOC-emulator introduced in Section 3.1.3 is critical in this context to avoid the need to monitor detector conditions due to the involvement of real detector modules.

As more functionalities were added to the RODs over time, the demands of the FW outgrew available HW capabilities. Therefore, separate FW versions are used for calibration and data-taking in the FPGAs in charge of histogramming and event fragment building. This split is reflected in the structure of the system testing framework, whose development is described in the following. Common components of the testing infrastructure are discussed in Section 4.1, while Sections 4.2 and 4.3 introduce the specific tests for the calibration and data-taking read-out chain, respectively.

4.1. Common Testing Infrastructure

In order to provide system tests in a standardised and repeatable format, the tests are implemented within the Pixel detector Continuous Integration (CI) pipelines – a software development practice centred around the merging of many small code changes into a central codebase whose continued correct operation is ensured at all times. The common `atlaspixeldaQ` Pixel DAQ SW framework [305] utilises CI for compilation, unit tests, and SW deployment to the read-only CERNVM File System (CVMFS) [306]. Similarly, CI pipelines are routinely used to synthesise the FW builds for the BOC and ROD FPGAs¹. Therefore, embedding the system tests as an additional stage in the current CI pipelines was deemed a natural extension of the existing workflows.

Within the GitLab infrastructure hosting the repositories, CI pipelines are realised via configuration files written in YAML and allow specifying, among others, job dependencies within the pipeline, artifacts to keep after job completion, and the shell scripts to run for each job. The pipelines themselves are operated via so-called *runners* and can be configured to run manually or automatically in specific circumstances or schedules. Additionally, environment variables for manual pipeline runs can be specified via the web-based GitLab Graphical User Interface (GUI), while default parameters can be defined for automatic pipeline runs. Since the machines capable of controlling the SR1 test system are not directly accessible from the web, the GitLab runners for the Pixel system tests connect to them via `ssh`.

In addition to the test implementation in the CI, several utilities are required for the system test infrastructure. To avoid code duplication when running the tests in the different SW and FW repositories of the Pixel detector, these utilities are contained within the `pixelDAQtest` [307] repository together with the testing framework itself. Currently, `pixelDAQtest` interfaces with these repositories through individual CI configurations, but it is planned to be fully modularised in the near future. The utilities themselves consist of the following components: a script-based alternative supplants the usual Integrated GUI (IGUI) application [308] of the central ATLAS Trigger and Data Acquisition (TDAQ) framework [309] used for the configuration of the Pixel detector. It controls the automated TDAQ startup before and shutdown after the system tests. Similarly, the so-called *detector tags*, which encode configuration settings such as the number of bunch crossings read out during data-taking, are loaded using a scripted tool instead of the GUI used during manual detector operation. The remaining utilities cover smaller general tasks. For instance, translations between ROD naming schemes are required for establishing `ssh` connections, and Python-based interfaces provide connectivity information between RODs and emulated modules.

¹ CI pipelines are also used to compile the ATLAS paper drafts automatically – as well as this thesis.

4.2. Calibration Tests

The calibration side of the system testing framework is concerned with validating basic operations related to running the calibration scans and tunings described in Section 3.2. To this end, two types of tests were introduced. The first of them uses a dedicated executable located in the DAQ SW framework to carry out scanning procedures without the GUI used during manual detector tuning. Configuration files are used to steer the behaviour of the scans, such as the type of scan to be performed or whether to use the BOC emulator in place of real modules.

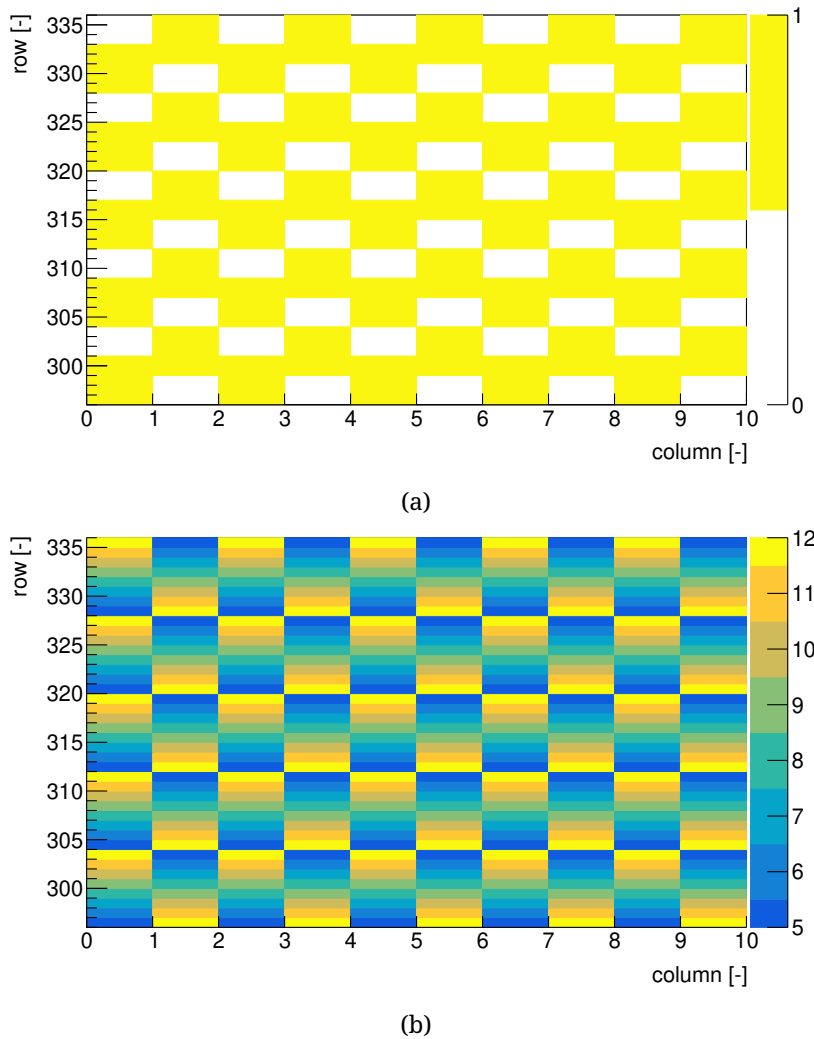


Figure 4.1: Plots showing (a) the occupancy and (b) the ToT of digital scans generated for parts of an FE-I4B with the augmented BOC emulator. A detailed explanation is contained in the text. The occupancy plot was produced for five out of eight mask steps.

The initial version of the BOC emulator only provided basic digital scans with random hits for FE-I3 emulation and a fixed pattern for FE-I4B emulation, so that no tests beyond checks on histogram existence could be carried out. Later, an improved

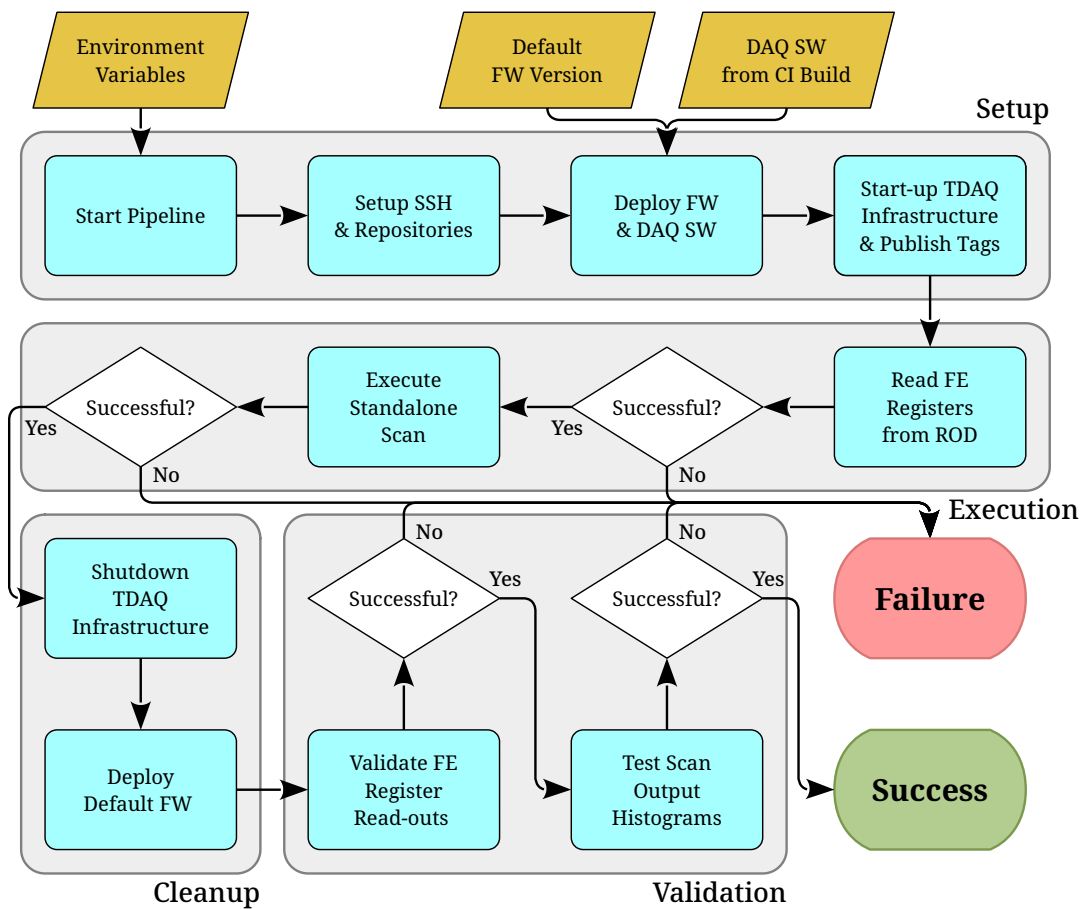


Figure 4.2: Flowchart of the calibration test pipeline grouped in stages. The test could alternatively be conducted using the default DAQ SW and a FW version from a CI build.

Boc emulator version also implemented the mask stepping present in real module scans via the strobing signal mentioned in Section 3.2, as shown in Figure 4.1. The number of mask steps and emulated injections are adjustable parameters of this augmented emulator version. Additionally, the mask steps are encoded via different ToT values in the corresponding histograms. Using the deterministic output of the improved emulation, additional tests validating the mask stepping via histogram contents were implemented in the `pixelDAQtest` repository.

The other kind of tests instead considers the FE register values stored on the RODs, which are loaded into the on-detector components during module configuration. By reading out the register values sent by the ROD during configuration and comparing them to references, the module configuration procedure can be tested on the ROD side. To validate these register config tests, several test bugs, which can be enabled using a command during compilation, were implemented in the ROD PPC code. They increment or decrement the FE-level GDAC by a value of one, and increment or decrement TDAC and FDAC of a random pixel, while flipping its enable bit². In tests, these

² The enable bit is used to mask single pixels.

bugs were consistently identified by the calibration system tests.

As the calibration tests were initially foreseen mainly in the context of the Pixel DAQ software, they are currently implemented primarily in the `atlaspixeldaq` CI. The corresponding system test CI pipeline is shown schematically in Figure 4.2. In general, each system test deploys the default Fw and Sw versions of all components except for the one under test. Therefore, for tests of the DAQ Sw framework, the standard Fw is loaded into the RODs used for the tests, while the Sw build under test is taken from the compiler stage of the CI pipeline. Meanwhile, for tests of the Fw, the synthesised Fw is taken from a previous CI stage, while the most recent `atlaspixeldaq` version on CVMFS is loaded. Since the augmented BOC emulator version is not yet available in the Pixel testbed BOCs, a simplified version of the automated calibration tests, which only carries out an emulated digital scan instead of the mask-checking procedure and FE register validation, is regularly running in the `atlaspixeldaq` CI pipelines at the moment.

4.3. Data-Taking Tests

While the calibration read-out chain is validated using one-time scans, the data-taking tests instead emulate data-taking conditions with simulated L1 triggers fired for a fixed duration at rates similar to those in ATLAS runs. Simultaneously, the BOC emulator returns data mimicking those generated in the modules during data-taking. In addition to the TDAQ instance taking care of detector configuration, a so-called DAQSlice instance for steering the data-taking is required in the tests making use of the simulated triggers. During manual data-taking tests with the Pixel detector, another IGUI application³ provides the control interfaces for the DAQSlice. For the automated tests, it is contained in a scripted final state machine instead.

The desired data-taking tests are outlined together with implementation details of a first version in a dedicated note written by the Pixel operations team during Long Shutdown 2 [310]. The tests are run at regular intervals during the simulated data-taking period, where fast and slow tests can be distinguished: fast tests can be executed while the simulated triggers are active. In contrast, slow tests require a temporary trigger hold to avoid race conditions in quantities such as the number of triggers fired in total. Regular checks on busy and timeout signals pertaining to the ROD under test primarily constitute the fast tests. They are therefore executed very frequently during the test duration by reading the corresponding variables out from the TDAQ information system. Meanwhile, the slow tests are further split up: they check that certain internal BOC and ROD registers agree with reference values, compare the amount of simulated fired L1 triggers with those sent in each Tx channel and the number of event fragments received back via the emulated Rx plugins,

³ Similar to the one used by the Run Control shifter in the ATLAS control room, but only containing the Pixel subdetector.

and compare the number of other commands sent over the individual Tx channels. Additionally, they verify that the number of event fragments received aligns with the number of triggers on channels connected to enabled modules, and perform sanity checks on the number of fragments corrupted for different reasons.

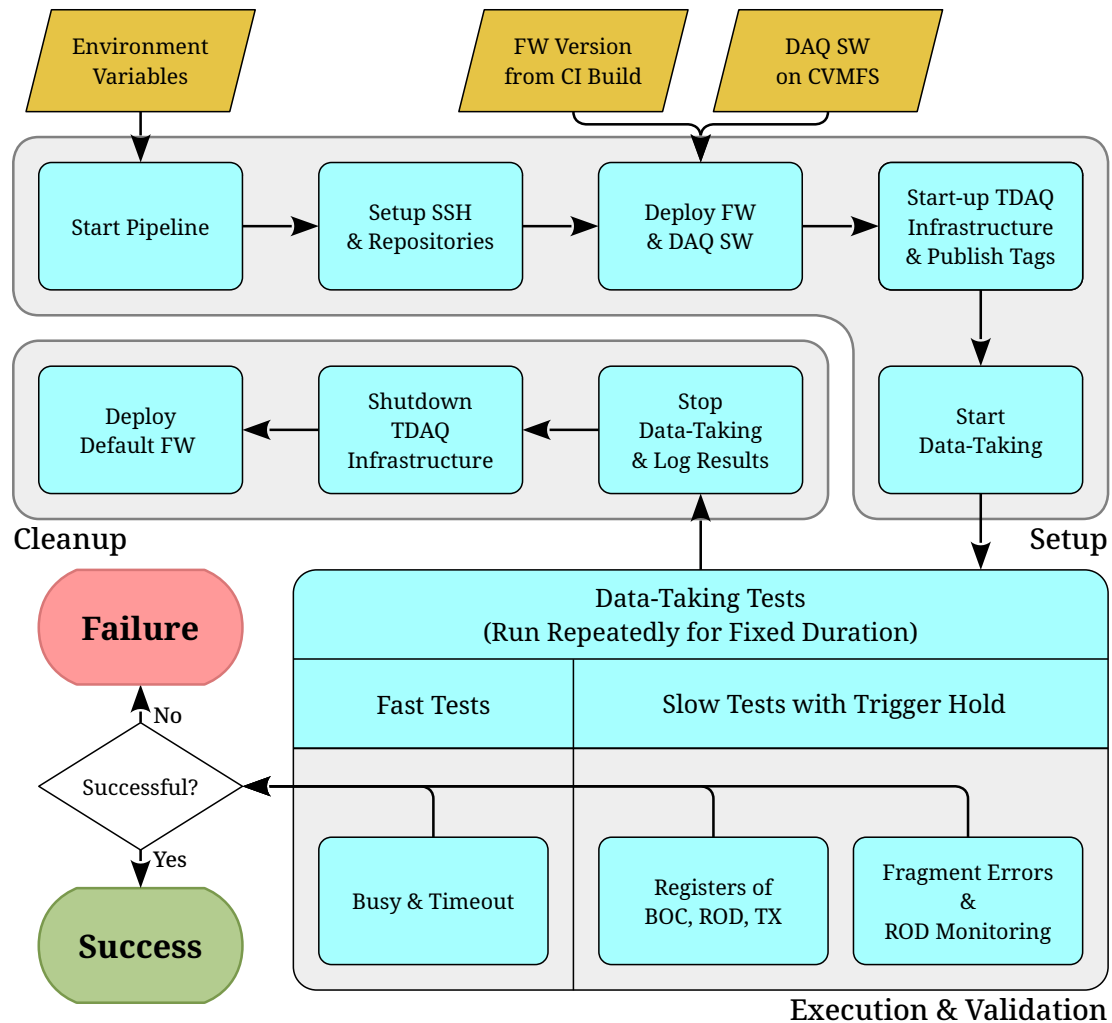


Figure 4.3: Flowchart of the data-taking test pipeline grouped in stages. The test could alternatively be conducted using the default ROD FW and a DAQ SW version from a CI build. While the fast tests can be run every couple of seconds during active triggers, the slow tests are run at a lower frequency during trigger holds.

A scheduler is responsible for executing the fast tests, trigger holds, slow tests, and the eventual shutdown of the data-taking infrastructure after test completion. Additionally, logging information and all data dumps used by the tests are stored for debugging purposes. A flowchart depicting the CI pipeline steps and the data-taking tests run is shown in Figure 4.3. As opposed to the calibration tests, the data-taking tests were mainly developed for the ROD FW repositories. Therefore, the corresponding pipeline using the default DAQ SW framework is depicted.

Part III.

The $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 Legacy Analysis

Machine Learning Fundamentals

An increasingly diverse set of machine learning methods is becoming available for performing data analysis in high-energy physics. As a result, established multivariate analysis techniques such as Boosted Decision Trees and kinematic likelihoods are fast being replaced by more sophisticated methods such as Recurrent Neural Networks, Graph Neural Networks, and Normalising Flows. The tool of choice in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis is the Transformer architecture developed initially in the context of natural language processing [311].

In the following, the Multilayer Perceptron (MLP) will be introduced as the basic building block of neural networks in Section 5.1. Afterwards, the concepts of back-propagation and gradient descent methods relevant to many machine learning models will be introduced in Section 5.2. Finally, the Transformer architecture will be discussed in detail in Section 5.3.

5.1. The Multilayer Perceptron

At the core of artificial neural networks, which are ubiquitous in machine learning applications, lie Multilayer Perceptrons (MLPs) consisting of multiple feed-forward layers. Based on the McCulloch-Pitts model [312], the *perceptron* [313] was implemented initially by Frank Rosenblatt to behave similarly to neurons in the brain. It consisted of an array of photocells connected to an analogue computer made up of potentiometers in order to learn to distinguish different simple geometric shapes.

Mathematically, a perceptron with a vector of d input variables x_j and a single output z can be described as

$$z = h \left(\sum_j^d w_j x_j \right), \quad (5.1)$$

where w_j is a vector of d adjustable weights and h a non-linear function called the *activation function*. While the original perceptron used a step function to emulate the binary neuron activation of the McCulloch-Pitts model, a continuous non-linear activation function is commonly used in modern machine learning to enable a continuous range of outputs. Additionally, this allows calculating a gradient across the function.

A fully-connected feed-forward layer, also called a linear layer, with m output nodes can be built out of m parallel perceptrons. With m additional bias terms b_i , this results in

$$z_i = h \left(\sum_j^d W_{ij} x_j + b_i \right),$$

where $\mathbf{W}$ is an $m \times d$ matrix with each row i made up of the vector $\mathbf{w}^{(i)}$ of d adjustable parameters for a single perceptron as defined in Equation (5.1).

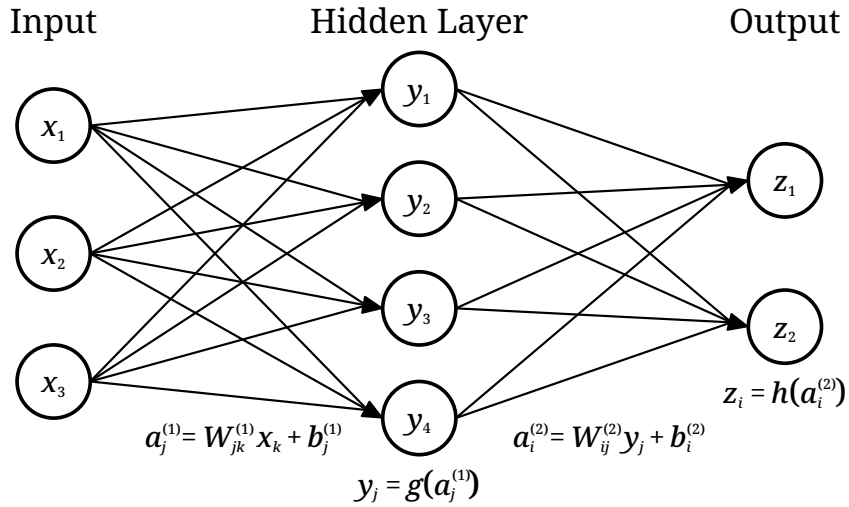


Figure 5.1: Schematic of a Multilayer Perceptron with a single hidden layer. The parameters used for feature calculations and the node contents are listed underneath. Activation functions are denoted by g and h . The superscripts (l) refer to the l -th layer counted from the input. For brevity, the Einstein summation convention is used on duplicate indices.

Using the output of such a feed-forward layer as the input of another one then generates an MLP. As the outputs of the internal layers of an MLP are usually not used directly, they are often referred to as *hidden layers*. Figure 5.1 illustrates an example of an MLP with a single hidden layer.

MLPs address a problem observed in perceptrons: while the training of perceptrons eventually converges for linearly separable problems, they are unable to approximate non-linear cases [314]. Meanwhile, MLPs with as few as one single hidden layer can approximate arbitrarily complex functions to an arbitrary precision [315]. In practice, the number of layers and nodes in each layer is optimised for each specific model during architecture development, making them part of the set of so-called *hyperparameters*. Non-linear activation functions in the hidden layers also play a crucial role: without them, multiple feed-forward layers could be represented by a single ‘effective’ layer.

5.2. Learning via Backpropagation and Gradient Descent

Having an architecture that is theoretically capable of approximating arbitrarily complex functions is not enough for practical machine learning applications. A method for optimising the model weights is also required to facilitate learning. In general, this is achieved by defining a loss function $L(\mathbf{z})$ dependent on the model outputs $\mathbf{z}$. The model then tries to minimise $L(\mathbf{z})$ during training via repeated weight updates. The focus in the following will be on supervised learning, where the model prediction target t , for instance, the geometric shape shown to a perceptron or the type of a particle collision event in an MC sample, is known during training.

In Rosenblatt’s original perceptron, the randomly initialised parameters are optimised by repeatedly comparing expected outputs with the prediction: In case of a mismatch between prediction and target, the weights of enabled inputs are increased if an output $z = 1$ is expected, and decreased if $z = 0$ is expected. While this procedure works for a single-layer network, it cannot be generalised easily to models with multiple layers. Instead, another approach called *backpropagation* (c.f. Reference [316]) is commonly used to optimise models with hidden layers. This method aims to minimise the error between the training target and the model prediction by using the gradients of the loss with respect to the model parameters. The parameters are then updated accordingly, and the procedure repeated until a minimum in the loss function is reached.

As a concrete example, backpropagation for an MLP with several outputs $z_i \in \mathbf{z}$ and a Mean-Squared Error (MSE) loss function,

$$L(\mathbf{z}) = \frac{1}{2} \sum_i (z_i - t_i)^2, \quad (5.2)$$

is shown in the following. Here, $t_i \in \mathbf{t}$ are the prediction targets of the individual output nodes.

Backpropagation operates by sequentially differentiating the loss function in Equation (5.2) with respect to the learnable model weights via the chain rule. Starting with

the differentiation up to a_i , the output of the last layer before applying activation, yields

$$\frac{\partial L}{\partial a_i} = \frac{\partial L}{\partial z_i} \frac{\partial z_i}{\partial a_i} = (z_i - t_i) \cdot h'(a_i) \quad (5.3)$$

in the notation of Figure 5.1, where h' is the derivative of the activation function h . Differentiating further allows the determination of the loss function gradients with respect to the model parameters of the last layer, W_{ij} and b_i , as

$$\frac{\partial L}{\partial W_{ij}} = \frac{\partial L}{\partial z_i} \frac{\partial z_i}{\partial a_i} \frac{\partial a_i}{\partial W_{ij}} = (z_i - t_i) \cdot h'(a_i) \cdot y_j, \quad \text{and} \quad (5.4)$$

$$\frac{\partial L}{\partial b_i} = \frac{\partial L}{\partial z_i} \frac{\partial z_i}{\partial a_i} \frac{\partial a_i}{\partial b_i} = (z_i - t_i) \cdot h'(a_i), \quad (5.5)$$

respectively. As can be seen in Equations (5.4) and (5.5), backpropagation directly makes use of the intermediate values calculated for the model prediction. Therefore, backpropagation can be easily carried out after running model predictions if these intermediate outputs are retained. To backpropagate into earlier hidden layers, terms in the loss gradients are summed together to carry through the chain rule fully. The derivative h' is the main reason for the usage of continuous differentiable activation functions.

After backpropagation, the parameters are updated via *gradient descent* to minimise the loss during model training successively. In its simplest form, gradient descent involves directly following the loss gradient for the updates of each model parameter w :

$$w^{(t)} = w^{(t-1)} - \lambda \frac{\partial L^{(t-1)}}{\partial w},$$

where the training step is indexed by t and λ is the *learning rate*, one of the most important hyperparameters of a model. The process is also called Stochastic Gradient Descent (SGD) because the loss L is usually calculated on so-called *mini-batches* of training data encompassing a limited number of data points for each backpropagation step to speed up training and allow the model to escape local loss minima. To enhance training performance, simple SGD can additionally be improved by introducing an exponentially decaying momentum term¹ [317–319] and by using adaptive learning rates implemented for instance in *AdaGrad* [320] and *RMSProp* [321]. Nowadays, the *Adam*² [322] optimiser is commonplace in machine learning. It combines RMSProp and momentum with improved bias corrections on the momenta used for the parameter updates.

Because backpropagation can be extended to determine loss gradients for weights in earlier layers, it enables the training of deep neural networks. In practice, models are trained after randomly initialising the weights in the following manner:

¹ This is analogous to a physical system with momentum and viscous drag, in which the negative gradient acts as a force.

² The name derives from ‘adaptive moments’.

5.2. Learning via Backpropagation and Gradient Descent

1. Carry out some model predictions by sampling from the input data distribution. This constitutes the *forward pass* of the information.
2. Use the intermediate results of the forward pass to calculate the gradient terms needed for backpropagation.
3. Backpropagate the loss to calculate the gradients with respect to the model parameters.
4. Update the model parameters via gradient descent using the backpropagated loss gradients.

These steps are then repeated until the loss minimisation procedure converges. Care has to be taken to ensure that the batched input data represents the overall population distribution. Otherwise, the trained model might give biased predictions.

In practice, training is also not carried out on the same data as used for inference with the converged model to avoid overfitting to the specific dataset used for training. Instead, the data are usually split into three separate sets:

- The *training* set is used to update the model weights during backpropagation.
- The *validation* set is used to validate the training and to select the optimal point of convergence, which avoids so-called *overtraining*, where the model becomes too specialised on patterns specific to the training set.
- The *test* set encompasses the data on which subsequent inference is carried out after the training has converged.

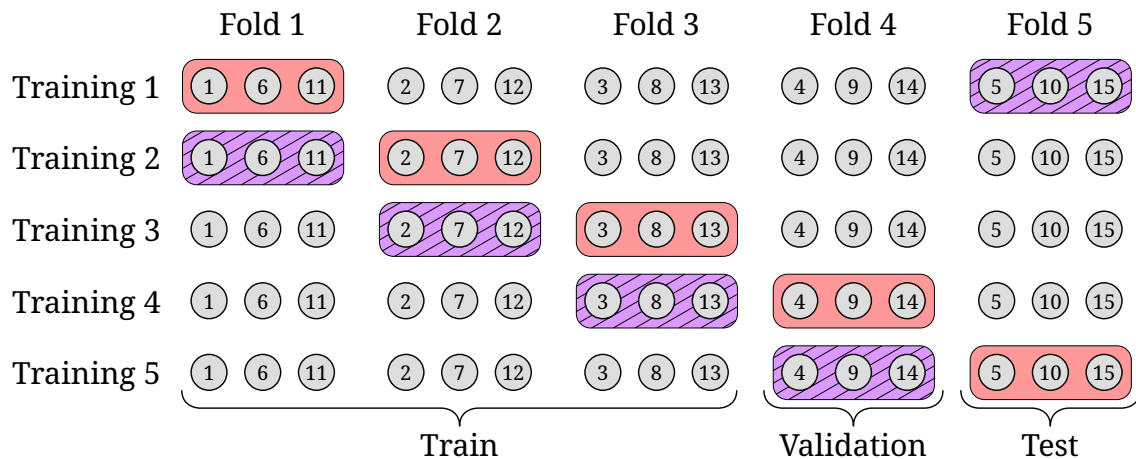


Figure 5.2: Schematic of a k -fold cross-validation setup with five folds, whose data is assigned based on an index modulo five. The folds used as validation and test sets, represented by the purple hatched and red boxes, respectively, are cyclically permuted between training iterations. The bottom brackets show the fold assignment for training iteration five.

One way to split the input data into training, validation, and testing sets without sacrificing a fraction of the input data for training is the so-called *k-fold cross-validation*: for each of the k folds, a separate network is trained and validated with $(k - 2)/k$ -th and $1/k$ -th of the input data, respectively, while inference is performed with a testing set consisting of the last $1/k$ -th of the input data. Between the training iterations, the assignment of the folds to train, validation, and test sets is cyclically permuted as shown in Figure 5.2.

The setup of the model output layer and the loss function $L(\mathbf{z})$ is dependent on the training objective: the model parameters are optimised such that the outputs $\mathbf{z}$ approach the true posterior probability distribution $p(\mathbf{t} | \mathbf{x})$ of the specific model task. So-called *canonical link functions* are then chosen as the accompanying loss functions such that the following holds for the loss function $L(\mathbf{z})$:

$$\frac{\partial L}{\partial a_i} = z_i - t_i. \quad (5.6)$$

In the following, setups relevant for the tasks in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis are outlined. In regression tasks, such as the direct prediction of the Higgs boson p_T , $p(\mathbf{t} | \mathbf{x})$ is assumed to have the form of a Gaussian distribution. Building the maximum log likelihood as the function to be maximised immediately yields the MSE loss function already used above in Equation (5.2). Therefore, $h(x) = x$ already leads to the optimal loss gradient of Equation (5.6) as seen in Equation (5.3). As a result, the output layer in regression tasks should not have any activation function and use an MSE loss function.

In multi-class classification problems with m classes C_i , such as particle collision event classification, Bayes' theorem gives the posterior probability distribution as

$$\begin{aligned} p(C_i | \mathbf{x}) &= \frac{p(\mathbf{x} | C_i) p(C_i)}{\sum_j^m p(\mathbf{x} | C_j) p(C_j)} \\ &= \frac{\exp a_i}{\sum_j^m \exp a_j}, \end{aligned} \quad (5.7)$$

where the ratio of exponentials is the so-called *softmax* function and where, for reasons outlined in the following, $a_i = \ln(p(\mathbf{x} | C_i) p(C_i))$ reuses the notation a_i for the output of the last layer before activation is applied. Simultaneously, the model outputs $\mathbf{z}$ and the one-hot encoded³ target vector $\mathbf{t}$ form the likelihood function as

$$p(\mathbf{t} | \mathbf{z}) = \prod_j^m z_j^{t_j}. \quad (5.8)$$

³ In one-hot encoding, the true target class i gets assigned a value of one and all other classes a value of zero. The encoding of class probabilities is also possible.

Taking the negative logarithm of Equation (5.8) to be minimised yields the *categorical cross-entropy* loss function,

$$L(\mathbf{z}) = - \sum_j^m t_j \ln z_j,$$

with a gradient of

$$\frac{\partial L}{\partial a_i} = - \sum_j^m \frac{t_j}{z_j} \frac{\partial z_j}{\partial a_i},$$

using $\mathbf{z} = h(\mathbf{a})$. This expression turns into Equation (5.6) if the softmax function from Equation (5.7) is used for the activation function h . Therefore, multi-class classification tasks usually use a softmax function for the output layer activation and a categorical cross-entropy loss⁴.

Multiple loss terms can be combined during model training. For instance, additional linear or quadratic losses on the absolute values of trainable parameters can be used in L1 and L2 regularisation, respectively. This procedure helps in mitigating training problems from so-called *vanishing gradients*, where potentially large model parameters project latent features into a flat region of the activation function. Additionally, regularisation tends to ‘simplify’ models in analogy to Occam’s razor: while important parameters persist, unimportant ones decay from the model, so that overfitting of the model decreases.

5.3. The Transformer Architecture

The neural network architecture used for event classification and Higgs boson p_T reconstruction in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis is somewhat more complicated than an MLP: it makes use of the Transformer architecture [311] initially developed in the context of natural language processing.

As opposed to the fixed input variables of an MLP, Transformers process inputs organised into objects, each containing a fixed set of variables. At the same time, they are agnostic to the number and the ordering of the input objects. These properties make Transformers very suitable to process any data that can be structured as a point cloud. In Section 7.1, the specific benefits of this approach for the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis are discussed.

In the Transformer, latent features are computed per object using feed-forward layers with weights shared between and optimised using all objects. Transformers additionally hinge on so-called *attention layers* [311] to facilitate pair-wise combina-

⁴ In binary classification tasks with only two classes, a single classification output y can equally be used together with a *binary cross-entropy* loss $L(y) = t \ln y + (1 - t) \ln(1 - y)$ and a *sigmoid* activation function $h(x) = (1 + \exp(-x))^{-1}$ reminiscent of a Fermi-Dirac distribution mirrored around the x -axis.

tions of the objects. Attention layers achieve this by performing a linear combination across the o objects for each of the d latent features. For this purpose, they initially obtain three tensors: the *query* $\mathbf{Q}$, the *key* $\mathbf{K}$, and the *value* $\mathbf{V}$. This is commonly done by applying single feed-forward layers without bias or activation to the $o \times d$ dimensional input feature tensor of the attention layer. Afterwards, query and key are matrix multiplied to obtain a score matrix $\mathbf{S}$:

$$\mathbf{S} = \mathbf{Q}\mathbf{K}^\top. \quad (5.9)$$

This score matrix then undergoes row-wise normalisation to yield a matrix of linear combination coefficients $\hat{\mathbf{S}}$, for instance by applying a softmax function. As a last step, this normalised score matrix is matrix multiplied with the value tensor to obtain the output $\mathbf{Z}$:

$$\mathbf{Z} = \hat{\mathbf{S}}\mathbf{V}. \quad (5.10)$$

Combining Equations (5.9) and (5.10) yields the overall action of the attention layer,

$$\mathbf{Z} = \text{softmax}_{\text{row}} \left(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{d}} \right) \mathbf{V}, \quad (5.11)$$

where the additional factor $\sqrt{d}$ was found to improve performance since it avoids low gradients across the softmax function in models with a large number d of latent features during training [311].

Two variations on the query tensor for attention layers exist:

- In *self-attention* layers, the query tensor is generated from the feature tensor using a feed-forward layer. The features of the individual objects then decide which parts of the key tensor are important for the linear object combination.
- In *cross-attention* layers, the query tensor is generated externally, for instance, as a learnable tensor. This allows the query tensor to have a different dimensionality with respect to the number of objects o . As a result, the object dimension of the output feature tensor can be modified in cross-attention layers, for instance, to perform pooling of the object parameters.

It is also possible to split the attention layer into n multiple heads [311]. The heads allow a single attention layer to process information from multiple lower-dimensional representations of the feature space simultaneously, instead of relying on a single high-dimensional representation. The low-dimensional outputs of the individual heads are concatenated after the attention layer for the output feature tensor. Schematics of the information flow in a multi-headed self-attention and a single-headed cross-attention layer are shown in Figures 5.3 and 5.4.

Overall, the attention mechanism allows the model to combine objects based on how well their features fit together. As a simple example of this, one could consider a

model that tries to find a Higgs boson decaying into two b -jets: the two jet input objects corresponding to these decay products will likely have high b -tagging scores in their input features. The attention layer could then select the b -tagging score as a query and as a key in one of the heads, resulting in high scores between objects with high b -tagging scores. As a result, the linear object combinations output by the attention layer will mainly consist of the b -tagged input objects.

In addition to attention layers, the Higgs boson p_T reconstruction Transformers use a *tensor attention* [323] layer (called *pairing layer* in the $t\bar{t}(H \rightarrow b\bar{b})$ analysis) to give each combination of two input jet objects i and k a pairing score S_{ik} . In its simplest form, this layer computes the inner products between the latent features of all pairs of the j jets. Using the $j \times d$ dimensional jet feature tensor $\mathbf{J}$,

$$\mathbf{S} = \mathbf{J}\mathbf{J}^\top. \quad (5.12)$$

For an unambiguous assignment using every jet only once, the score is further processed by masking its lower diagonal part and applying a softmax function over the unmasked elements to yield $\hat{\mathbf{S}}$ with elements $\hat{s}_{ik}$:

$$\hat{s}_{ik} = \text{softmax} \begin{cases} \text{mask} & \text{if } i \geq k, \\ s_{ik} & \text{if } i < k. \end{cases} \quad (5.13)$$

The highest score $\hat{s}_{ik}$ then yields the i -th and k -th jet as the most likely object combination. As $\mathbf{S}$ is symmetric by construction, no information is lost during the masking step. Tensor attention can be improved by using a tensor $\mathbf{T}$ with learnable parameters to compute a score matrix as

$$\mathbf{S} = \mathbf{J}\mathbf{T}\mathbf{J}^\top, \quad (5.14)$$

where the score is then processed analogously to its treatment in direct tensor attention via Equation (5.13). In the case of the two interchangeable b -jets originating from a Higgs boson, $T_{ik} \stackrel{!}{=} T_{ki}$ to encode the exchange symmetry in the pairing layer. In practice, this can be achieved by using a trainable tensor $\mathbf{T}'$ with $\mathbf{T} = \frac{1}{2} (\mathbf{T}' + \mathbf{T}'^\top)$. A schematic of the pairing layer with a trainable tensor is shown in Figure 5.5. The procedure can, in principle, be extended to an arbitrary number of jets as shown in Reference [323]. However, for the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis, the reconstruction of the two jets containing the Higgs boson decay products is considered sufficient for a differential cross-section measurement of the Higgs boson p_T in the STXS formalism.

In addition to MLPs and the attention mechanism, the Transformer architecture as outlined in Reference [311] also makes use of *skip connections* and *layer normalisation*. These will be discussed in more detail in the context of the specific architecture employed for the $t\bar{t}(H \rightarrow b\bar{b})$ analysis in Section 7.1.

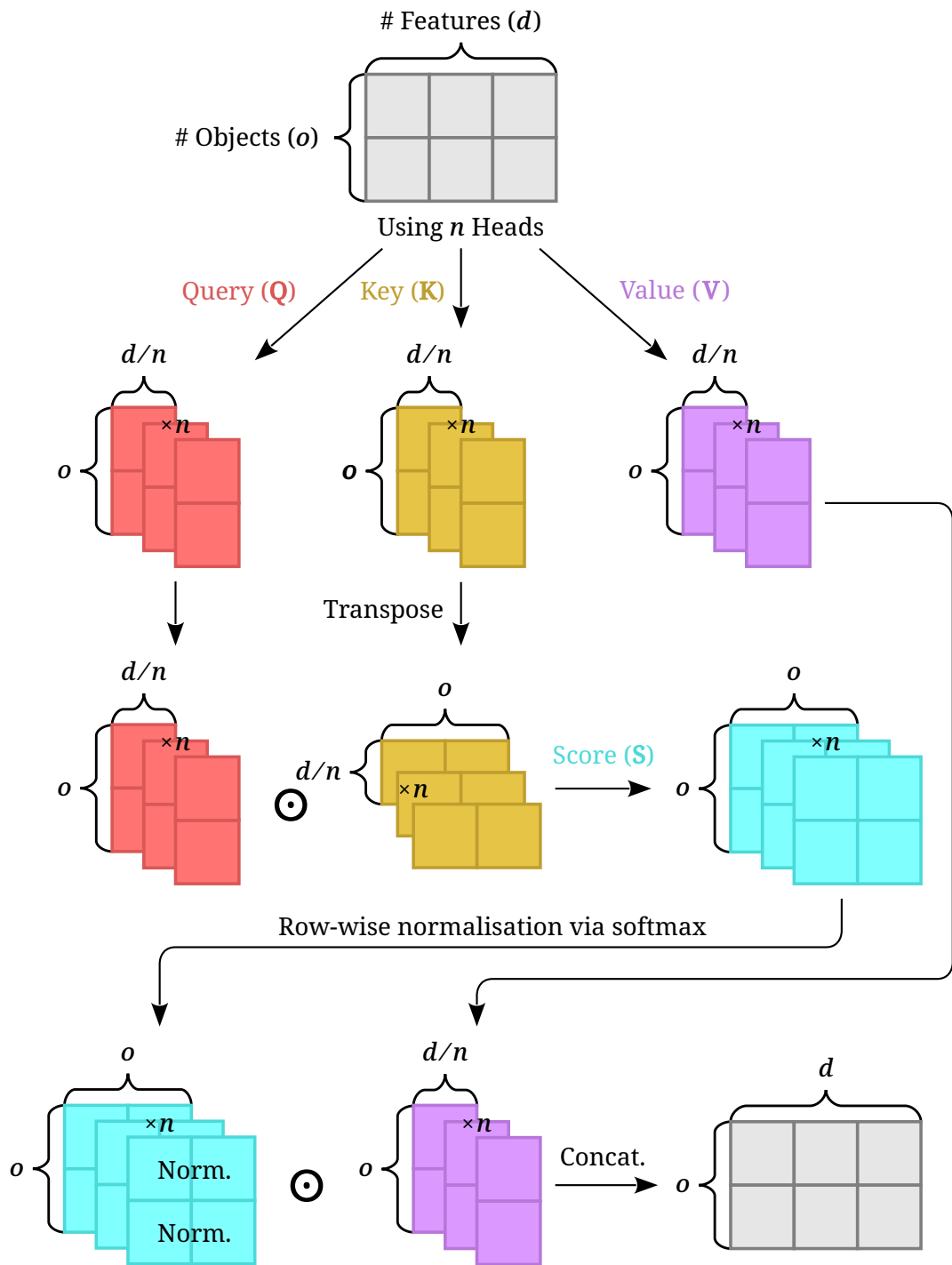


Figure 5.3: Schematic of the information flow in a multi-headed self-attention layer. The query, key, and value tensors are obtained from the $o \times d$ dimensional input feature tensor via three feed-forward layers. However, they are split up into n heads prior to carrying out the attention calculation (c.f. Equation (5.11)). After the calculation has been performed on each head, the n output feature tensors are concatenated along the feature dimension again. In principle, the individual heads could have a different number than d/n features each.

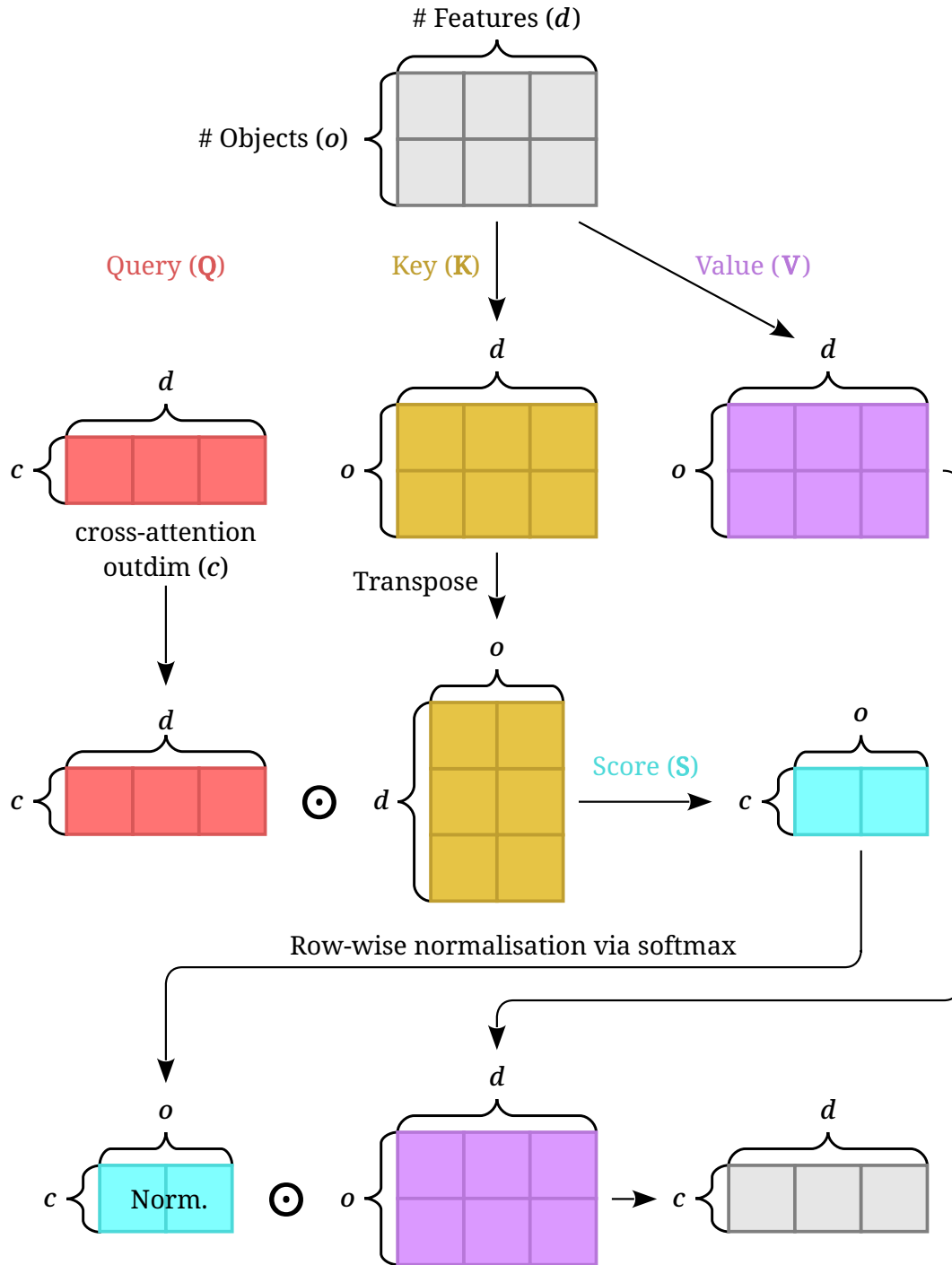


Figure 5.4: Schematic of the information flow in a cross-attention layer. The key and value tensors are obtained from the $o \times d$ dimensional input feature tensor via feed-forward layers. The query tensor is a $c \times d$ dimensional external input. Because of the different dimensionality of the query tensor, the output feature tensor will be $c \times d$ dimensional instead of $o \times d$ dimensional. The computation involving the three tensors follows Equation (5.11).

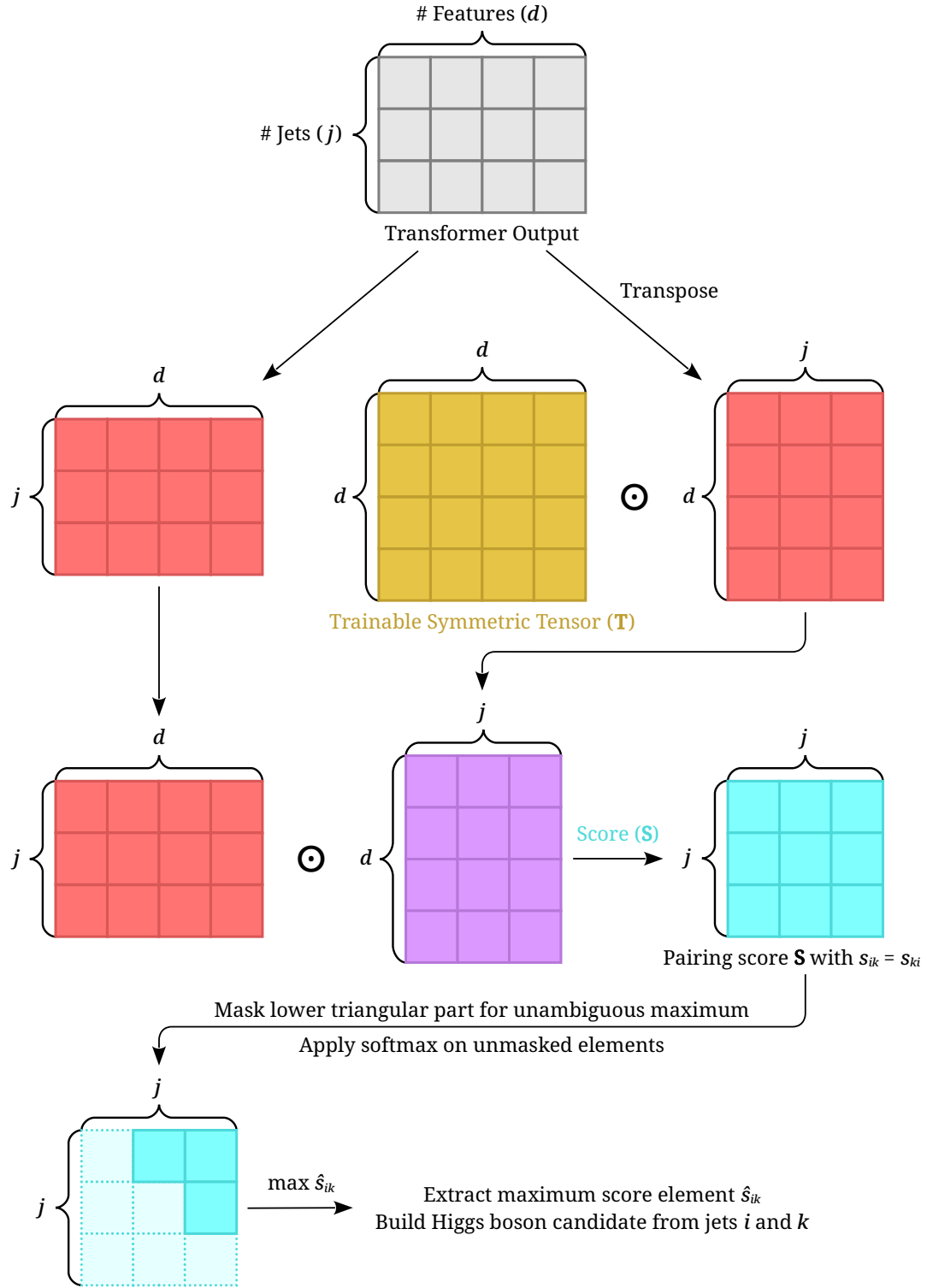


Figure 5.5: Schematic of the information flow in a pairing layer involving a symmetric tensor with trainable parameters. The computation follows Equations (5.13) and (5.14).

The $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 Legacy Analysis

After a general overview of the interplay between the top quark and the Higgs boson and previous $t\bar{t}H$ analyses has been provided in Section 1.2, the ATLAS $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis [1] will be outlined in the following. As a re-analysis of the full ATLAS Run 2 dataset, it supersedes a previous analysis published as Reference [211]. A discussion of the improvements carried out by the legacy analysis with respect to the first full Run 2 analysis is contained in Section 6.1. Afterwards, Section 6.2 presents the collider data and simulated Mc samples used by the analysis, while Section 6.3 outlines the object and event selection applied for the analysis. Section 6.4 introduces the likelihood fit setups used for the final cross-section extraction, but also for the development of the overall analysis strategy described in Section 6.5. A detailed discussion of the MVA techniques developed for the legacy analysis is deferred to Chapter 7, while the uncertainty model and the cross-section results of the legacy analysis are presented in Chapter 8.

6.1. Comparison with First Full Run 2 Analysis Setup

The $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis improves on the first full Run 2 analysis in several critical ways. First, the legacy analysis makes use of the consolidated ATLAS Run 2 object recommendations, leading to improved object definitions and associated uncertainties. In this context, the switch from the BDT-based Mv2c10 b -tagger to the RNN-based DL1r tagger is especially significant in the analysis phase space due to improvements of up to a factor of two in the light and c -jet rejection rates [324].

For event classification and reconstruction, the previous analysis utilised BDTs initially developed for a partial Run 2 analysis published as Reference [210]. These

operated mainly on high-level event features, such as invariant masses and angular separations between reconstructed objects. Meanwhile, the legacy analysis relies on Transformer networks, which use low-level kinematic and b -tagging object information as inputs.

The improvement in the MVA techniques allows for a looser kinematic event pre-selection in the legacy analysis and leads to a $t\bar{t}(H \rightarrow b\bar{b})$ event acceptance¹ of 6.3 %, corresponding to an increase of about a factor of three compared to the previous analysis. Furthermore, the definition of Signal Regions (SRs) and Control Regions (CRs) based on the Transformer event classification output in the legacy analysis presents an improvement in the discrimination of signal and backgrounds compared to the previous analysis, where regions were defined in terms of jet and b -jet multiplicities.

Additionally, a dedicated study on the modelling of the dominant $t\bar{t} + \geq 1b$ background [219], which represented the largest source of uncertainty in the first full Run 2 analysis, was performed in between the analyses to refine the $t\bar{t} + \geq 1b$ systematic model. Following Reference [219], the legacy analysis uses a consistent $t\bar{t} + \geq 1b$ model fully simulated in the 4FS, while the previous analysis employed a 4FS nominal $t\bar{t} + \geq 1b$ prediction and 5FS extrapolations for NLO matching, Ps generator, and $t\bar{t} + \geq 1b$ fraction uncertainties². With the improved Transformer-based event classification and increased event statistics, the $t\bar{t} + \geq 1b$ background could additionally be split into the three components introduced in Section 1.2.3 in the legacy analysis to provide better control over component-specific mis-modelling to the statistical fit. Simultaneously, the fixed $t\bar{t} + \geq 1c$ and $t\bar{t} + \text{light}$ cross-sections of the previous analysis are allowed to float for an in-situ nuisance parameter determination in the current analysis as well.

Both the first full Run 2 and the legacy analysis are set up for an inclusive $t\bar{t}H$ production measurement and a measurement differential in the Higgs boson transverse momentum p_T^H in line with stage 1.2 of the STxs formalism [168] outlined in Figure 1.16. However, while the previous analysis measured the $t\bar{t}H$ signal strengths, the legacy analysis is measuring cross-sections, which results in the cancellation of theoretical uncertainties in the $t\bar{t}H$ cross-section. The legacy analysis also conducts an inclusive signal strength measurement. Due to the improved event reconstruction provided by the Transformer networks used in the legacy analysis, the lower two STxs bins at $0 \leq p_T^H < 60$ GeV and $60 \text{ GeV} \leq p_T^H < 120$ GeV, which were combined in the previous analysis, can be measured individually. Therefore, the full STxs binning scheme can be realised in the legacy analysis.

¹ The event acceptance is defined here as the total number of events passing the analysis pre-selection divided by the total number expected from cross-section and luminosity.

² As a result of the single $t\bar{t} + \geq 1b$ cross-section nuisance parameter used, sub-component fractions for $t\bar{t} + 1b$ and $t\bar{t} + \geq 2b$ were required. No dedicated $t\bar{t} + 1B$ category was employed by the first full Run 2 analysis. Instead, the associated events are rolled into $t\bar{t} + 1b$.

6.2. Data and Simulation Samples

While experimental detector data from particle collisions is essential for the $t\bar{t}H$ cross-section measurement presented in this thesis, accurate simulations of significant SM processes also play an important role in refining the analysis strategy. Additionally, since the signal-over-background ratio in the analysis phase space does not allow a simple association of the detector data with the target process, MC simulations are required for a fair assessment of the data composition.

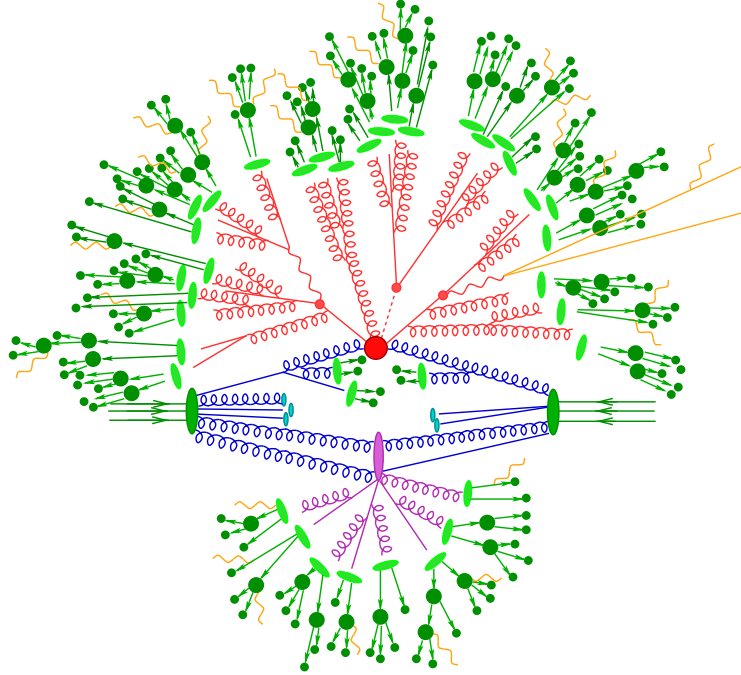


Figure 6.1: Schematic depiction of semileptonic $t\bar{t}(H \rightarrow b\bar{b})$ production in proton–proton collisions, as modelled by a Monte-Carlo generator. The hard interaction (large dark red blob in the middle) of two partons from the incoming protons (three parallel lines and blob on the left and right) is followed by the subsequent decays of the two top quarks and the Higgs boson (smaller red blobs above the hard scattering). A secondary interaction (elongated purple blob below the hard scattering) also takes place. Hard QCD radiation (gluon lines) is produced before the final-state partons hadronise (elongated light green blobs around the tree-like structure) and lead to further hadron decays (dark green round blobs on the outside). One of the W bosons decays into a lepton and a neutrino (long orange lines on the right). Some photon radiation (orange final-state boson lines) also occurs.

JHEP 12 (2024) 156 (CC-BY-4.0 license), modified.

Generally, the simulation of particle physics processes can be split into different parts relying on various approximations and assumptions: the hard scattering process at very high energy scales is modelled in so-called Matrix Element (ME) generators, which consider the sum of all perturbative Feynman diagrams to some order in perturbation theory. The subsequent Parton Shower (Ps) is modelled in a separate

step to approximate missing orders of perturbation theory by the emission of gluons and quarks from initial and final state particles at intermediate energy scales. Afterwards, hadronisation is modelled as the process of turning the asymptotically free high-energy quarks into colour-confined hadrons. For this stage, simulations commonly rely on either the cluster model [325, 326] or the Lund string model [327, 328]. In the former, gluons are forced to decay into $q\bar{q}$ pairs, which are subsequently collected into clusters decaying into hadrons. Meanwhile, the latter models the potential between quarks as colour strings, which subsequently break once the quark distance is sufficiently large and give rise to pair production, finally resulting in bound hadrons. In addition to the primary interaction giving rise to the hard scattering, secondary interactions from additional partons of the colliding protons to form the so-called *underlying event* are also possible. These are usually included in Mc simulations via phenomenological models with parameters matched to collider data due to the non-perturbative nature of the underlying event [329]. Furthermore, pile-up also needs to be included in Mc simulations for a complete approximation of the processes occurring in hadron colliders. Last but not least, the response of the ATLAS detector to the simulated collision processes also has to be modelled as accurately as possible. A depiction of a semileptonic $t\bar{t}(H \rightarrow b\bar{b})$ event with the different Mc simulation stages – excluding the detector response – is shown in Figure 6.1.

Although the same reconstruction algorithms are used to process both collision data and simulated Mc samples, the simulated samples additionally experience some degree of standard treatment beforehand, as described in the following. The Mc sample specifics for the processes considered in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis are detailed in Sections 6.2.2 to 6.2.4.

Nominal Mc samples in the analysis are generally produced using the full ATLAS detector simulation [330] based on the GEANT4 framework [331], except where stated otherwise. Since the detector simulation is the most computationally demanding part of Mc simulations, alternative samples are produced using the fast AFII [332, 333] simulation, which replaces the full simulation of the calorimeter response with a detailed parametrisation of shower shapes. No significant modelling differences between the full and fast simulations were observed.

Pile-up effects are included in the Mc samples via additional interactions generated in PYTHIA8.186 [334] with the NNPDF 2.3 LO [142] PDF set and the A3 Ps parameter tune [335]. These are then overlaid onto the dedicated process simulations. A reweighting of the simulated events in each data-taking period³ is performed to match the pile-up conditions present in the data.

EVTGEN1.6.0 [336] is used to simulate b - and c -hadron decays in samples showered via PYTHIA8 and HERWIG7 [337, 338]. PYTHIA8 additionally uses the A14 Ps parameter tune derived in studies with Run 1 data [339] and the NNPDF 2.3 LO PDF set.

³ The simulations used in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis are split into the mc16a, mc16d, and mc16e sub-campaigns to model data-taking conditions encountered in 2015–2016, 2017, and 2018, respectively.

The H7UE tune provided by the HERWIG authors [338] and the MMHT 2014 LO [340] PDF set are used for HERWIG7 showered samples. Samples generated via MADGRAPH5_AMC@NLO [341] use MADSPIN [342, 343] for the simulation of top quark and weak boson decays at LO to preserve spin correlations.

All simulations in the analysis use a top quark mass $m_t = 172.5$ GeV, a Higgs boson mass $m_H = 125$ GeV, and the corresponding Higgs boson branching ratios recommended by the LHC Higgs Cross Section Working Group in Reference [168]. Correspondingly, a bottom quark mass $m_b = 4.80$ GeV (4.50 GeV) is used for PYTHIA8 (HERWIG7) Higgs samples, while $m_b = 4.95$ GeV is used for b -quarks from the top decays. No fully hadronic $t\bar{t}$, $t\bar{t}b\bar{b}$, s - and t -channel single-top, and V + jets samples are used.

On the technical side, the raw detector data and Mc simulations stored in file formats of the ROOT [344] data analysis framework are processed into central Analysis Object Data (AOD) samples, which are further transformed into Derived Analysis Object Data (DAOD) samples for analysis usage by the individual physics and performance groups. TOPQ1 DAODs⁴ produced by the ATLAS Top Working Group are used as inputs to the two-stage TTHbbAnalysis [345] data processing framework in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis. The first stage herein consists of a modified event saver on top of version 21.2.238 of the AnalysisTop [346, 347] framework and is run on the Worldwide LHC Computing Grid (WLCG) [348], commonly referred to as the *grid*. Afterwards, the $O(40$ TB) output ntuples are processed locally on a batch system via the second stage to produce the $O(2.5$ TB) ntuple inputs to analysis studies and fits. The two-stage approach requires only a single run of compute-heavy algorithms and event skimming on the grid during the so-called L1 processing, while permitting fast offline sample re-processing during the second L2 stage.

6.2.1. Collision Data

Table 6.1: Amount of data recorded by ATLAS in each year of Run 2 compared to the amount of data remaining after applying data quality criteria [349]. The overall uncertainty of the combined luminosity amounts to $\pm 0.83\%$. The values are taken from Reference [277].

Year	Delivered by LHC [fb^{-1}]	Good for Physics [fb^{-1}]
2015	4.0	3.2
2016	39.0	33.4
2017	50.6	44.6
2018	63.8	58.8
Combined	157.4	140.1

⁴ Events are required to have at least one high-energy lepton to pass the TOPQ1 event skimming.

The data used by the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis encompass the full 140 fb^{-1} ATLAS Run 2 dataset collected in pp collisions between 2015 and 2018 at $\sqrt{s} = 13 \text{ TeV}$. The data quality is assessed for the individual detector subsystems and combined performance groups. As a result, data are designated unusable for physics analysis in case of unstable LHC running conditions, severe detector defects, or other issues on single lumi block granularity via the so-called Good Runs List (GRL) [349]. The amount of data collected by ATLAS and the amount remaining after applying the data quality criteria is shown in Table 6.1 for the individual years of Run 2. In total, the usable data amount to a combined luminosity of $140.1 \pm 1.2 \text{ fb}^{-1}$ [277].

6.2.2. $t\bar{t}H$ Signal Simulation

The hard scattering of the nominal $t\bar{t}H$ Mc samples is modelled in the 5Fs using the POWHEGBOX v2 [350–354] generator at NLO in QCD with the NNPDF 3.0 NLO [355] PDF set. The h_{damp} parameter, which regulates the p_T of the first high-energy gluon emission beyond LO in the Ps, is set to $3/4 \cdot (2m_t + m_H)$, while the geometric mean of the transverse masses⁵ $\sqrt[3]{\prod_i m_{T,i}}$, with $i \in \{t, \bar{t}, H\}$, is chosen as the functional form of the dynamic renormalisation and factorisation scales μ_R and μ_F , respectively. PYTHIA8.230 is used for showering.

Uncertainties due to Initial State Radiation (ISR) and Final State Radiation (FSR) are estimated via weights in the ME and Ps generators. The VAR3c up and down variations [339] of the A14 tune and the μ_R variations for QCD emissions by factors of two are used for ISR and FSR, respectively. Meanwhile, uncertainties from missing higher-order corrections are estimated via Ps weights by independently varying μ_R and μ_F by factors of two. Alternative samples showered with HERWIG7.0.4, but using the same hard scattering events, are used for a two-point systematic covering uncertainties in Ps and hadronisation. Similarly, the uncertainty in the NLO matching between the ME and Ps generator is assessed from events showered using a varied p_T^{hard} setting to change the strategy used to determine the emission veto scale in PYTHIA [356]. Uncertainties in the PDF set were found to be negligible.

6.2.3. $t\bar{t} + \text{jets}$ Background Simulation

The $t\bar{t} + \text{jets}$ background samples are split into two parts in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis: the $t\bar{t} + \geq 1b$ components are simulated through a dedicated 4Fs $t\bar{t}b\bar{b}$ sample, while the $t\bar{t} + \geq 1c$ and $t\bar{t} + \text{light}$ components rely on an inclusive 5Fs sample. As a result, $t\bar{t} + \geq 1b$ components in the inclusive sample have to be removed by a classification procedure outlined in Section 6.2.5 to avoid overlap leading to duplicate background contributions.

⁵ Defined for the Mc simulations described here as $m_T = \sqrt{m^2 + p_T^2}$.

The nominal 4FS $t\bar{t}b\bar{b}$ samples are generated with the POWHEGBOXRES [357, 358] generator and OPENLOOPS [359–361] at NLO in QCD using the NNPDF 3.1 NNLO 4FS [69] PDF set. The hard scattering events are interfaced to PYTHIA8.244 for showering. The QCD scales are set to $\mu_F = \frac{1}{2}(\sum_i m_{T,i})$ and $\mu_R = \frac{1}{2}\sqrt[4]{\prod_i m_{T,i}}$, where $i \in \{t, \bar{t}, b, \bar{b}, j\}$ and $i \in \{t, \bar{t}, b, \bar{b}\}$, respectively⁶. Meanwhile, $H_T = \sum_i m_{T,i}$, the scalar sum of transverse masses of final-state particles⁷ is used for $h_{\text{damp}} = H_T/2$ with the same index set as that of μ_R . Additionally, the so-called *bornzerodamp* parameter h_{bzd} , which regulates NLO emissions together with h_{damp} in POWHEG [358], is set to a value of 5.

Uncertainties in ISR, FSR, the QCD scales, and the NLO matching are evaluated in the same way as for the signal samples. An alternative sample showered with HERWIG7.1.6 is used to estimate Ps and hadronisation uncertainties. Additionally, an uncertainty on the recoil scheme used by PYTHIA ISR emissions is included: by default, a global recoil scheme, in which emissions rebound off the complete final state, is used by PYTHIA. This procedure is not well motivated for processes with colour charge flows beyond those of quarks between initial and final state [219], such as gluon-mediated $t\bar{t}b\bar{b}$ production. Therefore, an alternative sample using the dipole recoil scheme, in which initial-state emissions rebound off a single final-state particle, is produced as well⁸. The systematic model of $t\bar{t}b\bar{b}$ is motivated by dedicated modelling studies reported in Reference [219].

The inclusive 5FS $t\bar{t}$ samples are produced using POWHEGBOX v2 [353] at NLO in QCD and with the NNPDF 3.0 NLO PDF set interfaced to PYTHIA8.230 with the settings $h_{\text{damp}} = 3/2 \cdot m_t$ and $\mu_R = \mu_F = m_{T,t}$. An alternative sample with $h_{\text{damp}} = 3m_t$ is used as a two-point uncertainty in the damping.

The uncertainty prescription for ISR, the QCD scales, and the NLO matching follows that of the signal samples. For the FSR uncertainty, dedicated alternative samples varying μ_R for Ps emissions up and down by a factor of two are used. Events showered by HERWIG7.1.3 are used to assess Ps and hadronisation uncertainties.

6.2.4. Simulation of Other Background Processes

Higgs-associated single-top production is considered via the $tHjb$ and tWH processes. The $tHjb$ process is simulated in the 4FS via MADGRAPH5_AMC@NLO2.6.2, while tWH is generated in the 5FS via MADGRAPH5_AMC@NLO2.8.2 under usage of the DR1 scheme [150, 187] to treat interference with $t\bar{t}H$. NNPDF 3.0 NLO PDF sets for the respective flavour scheme are employed for both processes and dynamic QCD

⁶ The entry j here refers to an additional quark or gluon to include QCD radiation at NLO [358].

⁷ Later during the $t\bar{t}(H \rightarrow b\bar{b})$ analysis strategy discussion, H_T instead refers to the scalar sum of transverse momenta only, so that the definitions are only equivalent if particle masses are neglected. Additionally, the sum runs over all jets and leptons in the context of the analysis strategy.

⁸ Since the dipole-recoil sample was produced at an early stage of the analysis with $h_{\text{bzd}} = 2$, a reference sample with the same h_{bzd} setting, but otherwise using nominal parameters, is used for the associated two-point systematic.

scales of $H_T/2$ are used. The $tHjb$ (tWH) sample is showered with PYTHIA8.230 (PYTHIA8.245). No modelling uncertainties are considered for tH processes because of their small contributions in the analysis phase space. The detector simulation for the tH processes is performed in AFII.

Similarly, single-top production is split into s -channel, t -channel, and tW samples, all of which are generated using POWHEGBOX v2 [362–364]. The t -channel process is generated in the 4FS with the NNPDF 3.0 NLO 4FS PDF set, while s -channel and tW utilise the 5FS version. The QCD scales are set to $m_{T,b}$ for the t -channel process, m_t for the s -channel process, and use dynamic scales of $H_T/2$ [365] for tW . The Ps simulation for the t - and s -channel samples (tW samples) is done via PYTHIA8.230 (PYTHIA8.309). Additionally, the tW sample uses the DR1 scheme to remove interference with $t\bar{t}$.

In the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy phase space, t -channel and s -channel processes do not have significant contributions and are therefore not assigned modelling uncertainties. Meanwhile, the tW uncertainty prescription for NLO matching and the h_{damp} parameter (set to $3/2 \cdot m_t$ in the nominal tW sample) follows that of the 5FS inclusive $t\bar{t}$ sample. The tW Ps and hadronisation uncertainty is estimated via an alternative sample interfaced to HERWIG7.2.3. Finally, uncertainties arising from the choice of the interference removal scheme are covered via an alternative sample, in which the Ds scheme [150, 366] is employed⁹.

The production of $t\bar{t}W$ is generated using the SHERPA2.2.10 [368, 369] generator with the NNPDF 3.0 NNLO PDF set. The ME is calculated using COMIX [370] and OPEN-LOOPS at NLO in QCD for events containing up to one additional parton and at LO in QCD for up to two additional partons. Additional partons are matched and merged with the SHERPA Ps [371], which is based on Catani-Seymour dipole factorisation. The CKKM matching procedure [372, 373] is extended to NLO via the MEPS@NLO prescription [374] at a merging cut of 30 GeV. Dynamic $H_T/2$ scales are used for μ_R and μ_F . Meanwhile, $t\bar{t}Z$ production is modelled in MADGRAPH5_AMC@NLO2.8.1 at NLO in QCD with the NNPDF 3.0 NLO PDF set and the same dynamic $H_T/2$ scales as $t\bar{t}W$. $t\bar{t}Z$ events are showered with PYTHIA8.244

For $t\bar{t}W$, an estimate for the uncertainties due to the MC generator is obtained via alternative samples produced with MADGRAPH5_AMC@NLO2.3.3 and showered with PYTHIA8.210. The Ps uncertainties in $t\bar{t}Z$ are evaluated via an alternative sample generated in MADGRAPH5_AMC@NLO2.8.1 and showered in HERWIG7.2.1. The $t\bar{t}W$ uncertainty prescription represents a simplified version of the model used by the ATLAS full Run 2 $t\bar{t}(H \rightarrow \text{ML})$ analysis, which has been motivated in dedicated studies [220].

⁹ The Ds sample was produced with an older generator version containing a bug in the top line-shape [367]. Therefore, a DR sample behaving analogously is used as a reference for the Ds sample.

Table 6.2: Mc generator setting for small background processes considered in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis.

Process	ME Generator	Ps Generator	PDF	QCD Perturbation Order
$t\bar{t}t\bar{t}$ ¹⁰	MADGRAPH5_AMC@NLO2.3.3	PYTHIA8.230	NNPDF 3.1 NLO	NLO
$t(Z \rightarrow \ell\ell)q$ ¹¹	MADGRAPH5_AMC@NLO2.6.0	PYTHIA8.230	NNPDF 3.0 NLO 4Fs	NLO
$tW(Z \rightarrow \ell\ell)$ ¹²	MADGRAPH5_AMC@NLO2.6.2	PYTHIA8.235	NNPDF 3.0 NLO	NLO
V + jets	SHERPA2.2.11 ¹³	SHERPA	NNPDF 3.0 NNLO	NLO (≤ 2 partons) Lo (≤ 5 partons)
Diboson ¹⁴	SHERPA2.2.2	SHERPA2.2.2	NNPDF 3.0 NNLO	NLO (≤ 1 parton) Lo (≤ 3 partons)

Small backgrounds due to the $t\bar{t}t\bar{t}$, tZq , tWZ , V + jets, and diboson processes are also considered in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis. Their corresponding Mc generator settings are summarised in Table 6.2. No modelling uncertainties are considered for these smaller Mc samples. Lastly, cross-section uncertainties for all processes considered in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis are provided with the corresponding perturbation order of the calculation in Table 6.3.

Table 6.3: Uncertainties and calculation order for the cross-sections of processes used in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis. The Mc samples are normalised to the associated cross-sections. No uncertainties are listed for the $t\bar{t}$ + jets processes, as the associated processes have free-floating normalisation factors in the fit setups. Similarly, the $t\bar{t}H$ cross-section uncertainties are only used for the signal strength measurement presented in Chapter 8. The uncertainties on the $tW(Z \rightarrow \ell\ell)$, V + jets, and diboson processes are conservative estimates.

Process	Cross-section Uncertainties [%]	Calculation Order	Reference(s)
$t\bar{t}H$	± 9.2 (QCD scales) ± 3.6 (PDF+ α_S)	NLO+EW NLO	[168]
$t\bar{t}$	None used	NNLO+NNLL	[139]
$t\bar{t}b\bar{b}$	None used	NLO+Ps	[358]
$tHj\bar{b}$	± 14.9 (QCD scales) ± 3.7 (PDF+ α_S)	NLO	[168]
tWH	± 6.7 (QCD scales) ± 6.3 (PDF+ α_S)	NLO	[187]
Single-top	± 5.0 (s- & t-channel)	NLO	[145, 146]
tW	$^{+2.4}_{-2.3}$ (QCD scales) ± 2.8 (PDF+ α_S)	NLO+NNLL	[148]
$t\bar{t}Z$	± 11.3 (QCD scales) ± 4.0 (PDF+ α_S)	NLO+EW NLO	[168]
$t\bar{t}W$	± 6.7 (QCD scales) ± 2.6 (PDF+ α_S) ± 1.8 (2-loop approx.)	NNLO+EW NLO	[375]
$t\bar{t}t\bar{t}$	$^{+70}_{-15}$ (ATLAS observation) ¹⁵	NLO+EW NLO	[376]
$t(Z \rightarrow \ell\ell)q$	± 7.9 (QCD scales) ± 0.9 (PDF+ α_S)	NLO+Ps	[341]
$tW(Z \rightarrow \ell\ell)$	± 50	NLO+Ps	[341]
W + jets	$\pm 40, \pm 30$ (≥ 2 heavy-flavour jets)	NNLO	[377]
Z + jets	± 35	NNLO	[377]
Diboson	± 50	NLO	[378]

¹⁰ Detector simulation in AFII.¹¹ Generated in the 4Fs.¹² Using the DR1 scheme for treating $t\bar{t}Z$ interference.¹³ ($Z \rightarrow \tau\tau$) + jets samples are generated in SHERPA2.2.14.¹⁴ $ZZ \rightarrow 4\nu$ and samples with hadronic boson decays are generated in SHERPA2.2.1.¹⁵ The up variation covers the ATLAS $t\bar{t}t\bar{t}$ observation reported in Reference [379].

6.2.5. $t\bar{t}$ + jets Classification

The inclusive 5Fs $t\bar{t}$ samples described in Section 6.2.3 can contain additional b -jets not originating from the top decays from higher-order effects. To avoid overlap of these events with the dedicated 4Fs $t\bar{t}b\bar{b}$ samples, and for fine-grained control over the different $t\bar{t}$ + jets components in general, events are classified into five $t\bar{t}$ + jets categories: events in the three $t\bar{t} + \geq 1b$ categories $t\bar{t} + \geq 2b$, $t\bar{t} + 1B$, and $t\bar{t} + 1b$ introduced in Section 1.2.3 come from the 4Fs samples exclusively. Meanwhile, events in the additional $t\bar{t} + \geq 1c$ and $t\bar{t}$ + light categories with vetos on additional b -jets and additional b - and c -jets, respectively, come from the 5Fs inclusive samples. The $t\bar{t}$ + light category contains all remaining events, in which no jets originate from heavy-flavour hadrons.

In the particle-level¹⁶ Mc samples, the so-called *truth jets* are reconstructed from stable particles with a mean lifetime above $c\tau = 10$ mm using the anti- k_T clustering algorithm [380] with a radius parameter $R = 0.4$. Heavy-flavour particle-level hadrons not originating from Higgs boson, top quark, and W boson decays are then matched with a maximum distance of $\Delta R = 0.4$ to the closest truth jet satisfying $p_T > 15$ GeV and $|\eta| < 2.5$ to classify individual jets based on their b -hadron and c -hadron multiplicity. Identification and classification of the hadrons satisfying $p_T > 5$ GeV and $|\eta| < 2.5$ is performed via the `HadronOriginClassifier` tool [381] during the TOPQ1 derivation step of Mc events. Based on their jet matches, events are assigned to the $t\bar{t}$ + jets categories as follows:

- $t\bar{t} + \geq 2b$: Events with at least two additional jets matched to at least one b -hadron each.
- $t\bar{t} + 1B$: Events with exactly one additional jet matched to at least two b -hadrons.
- $t\bar{t} + 1b$: Events with exactly one additional jet matched to exactly one b -hadron.
- $t\bar{t} + \geq 1c$: Events with at least one additional jet matched to at least one c -hadron and no additional jets matched to b -hadrons.
- $t\bar{t}$ + light: Events with no jets matched to b -hadrons or c -hadrons.

Events in the $t\bar{t} + \geq 1c$ and $t\bar{t}$ + light categories are removed from the 4Fs $t\bar{t}b\bar{b}$ sample, while events in the $t\bar{t} + \geq 1b$ categories are removed from the 5Fs inclusive $t\bar{t}$ samples. The event classification Transformers also target the five $t\bar{t}$ + jets classes in addition to a separate $t\bar{t}H$ signal class to improve the background yield estimation performed in the statistical fits.

¹⁶ Detector and reconstruction effects are not included on particle level.

6.3. Object & Event Selection

6.3.1. Event Triggers

Before the kinematic event selection, a first selection on data and simulated MC events is made via single-lepton event triggers suitable for the semi- and dileptonic top decay channels. The HLT triggers in ATLAS are set up in so-called *trigger chains*, where a single L1 trigger item is fed to multiple HLT algorithms with different trigger conditions. The resulting HLT single-electron triggers follow the nomenclature

$$\text{HLT_e}\{\min p_T\}_{\{\text{ID}\}}_{\{\text{Isol}\}},$$

where the minimum HLT p_T is given in GeV, the likelihood-based identification working points ID emulate the ones used in the offline reconstruction outlined in the next subsection¹⁷, and the isolation requirement Isol is only present in some HLT triggers [382]. Single-muon triggers use an analogous nomenclature to the single-electron ones but do not impose identification requirements [383].

Table 6.4: List of the single-lepton triggers used for event selection in the different years of Run 2. The naming schema is explained in the text.

Period	Single-Electron Triggers	Single-Muon Triggers
2015	HLT_e24_lhmedium_L1EM20VH	HLT_mu20_iloose_L1MU15
	HLT_e60_lhmedium	HLT_mu50
	HLT_e120_lhloose	
2016–2018	HLT_e26_lhtight_nod0_ivarloose	HLT_mu26_ivarmedium
	HLT_e60_lhmedium_nod0	HLT_mu50
	HLT_e140_lhloose_nod0	

In the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis, the single-lepton triggers of the lowest p_T threshold without prescaling¹⁸ are used for both decay channels. In 2015, lower trigger p_T thresholds than in later years were used because of lower event rates [384]. These changing conditions are reflected in the L1 items used by the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis during the different years in Table 6.4. The triggers are combined with a logical OR, so that each event only has to pass one of the listed triggers to be considered. In 2015 (2016–2018), the lowest L1 trigger lepton p_T threshold was 24 GeV (26 GeV) for electrons and 20 GeV (26 GeV) for muons. These inconsistencies are mitigated by the uniform lepton p_T requirements of the analysis event selection outlined in Section 6.3.3.

¹⁷ A nod0 suffix is added if the impact parameter and its significance are excluded from the likelihood.

¹⁸ Trigger prescaling with a factor n results in every triggered event having a $1/n$ probability of being recorded to budget the limited detector read-out bandwidth.

6.3.2. Object Selection

Jets are reconstructed from particle flow objects built from calorimeter deposits and matched ID tracks [385] via the anti- k_T algorithm [380] in the FASTJET implementation [386] with a radius parameter of $R = 0.4$. They need to fulfil the Tight criterion of the jet vertex tagger algorithm [387], which employs a discriminant technique to differentiate between pile-up jets and those coming from the primary collision vertex defined as the vertex with at least two associated tracks and the highest scalar sum of squared track p_T [388]. Additionally, quality criteria are imposed via the BadLoose WP to reduce non-collision backgrounds and jets originating from detector noise [389]. Only jets with $p_T > 25$ GeV and $|\eta| < 2.5$ are selected.

The DL1r b -tagger [324] is used to assign every jet a Pseudo-Continuous b -Tagging (PCBT) score corresponding to the four WPs at 60 %, 70 %, 77 %, and 85 % b -tagging efficiency and the so-called *untagged* jets, which do not pass any PCBT WP. To this purpose, distinctive features of b -jets are used: since b -hadrons have a reasonably long lifetime¹⁹ of $\langle c\tau \rangle \sim 470$ μm , large track impact parameters and displaced secondary vertices as reconstructed by particle hits in the ID can be used as their discriminating features in jets. DL1r employs five low-level reconstruction algorithms: two of them, ID2P and ID3P, output likelihood ratio discriminants between the different jet flavours based on the signed transversal (ID2P and ID3P) and longitudinal (ID3P only) impact parameter significances. RNNIP is instead based on an RNN to also take into account correlations between the impact parameter significances of multiple tracks belonging to the same jet. SV1 instead reconstructs characteristic secondary vertices in an iterative χ^2 fit, while JetFitter attempts to reconstruct the heavy hadron decay chain using a Kalman filter. The discriminating features output by these complementary algorithms are then combined in a Deep Neural Network (DNN) to predict the probability of each jet to be a light, c -, and b -jet, before the probabilities are combined into a final discriminant forming the basis of the PCBT WPs.

Electrons are reconstructed from energy deposits in the ECAL, which have to match ID tracks with loosened fit criteria accounting for the effects of Bremsstrahlung [390]. The track matching hereby is performed to differentiate electrons from photons, which generate calorimeter clusters with either no associated tracks or with partial tracks from secondary electrons produced in photon conversions with the ID material. Electrons are selected at a p_T threshold of 10 GeV and with $|\eta| < 2.47$, where the region of $1.37 < |\eta| < 1.52$ containing the transition from the barrel to the end-cap region of the ECAL is excluded. Meanwhile, muons are reconstructed from separate track fits in the MS and the ID, which are then matched for global track fitting [391]. They have to pass the same p_T threshold as electrons and lie in the pseudorapidity range of $|\eta| < 2.5$. In order to differentiate prompt leptons from those produced in non-prompt hadronic decays and from other backgrounds, elec-

¹⁹ Averaged over the admixture of b -hadron species [37].

tron (muon) candidates have to pass the MediumLH (Loose) particle identification WP, which is likelihood-based for electrons and cut-based for muons. Additionally, a match with the primary event vertex is needed with a longitudinal impact parameter $|z_0 \sin \theta| < 0.5$ mm and a transverse impact parameter significance $|d_0|/\sigma(d_0)$ of less than five (three) for electrons (muons). Following the overlap removal procedure described below, the identification criteria for the lepton candidates are tightened, and additional isolation criteria are introduced.

Hadronic tau lepton candidates are not used in the analysis itself, but instead serve to veto events for ensuring orthogonality with the $t\bar{t}(H \rightarrow \text{ML})$ analysis in ATLAS Higgs combination measurements. They are reconstructed as jets with radius parameter $R = 0.4$ from energy clusters in the calorimeters [392]. Owing to their principal single-prong and three-prong decay modes leading to a low track multiplicity and a characteristic secondary vertex, hadronically decaying tau leptons are discriminated against hadronic jets by an RNN-based identification algorithm using tracks, calorimeter clusters, and global jet features as inputs [393]. They are required to have $p_T > 25$ GeV and $|\eta| < 2.5$ in addition to passing the Medium WP of this algorithm.

The missing transverse energy E_T^{miss} is a proxy for the p_T carried off in the event by undetected particles, such as neutrinos produced in the leptonic top quark decays. It is defined as the negative vector sum of the transverse momenta of all calibrated objects originating from the primary vertex while avoiding double counting of detector signals [394]. In addition, it involves a so-called *soft term*, which contains all ID tracks associated with the primary vertex but not matched to any calibrated object to account for momenta not matched to reconstructed objects.

To avoid assigning detector signals to multiple objects, an overlap removal procedure is carried out using the distance $\Delta R_y = \sqrt{\Delta y^2 + \Delta \phi^2}$ in y - ϕ space²⁰: initially, the closest jet within $\Delta R_y = 0.2$ of each electron is removed to avoid double-counting of the electron energy deposits in the ECAL. Afterwards, all electrons within $\Delta R_y = 0.4$ of the remaining jets are removed. Additionally, jets with fewer than three tracks and within $\Delta R_y = 0.2$ of muons are removed to retain efficiency for high- p_T muons with significant energy deposition in the calorimeters. However, muons within $\Delta R_y = 0.4$ of the remaining jets are removed to decrease the contribution of non-prompt muons from heavy-flavour jet decays. Lastly, hadronic tau candidates are removed if they are within $\Delta R_y = 0.2$ of electrons or muons, which can fake the single- and three-prong tau lepton decays. Following the overlap removal, additional tighter electron and muon criteria are used during the event selection described in Section 6.3.3: electrons are required to fulfill the TightLH and muons the Medium identification WP. Additionally, electrons (muons) are required to fulfil the FCTight (TightTrackOnly_VarRad) isolation WP to reduce contributions from non-prompt leptons produced in hadronic decays and, in the case of electrons, contributions from mis-identified light hadrons. For this purpose, lepton candidates in a lepton- p_T

²⁰ This follows the ATLAS Run 2 recommendations, as η is not a good approximation of y for jets in general [395].

dependent cone with many high- p_T ID tracks are rejected. The electron isolation criteria additionally reject electron candidates in environments with large additional energy deposits in the ECAL using a fixed cone of radius $R = 0.2$ around the electron.

In addition to the muons described previously, the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis additionally makes use of non-prompt muons, the so-called *softmuons*, which do not enter the event selection or the overlap removal procedure. In fact, they are instead specifically required to lie within $\Delta R < 0.4$ of jets b -tagged at a 85 % DL1r WP for b -jet momentum corrections, as outlined in Section 6.5.3. Softmuons need to pass a p_T threshold of 4 GeV, fall within $|\eta| < 2.7$, and pass the Medium identification WP. Furthermore, softmuons lying within $\Delta R < 0.1$ of loose muons passing the overlap removal are removed to avoid double-counting event objects. The Softmuon collection [346] of AnalysisTop is used to facilitate the softmuon selection.

To enhance sensitivity in the high- p_T Higgs boson region, the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis targets boosted topologies in addition to a resolved selection. For the corresponding events, reclustered jets [396] with a radius parameter $R = 1.0$ are initially selected with a minimum p_T of 300 GeV, a mass in the range $100 \text{ GeV} < m_{\text{RC-jet}} < 140 \text{ GeV}$ around that of the SM Higgs boson, a pseudorapidity of $|\eta| < 2.0$, exactly two sub-jets b -tagged at the 85 % DL1r WP, and a minimum distance of $\Delta R > 1.0$ from all electrons in the event which pass the overlap removal and tightened selection criteria. Following the strategy used in the first full Run 2 analysis, reclustered jets are passed through a multi-class DNN, which assigns probability scores for the reclustered jet to be a Higgs boson, a top quark, or a QCD multi-jet based on 17 jet substructure variables. Boosted Higgs candidates are required to have a Higgs probability score of at least 0.4. The DNN tagger was retrained after the first full Run 2 analysis to be robust against modelling differences in the jet substructure between full simulation and AFII as described in detail in Reference [397].

6.3.3. Event Selection

Table 6.5: Event selection criteria employed for the different channels of the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis. Only b -tagging requirements not implicitly fulfilled by requirements at tighter working points are listed. Events passing the single-lepton boosted selection are vetoed for the single-lepton resolved selection.

Channel	# Jets	# DL1r b -Tags			# e/μ	# τ_{had}	# Boosted Higgs Candidates
		@ 70 %	@ 77 %	@ 85 %			
Dilepton	≥ 3	≥ 2	–	≥ 3	2	0	n/a
Single-Lepton Resolved	≥ 5	≥ 3	–	–	1	≤ 1	–
Single-Lepton Boosted	≥ 4	–	≥ 2	$\geq 4^{21}$	1	≤ 1	≥ 1

²¹ Exactly two b -jets have to be contained in the boosted Higgs candidate, the other two come from the requirement of ≥ 2 b -jets at 77 %.

Since the analysis targets the semileptonic and dileptonic $t\bar{t}$ decay modes, events are primarily required to contain either exactly one electron or exactly one muon fulfilling the tight criteria in the so-called single-lepton channel, and exactly two tight electrons and/or muons of opposite electric charge in the so-called dilepton channel, without any additional loose leptons of $p_T > 10$ GeV in either channel. A same-flavour trigger match for at least one reconstructed lepton is required within $\Delta R < 0.15$ and the leading lepton needs a minimum p_T of 27 GeV in both channels. Additionally, the p_T threshold for the subleading electron in e^+e^- events is increased to 15 GeV. To further improve signal purity, the invariant dilepton mass for e^+e^- and $\mu^+\mu^-$ events has to lie above 15 GeV to suppress heavy-flavour resonances and low-mass Drell-Yan processes, while an additional veto in the Z boson mass window of $83 \text{ GeV} < m_{\ell\ell} < 99 \text{ GeV}$ is used to suppress processes containing Z bosons produced at resonance.

The single-lepton channel is further split into the non-overlapping *resolved* and *boosted* channels. A boosted Higgs candidate as defined in Section 6.3.2 is required for events in the single-lepton boosted channel. Additionally, the channel requires at least four jets (including those in the boosted Higgs candidate), while at least two jets not contained in the boosted Higgs candidate have to pass the 77 % DL1r b -tagging WP. Single-lepton events not passing the boosted criteria can instead be included in the resolved channel, if they fulfil the jet multiplicity (N_{jet}) and b -tagging requirements listed in Table 6.5, which also contains the analogous requirements for the dilepton channel. Lastly, events with more than one (more than zero) hadronic tau candidates are vetoed in the single-lepton (dilepton) channel to avoid overlap of the $t\bar{t}(H \rightarrow b\bar{b})$ and $t\bar{t}(H \rightarrow \text{ML})$ analyses in later Higgs combination measurements.

Overall, 5.3 % of all $t\bar{t}(H \rightarrow b\bar{b})$ events produced in the collisions are expected to be selected by these criteria in the two single-lepton channels, while the dilepton channel is expected to accept an additional 1.0 %. Additionally, $H \rightarrow b\bar{b}$ events are expected to account for 92.8 %, 96.8 %, and 82.4 % of all selected $t\bar{t}H$ events in the single-lepton resolved, single-lepton boosted, and dilepton channel, respectively. The individual analysis channels are further split into Signal Regions (SRs) enriched in $t\bar{t}H$ events and Control Regions (CRs) dominated by the individual $t\bar{t} + \text{jets}$ components via the classification and reconstruction Transformers as outlined in Section 6.5.4. While the CRs help to constrain the background rates, the SRs extract the $t\bar{t}H$ cross-section in a simultaneous fit of all regions as presented in Chapter 8.

6.4. Profile Likelihood Fits

Both the analysis strategy optimisation studies outlined below and the cross-section extraction presented in Chapter 8 heavily rely on so-called *profile likelihood fits* in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis – similar in fact to a large part of experimental particle physics. In general, a likelihood is the joint probability density of obtaining a set of

results $x_i \in \mathbf{x}$ given a statistical model with underlying parameters λ as

$$L(\mathbf{x} | \lambda) = \prod_i P(x_i | \lambda). \quad (6.1)$$

The likelihood can then be maximised to obtain the optimal parameter values $\hat{\lambda}$, although for numerical reasons the negative log likelihood is generally minimised instead.

However, models have to fit the data in experimental science, while the likelihood of Equation (6.1) instead operates the other way around. As a solution, the uncertainties of repeated measurements on θ , the so-called *nuisance parameters* (NPs), are encoded in additional constraint terms $P(\theta_j)$ in the frequentist picture commonly adopted in particle physics. Meanwhile, the *parameters of interest* (POIs) μ are the remaining free parameters. The resulting negative log likelihood to be minimised is

$$-\ln L(\mu, \theta) = -\sum_i \ln P(x_i | \mu, \theta) - \sum_j \ln P(\theta_j), \quad (6.2)$$

where the explicit dependence on the measurements x_i has been dropped in the overall likelihood since these are assumed to be fixed.

In particle physics, binned distributions of data and Mc simulations are typically used instead of individual measurements. Additionally, the data consists of signal processes x and background processes y which are predicted in Mc simulations with NP-dependent event counts $S_{x,i}(\theta)$ and $B_{y,i}(\theta)$, respectively, which for each process consist of the overall number of expected events and their probability to fall into each bin i . For the event count N_i in bin i , this results in

$$N_i = \sum_x \mu_x S_{x,i}(\theta) + \sum_y B_{y,i}(\theta), \quad (6.3)$$

where μ_x denotes the signal strength POIs. The event count N_i in Equation (6.3) can be interpreted as the expectation value of a Poisson distribution in bin i . Consequently, the $P(n_i | \mu, \theta)$ terms in Equation (6.2), where the measured bin-wise event yields n_i replace the individual measurements x_i , can be represented in the *extended log-likelihood* [398] as

$$\begin{aligned} -\ln P(n_i | \mu, \theta) &= -\ln \text{Poisson}(n_i | \sum_x \mu_x S_{x,i}(\theta) + \sum_y B_{y,i}(\theta)) \\ &= -n_i \ln(\sum_x \mu_x S_{x,i}(\theta) + \sum_y B_{y,i}(\theta)) + \sum_x \mu_x S_{x,i}(\theta) + \sum_y B_{y,i}(\theta), \end{aligned}$$

up to a constant combinatorial factor dependent on n_i , which has no effect during the likelihood maximisation.

Meanwhile, the NPs can be split into normalisation effects, which affect the overall rates of signal and/or background processes, and shape effects affecting the individual bins differently. Usually, the NPs are encoded in Gaussian constraint terms, in

which the alternative and reweighted MC samples introduced in Section 6.2 are used as estimates of the impact of modelling uncertainties in so-called *templates*. These are defined as one standard deviation away from the nominal NP parameter in the associated constraint term. Interpolation and extrapolation are used for other values of the NPs. Experimental uncertainties are treated similarly. Lastly, uncertainties due to the limited MC sample statistics are incorporated on a per-bin basis without correlations between the various bins and using Poissonian constraint terms. The $t\bar{t}$ + jets normalisation factors are special in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis in the sense that they are free-floating NPs determined via auxiliary measurements of the associated processes dominated by their dedicated analysis CRs. In the following, they will therefore be denoted as k_j for the $t\bar{t}$ + jets component j .

To obtain uncertainties on the POIs from the likelihood maximisation, the profiling method is used. Profiling initially determines the best-fit NP and POI values $\hat{\theta}$ and $\hat{\mu}$ in an *unconditional* likelihood fit. Afterwards, the likelihood is maximised in a series of fits with one POI μ fixed at a time to obtain *conditional* best-fit NPs²² $\hat{\theta}(\mu)$. The resulting test statistic $t(\mu)$ is based on the profile likelihood ratio as

$$t(\mu) = -2 \ln \frac{L(\mu, \hat{\theta})}{L(\hat{\mu}, \hat{\theta})}, \quad (6.4)$$

where the measured bin yields n_i have been dropped from the explicit parameters. Since $t(\mu)$ corresponds to the χ^2 distribution in the asymptotic large sample limit according to Wilks' theorem [399], it can be used to obtain p -value estimates, which allow the determination of uncertainties on $\hat{\mu}$. Additionally, the special $\mu = 0$ case of the likelihood ratio in Equation (6.4) representing the background-only hypothesis is used to evaluate the discovery significance of the $t\bar{t}H$ signal process.

A similar procedure to estimate the expected POI uncertainties and the expected signal sensitivity during the analysis strategy development instead uses the so-called *Asimov* dataset [400], in which the data yields in every bin are assumed to correspond to the yields from the MC simulations. As a result, the expected uncertainties on μ and k , and the expected significance against a background-only hypothesis can be estimated via profiling without relying on collider data.

The statistical profile likelihood fits used in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis are carried out with the TRExFitter [401] framework, which uses the ROOT-based RooFit [402] and RooStats [403] frameworks in the HistFactory [404] XML-format. The likelihood minimisation procedure is performed using the Minuit2 [405] library implemented in ROOT.

²² And also conditional best-fit values for the other POIs in multi-POI fits.

6.5. Analysis Strategy

6.5.1. Monte-Carlo Corrections

Prior to analysis strategy optimisations, the overall $t\bar{t}$ + jets yields were corrected to more accurately reflect the ATLAS data. The yield correction scale factors shown in Table 6.6 were derived in two preliminary single-channel fits of the CRs and parts of the SRS with low $t\bar{t}H$ admixture²³. Additionally, the $t\bar{t} + 1B$ and $t\bar{t} + 1b$ scale factors were combined, and the $t\bar{t}H$ sample yield was fixed to zero, resulting in so-called *background-only* fits. The corrections do not have a direct impact on the fit setups of Chapter 8 used for extracting the analysis results, since the $t\bar{t}$ + jets yields are left as free-floating parameters there. However, they are included in the $t\bar{t}$ + jets yields for the studies discussed in the remainder of this section unless specified otherwise.

Table 6.6: Per-channel $t\bar{t}$ + jets yield scale factors with uncertainties as derived for the analysis strategy optimisation in single-channel background-only fits.

Channel	$t\bar{t}$ + light	$t\bar{t} + \geq 1c$	$t\bar{t} + 1b$	$t\bar{t} + 1B$	$t\bar{t} + \geq 2b$
Single Lepton	$0.74^{+0.10}_{-0.10}$	$1.79^{+0.25}_{-0.23}$	$1.08^{+0.14}_{-0.13}$		$0.96^{+0.14}_{-0.12}$
Dilepton	$0.85^{+0.17}_{-0.15}$	$1.59^{+0.22}_{-0.20}$	$1.29^{+0.22}_{-0.20}$		$0.93^{+0.15}_{-0.13}$

During analysis pre-fit studies, a significant shape discrepancy in the H_T distributions was observed. This discrepancy likely originates from modelling difficulties in the $t\bar{t}$ + jets Ps simulations in the high- N_{jet} $t\bar{t}(H \rightarrow b\bar{b})$ analysis phase space. Consequently, correction factors for an H_T -based event reweighting of the $t\bar{t} + \geq 1c$ and $t\bar{t}$ + light samples were derived. A single-lepton selection with at least five jets and exactly two b -tagged jets at the 70 % DL1r WP, and a dilepton selection with at least three jets and exactly two b -tagged jets at the 70 % DL1r WP without any additional jets passing the 85 % WP are used for this purpose to ensure orthogonality with the analysis pre-selection. Additionally, the regions are split by jet multiplicity from $N_{\text{jet}} = 5$ to $N_{\text{jet}} \geq 8$ and from $N_{\text{jet}} = 3$ to $N_{\text{jet}} \geq 6$ in the single-lepton and dilepton channels, respectively. After subtracting other processes, the data yields in these regions were compared to the $t\bar{t} + \geq 1c$ and $t\bar{t}$ + light yields in bins of H_T to extract initial bin-wise event reweighting factors. Subsequently, the reweighting factors were refined via quadratic splines defined piece-wise in each bin. These were fitted to the distributions under the conditions of continuity and differentiability at the bin boundaries, while ensuring that the total yields were conserved in each bin by the spline reweighting. Additionally, systematic uncertainties were derived for the procedure by varying the yield scale factors of the $t\bar{t} + \geq 1c$ and $t\bar{t}$ + light samples

²³ In ATLAS analyses, signal-rich regions are *blinded* to allow feedback from fits during the analysis strategy development while avoiding bias from premature exposure to results. In the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis, blinding thresholds of 7.7 % and 12 % in the signal-over-background ratio S/B are applied to variables used in fits and for validation, respectively.

within their uncertainties listed in Table 6.6, and by completely switching off the reweighting on a per- N_{jet} basis. Comparisons of pre-fit H_T distributions before and after the application of the H_T reweighting in the single-lepton resolved and the dilepton channel are shown in Figure 6.2.

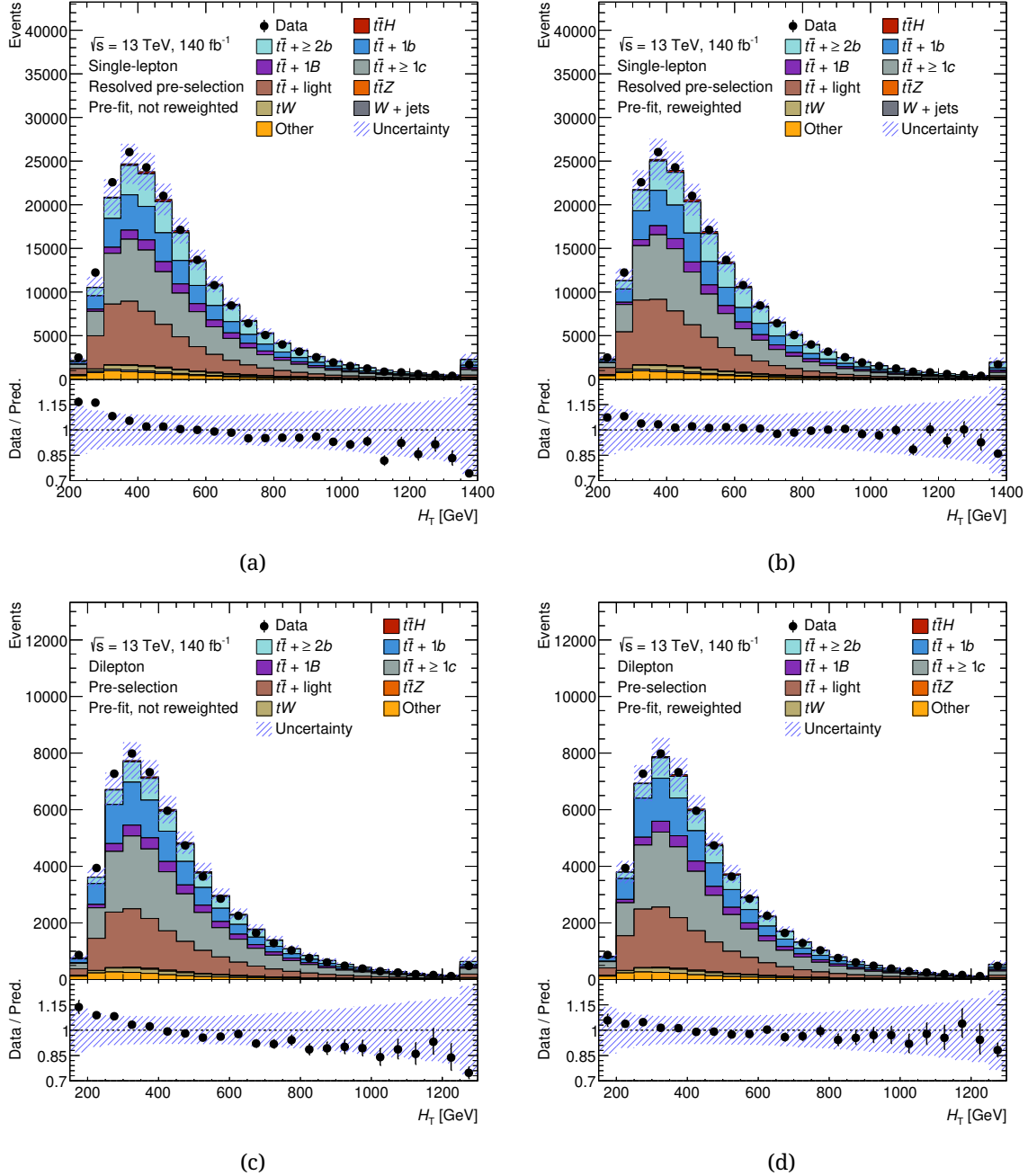


Figure 6.2: Comparisons of pre-fit H_T distributions in (a) + (b) the single-lepton resolved and (c) + (d) the dilepton channel. The distributions (a) + (c) before and (b) + (d) after application of the H_T reweighting are shown. All uncertainties are included in the uncertainty band.

6.5.2. Determination of Fake Lepton Backgrounds

Fake and non-prompt lepton backgrounds are modelled via dedicated samples in both the single-lepton and dilepton channels. For muons, the main participating processes are non-prompt lepton production from b -hadron or c -hadron decays, while fake electrons have additional significant contributions from mis-identified hadrons and converted photons. In the single-lepton channel, these processes are estimated with a data-driven *fake factor* method [406], which uses an orthogonal region enriched in fake leptons to derive a fake factor via tight and loose lepton definitions. Subsequently, this fake factor is used to estimate fake contributions in the analysis region by applying it to events in the analysis phase space that pass the loose lepton definition but fail the tight one. In the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis, the fake-enriched region is created by requiring at least two jets and at least two b -tags at the 70 % DL1r WP, and a sum of E_T^{miss} and the transverse W boson mass²⁴ $m_{T,W}$ below 60 GeV to suppress contributions from leptonic top quark and W boson decays. The fake factors are parametrised by the leading jet p_T and the $|\eta|$ of the lepton. Systematic uncertainties on the single-lepton fakes are estimated using an alternative fake factor parametrisation in addition to a flat normalisation uncertainty of 50 %.

In the dilepton channel, Mc simulations are used to estimate fake and non-prompt lepton backgrounds. A similar procedure was already in use during the first full Run 2 analysis, where a yield scale factor of 1.9 was derived for the fake yields using a fit in a fake-enriched region requiring two so-called *same-sign* leptons of the same electric charge [407]. Since the criteria used to define prompt leptons in Mc samples based on truth-level information evolved between the two analyses, the dilepton fakes prescription was revisited for the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis. As a first step, the new prompt lepton definition was validated by comparing the yields of events containing fakes with those only containing prompt leptons in processes dominated by fake and prompt leptons. Semi- and dileptonic $t\bar{t} + \geq 1c$ and $t\bar{t} + \text{light}$ events were chosen for this purpose due to their expected differences in fake contamination in the dileptonic channel. Since the semileptonic sample contained fakes with an event fraction of around 89 %, while the dileptonic sample was dominated by events without fakes at a fraction of more than 99 %, the new Mc prompt lepton definition gave no grounds for concern.

After the validation of the prompt lepton criteria, fits in the same-sign region were repeated for the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis to determine a new scale factor for dilepton fake processes. Apart from the requirement on the electric charge of the leptons, the fits use events selected with the analysis pre-selection criteria. Mc processes entering the fits were split up into events containing Mc fake leptons and those without fake leptons. However, V + jets and diboson processes were not considered in the fake sample because of their small contributions and suspected issues in the

²⁴ This is the ‘visible’ mass from the leptonic W boson defined via E_T^{miss} , the lepton p_T , and the azimuthal angle $\Delta\phi$ between them as $m_{T,W} = \sqrt{2p_{T,\ell} E_T^{\text{miss}} (1 - \cos \Delta\phi)}$.

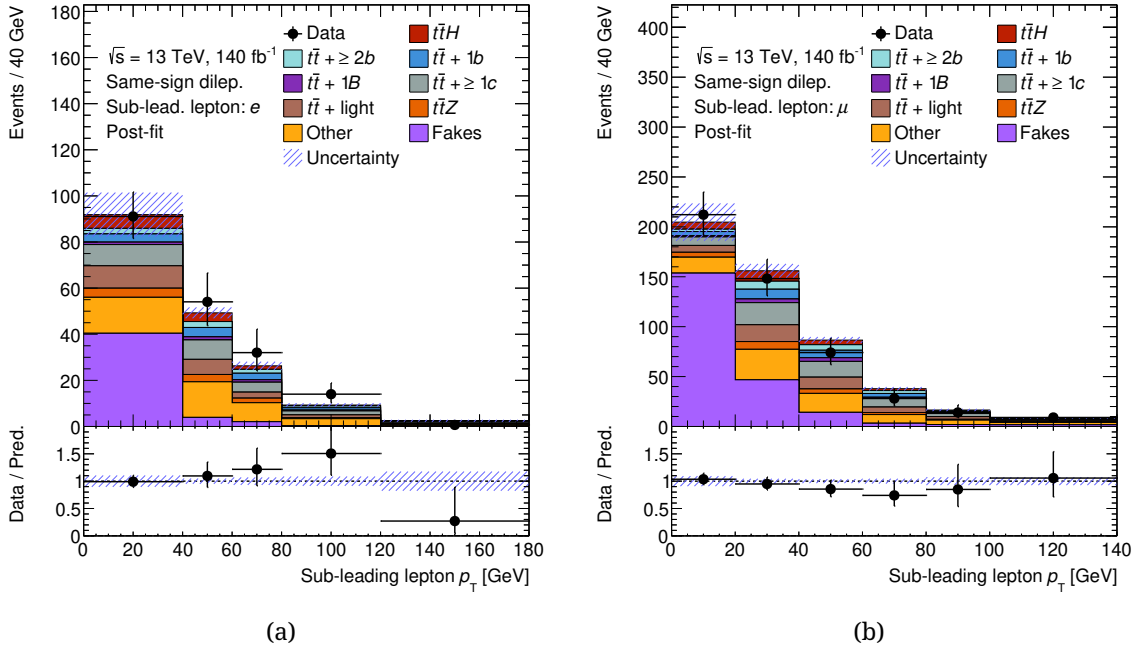


Figure 6.3: Post-fit distributions of the subleading lepton p_T in same-sign dilepton events with (a) subleading electrons and (b) subleading muons. The fit results are listed in Table 6.7, where additional fit results on N_{jet} distributions are also presented. Only statistical uncertainties are considered.

Table 6.7: Best-fit normalisations of the dilepton fake processes split by events with subleading electrons and subleading muons. The results of fits on the subleading lepton p_T and the N_{jet} distributions are shown.

Fit Setup	Subleading electron	Subleading muon
$p_{T,\ell}$ Fit	$0.70^{+0.18}_{-0.16}$	$1.34^{+0.18}_{-0.16}$
N_{jet} Fit	$0.79^{+0.20}_{-0.19}$	$1.22^{+0.20}_{-0.19}$

SHERPA parton history used for the prompt lepton categorisation. Similarly, $t\bar{t}H$ and tH events were also not considered in the MC fakes as target processes of the analysis. The yields of events containing at least one lepton fake with subleading electrons and subleading muons were fitted separately, as the subleading leptons represent the more likely candidate for fake leptons. Fits were performed in the distributions of the subleading lepton transverse momenta, while only considering statistical uncertainties, as shown post-fit in Figure 6.3. Additionally, separate fits on the N_{jet} distributions were performed for an alternative scale factor estimate. Since the fit results, shown in Table 6.7, are clustered around unity, the normalisation factor used in the first full Run 2 analysis was dropped. A flat 50 % normalisation uncertainty is considered for the dilepton fake contributions.

6.5.3. Muon-in-Jet Corrections

The b -quarks making up the hadrons of the b -jets in the analysis overwhelmingly decay into a c -quark and a W boson, which can subsequently decay either hadronically or leptonically. In 10.6 % of cases [37] the W boson decays into $\mu\nu_\mu$ ²⁵. Additionally, the tau lepton produced instead with a BR of 11.4 % also decays into $\mu\nu_\mu\nu_\tau$ in 17.8 % of cases²⁶. Therefore, around 12.7 % of b -jets contain a muon from a decay of the b -quark²⁷. The production of muons from subsequent c -quark decays is also possible. A Feynman diagram of a muon-in-jet decay is shown in Figure 6.4.

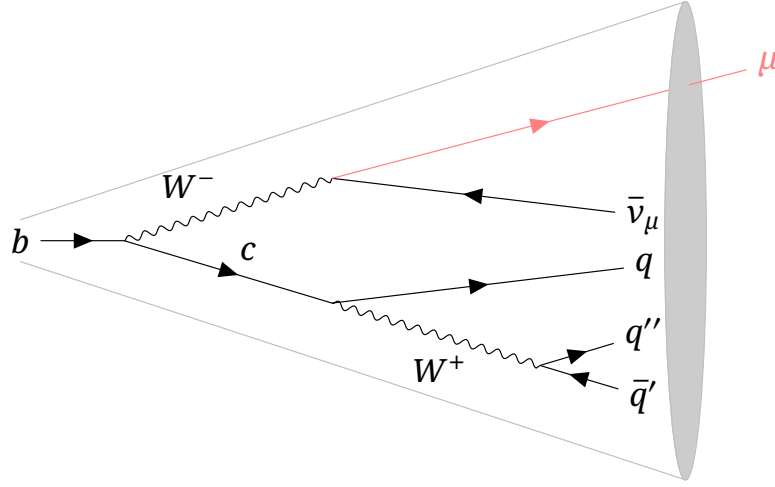


Figure 6.4: Feynman diagram of a semileptonic b -quark decay in a b -jet. The final-state muon, drawn in red, carries a significant fraction of the jet momentum. No spectator quarks are drawn for simplicity.

The lifetime of these muons is long enough that they mostly decay outside the detector. Additionally, they are minimally ionising particles and therefore only deposit a small fraction of their energy in the calorimeters. Therefore, in order to correct the b -jet four-momenta, the softmuon objects defined in Section 6.3.2 are used in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis to correct the four-vector of the closest jet passing the 85 % DL1r working point in the following manner:

$$\mathbf{p}_{\text{jet}}^{\text{corr}} = \mathbf{p}_{\text{jet}} + \mathbf{p}_{\text{muon}} - \mathbf{p}_{\text{muonCaloCluster}}.$$

Here, $\mathbf{p}_{\text{jet}}^{\text{corr}}$, $\mathbf{p}_{\text{jet}}$, and $\mathbf{p}_{\text{muon}}$ refer to the four-momenta of the corrected jet, the uncorrected jet, and the muon, respectively. The time-like four-momentum $\mathbf{p}_{\text{muonCaloCluster}}$ of the muon calorimeter deposit is subtracted to avoid double counting since the deposits are already taken into account in the uncorrected jet four-momentum. Each

²⁵ Explicit anti-fermion labels are neglected in this section for simplicity.

²⁶ Using the inclusive fraction, where additional particles, such as a photon, may be produced.

²⁷ This value neglects contributions from jets containing multiple b -quarks.

b -jet is only corrected by the muon closest to it to reduce fake lepton contributions. While electrons can be produced via in-jet b -hadron decays in a similar fashion, their momentum contribution is reconstructable in the electromagnetic calorimeter. Therefore, no electron-in-jet corrections are performed.

The muon-in-jet procedure mirrors a similar technique utilised for ATLAS $V(H \rightarrow b\bar{b})$ analyses since Run 1 [408, 409] following extensive studies on the procedure. In a study performed in the single-lepton channel of the $t\bar{t}(H \rightarrow b\bar{b})$ analysis with relaxed b -tagging requirements of at least two jets tagged at the 70 % DL1r WP, the muon-in-jet corrections led to a shift from 116 ± 24 GeV to 118 ± 24 GeV in the invariant mass of the two reco-level b -jets truth-matched to the b -hadrons originating in the Higgs boson decay. The muon-in-jet corrected b -jets are only used in the inputs to the Transformers trained for the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis, where the SM Higgs boson mass represents an important discriminating feature learned by the models. The corrections are also applied during the Higgs boson four-vector reconstruction from the two most likely decay jets, as determined by the reconstruction Transformers, in order to improve the object-based four-vector predictions. All other parts of the analysis rely on uncorrected jets.

6.5.4. Region Definitions via Multivariate Analysis Techniques

The $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis extensively relies on Transformer networks for both event classification and reconstruction as outlined in the flowchart contained in Figure 6.5. As the first step in the region assignment, two separate classification Transformers in the single-lepton resolved and the dilepton channel split events into the $t\bar{t}H$ signal category and the five $t\bar{t} + \text{jets}$ background categories defined in Section 6.2.5 using six classifiers – one for $t\bar{t}H$ and five for the $t\bar{t} + \text{jets}$ categories: events with a $t\bar{t}H$ classifier above a threshold chosen to maximise the signal-over-background noise ratio $S/\sqrt{B}$ in the $t\bar{t}H$ region get assigned to the SRs, while other events get assigned to the CR with the highest associated classifier value. Discriminant classifiers are used for the event categorisation as motivated in Section 7.3.1, where the optimisation procedure for the $t\bar{t}H$ classifier threshold is also described in detail.

Following the event categorisation, a second Transformer network infers the two jets in each event in the SRs most likely to originate from an $H \rightarrow b\bar{b}$ decay. The four-vectors of these two jets are then combined to reconstruct the Higgs boson four-vector and determine p_T^H . As a result, the events in the $t\bar{t}H$ region are further divided into six SRs, each corresponding to one of the STXS stage 1.2 bins, so that the sensitivity of the analysis in the differential cross-section measurement is maximised. In the single-lepton resolved channel, the $t\bar{t} + \geq 2b$ CR is also split into three parts via the reconstruction Transformer: to increase the control of the fit over the dominant $t\bar{t} + \geq 2b$ background process, events with a reconstructed p_T^H falling into the STXS bins 1+2, 3+4, and 5+6 are assigned to $t\bar{t} + \geq 2b$ CR 1, CR 2, and CR 3, respectively.

Events in the single-lepton boosted region are instead just categorised in a combined SR or a combined CR depending on whether the $t\bar{t}H$ classifier output by the single-lepton classification Transformer lies above or below threshold, respectively. Afterwards, the single-lepton reconstruction Transformer is used in an analogous way to the other channels. Both the boosted SR and the boosted CR use the reconstructed p_T^H as a fit variable.

The development, validation, and performance of the Transformer networks used in the analysis are described in detail in Chapter 7. After the event categorisation, the fraction of all $t\bar{t}H$ events decaying via $H \rightarrow b\bar{b}$ is 97.7 %, 98.4 %, and 94.7 % in the single-lepton resolved SRs, the single-lepton boosted SR, and the dilepton SRs, respectively.

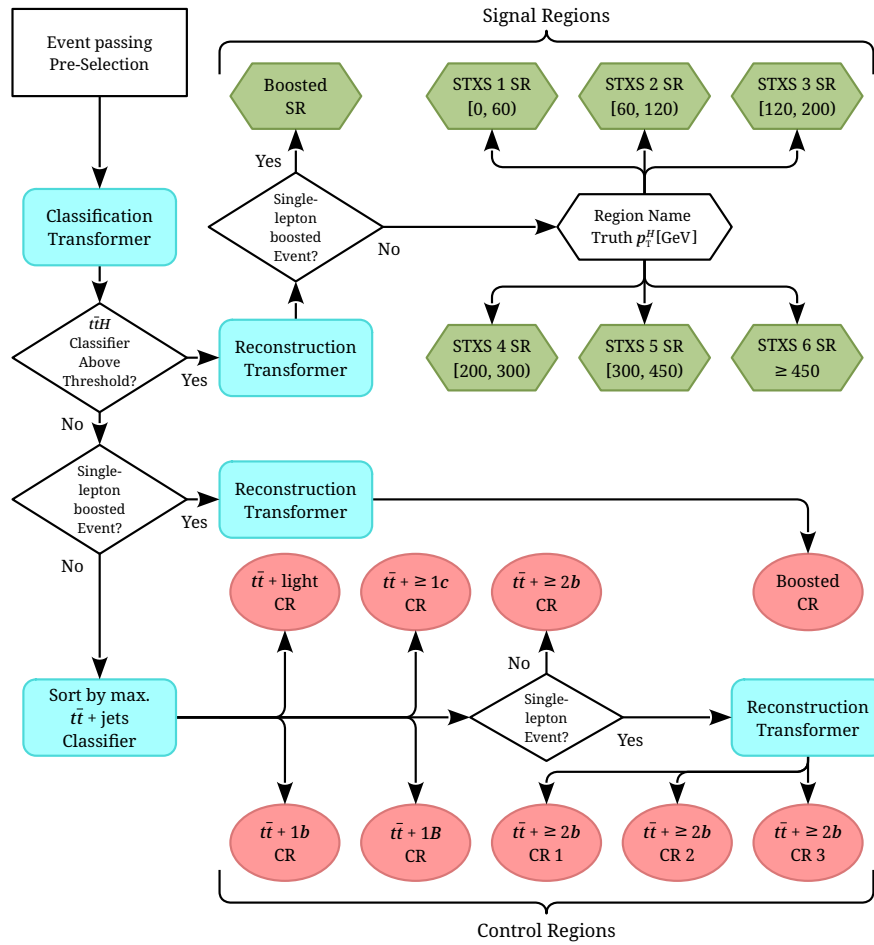


Figure 6.5: Flowchart of the joint event classification and p_T^H reconstruction strategy employed in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis. The discriminant classifiers d_i defined in Equation (7.2) are used for event classification with SR thresholds of 4.072 and 9.031 in the single-lepton and dilepton channels, respectively. The reconstruction Transformer is also used on events in the single-lepton boosted channel, where p_T^H is used as the fit variable, and the single-lepton $t\bar{t} + \geq 2b$ CR, which is further split into three regions CR 1, CR 2, and CR 3 corresponding to the p_T^H ranges of STXS bins 1+2, 3+4, and 5+6, respectively.

Multivariate Analysis Techniques for $t\bar{t}(H \rightarrow b\bar{b})$

Transformer networks are agnostic to the number and ordering of input objects. Simultaneously, they incorporate pair-wise interactions via the attention mechanism introduced in Section 5.3. As a result, the Transformer architecture is well suited for perturbative high-energy physics processes, where the number of particles generated in the inclusive process is not fixed and where a clear ordering of final-state particles is ambiguous at best – the symmetries and structure inherent in the physical system are matched well to the structure and processing of its data representation.

In the following, the development and validation of the networks used for event classification and reconstruction in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis will be outlined in Sections 7.1 to 7.3. Additionally, a comparison with previous MVA methods and alternative strategies is presented in Section 7.4.

7.1. Network Architecture

While the original iteration of the $t\bar{t}(H \rightarrow b\bar{b})$ Transformer models was developed in the course of another PhD project, two retraining campaigns incorporating significant changes to architecture and setup were undertaken during the course of this thesis in addition to extensive network validation studies. In the following, the three Transformer iterations are called Mk1, Mk2, and Mk3, respectively. Additionally, a separate Deep Set¹ network based on high-level input features [410] was developed

¹ Deep Sets can be thought of as Transformers without an attention mechanism, so that no pair-wise object interactions are possible prior to the final pooling operation.

for the single-lepton channel in parallel to the Transformers. It is included in the performance comparisons of Section 7.4.

7.1.1. Model Inputs

Table 7.1: List of all reco-level input features used by the MK3 Transformers in this analysis. Where appropriate, values given to the different input objects and transformations on the input feature distributions, shifting and scaling them to a specific mean μ and standard deviation σ , are included for reference. Overall, kinematic properties are calculated from the object four-vectors, where E_T^{miss} is assumed to have no rest mass and no momentum component in the z-direction. Redundant information is supplied to simplify the task of building the latent feature space in the fully-connected embedding layers set before the attention blocks.

Feature	Description	Feature Transformations
p_x	Object momentum in x-direction.	Rescaled to $\mu = 0, \sigma = 1$.
p_y	Object momentum in y-direction.	Rescaled to $\mu = 0, \sigma = 1$.
p_z	Object momentum in z-direction.	Rescaled to $\mu = 0, \sigma = 1$.
energy	Object energy.	Rescaled to $\mu = 0, \sigma = 1$.
p_T	Object transverse momentum.	Rescaled to $\mu = 0, \sigma = 1$.
mass	Object mass.	Rescaled to $\mu = 0, \sigma = 1$.
η	Object pseudorapidity.	Rescaled to $\mu = 0, \sigma = 1$.
ϕ	Object azimuthal angle.	Rescaled to $\mu = 0, \sigma = 1$.
$\cos \phi$	Sine of object azimuthal angle.	Rescaled to $\mu = 0, \sigma = 1$.
$\sin \phi$	Cosine of object azimuthal angle.	Rescaled to $\mu = 0, \sigma = 1$.
PCBT bin	DL1r pseudo-continuous b -tagging bin assigned to jets in the following manner. Set to 0 for leptons and E_T^{miss} .	None.
$\text{feature} = \begin{cases} 1, & \text{if un-tagged} \\ 2, & \text{if tagged at [85 \%, 77 \%]} \\ 3, & \text{if tagged at [77 \%, 70 \%]} \\ 4, & \text{if tagged at [70 \%, 60 \%]} \\ 5, & \text{if tagged at 60 \%.} \end{cases}$		
lepton type	Lepton type of input objects. Set to 1 for electrons, 2 for muons, and 0 for jets and E_T^{miss} .	None.
lepton charge	Charge of lepton objects in units of e . Set to 0 for jets and E_T^{miss} .	Rescaled to $\mu = 0, \sigma = 1$.
E_T^{miss} flag	Whether input object is E_T^{miss} (value of 1) or not (value of 0).	None.

The Transformer networks in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis use low-level four-vector information of the final-state objects as input features. Consequently, no feature engineering, which might be biased by prior assumptions, is required because the Transformer can learn the reconstruction of event-level information. In dedicated studies, it was found, however, that supplying redundant kinematic infor-

mation improved the training performance. This is likely due to the simplified latent feature construction in the Transformer, resulting from a more varied input feature selection. All kinematic input features are rescaled to fall within a normal distribution in order to aid regularisation in the Transformer. During the Mk2 retraining, a jet charge estimate was included as an additional input variable. However, due to validation issues of the variable and its negligible impact on the overall network performance, it was removed again for the Mk3 models. An exhaustive list of reco-level input features used in the Mk3 models with the corresponding rescaling is provided in Table 7.1.

The Mk1 Transformer only used jets and leptons present in the events as input objects. In the Mk2 retraining, the muon-in-jet corrections described in Section 6.5.3 were included to improve the b -jet four-vector inputs to the event classification and especially the Higgs boson p_T reconstruction based on object four-vectors. Finally, E_T^{miss} was added as an additional input object for the Mk3 retraining to make the complete kinematic information inherent in the events available for model inference.

7.1.2. General Structure

Table 7.2: Table of optimal hyperparameter values of the Mk3 Transformer architecture used in the analysis depicted in Figure 7.1. The total number of trainable parameters present in each resulting model is listed in the last row. Additional hyperparameters defining the training procedure are listed in the appendix.

Parameter	Classification		Reconstruction	
	Single Lepton	Dilepton	Single Lepton	Dilepton
d	256	256	128	128
h	4	4	8	8
n_{embed}	2	2	2	2
n_{attblock}	11	10	10	8
n_{dense}	5	4	3	3
Resulting total number of parameters	6 716 416	5 466 368	1 207 424	976 256

The networks developed for event classification and p_T^H reconstruction mainly differ in their final layers as shown in Figure 7.1. Therefore, the common aspects of the architecture are outlined in the following, while the specific differences will be discussed afterwards. The optimal architecture hyperparameters for the Mk3 models are contained in Table 7.2.

Both flavours of Transformer initially perform a *feature embedding* operation. This step involves n_{embed} linear layers for transforming the input features object-wise into d latent features. In addition to coupling the number of input features to the

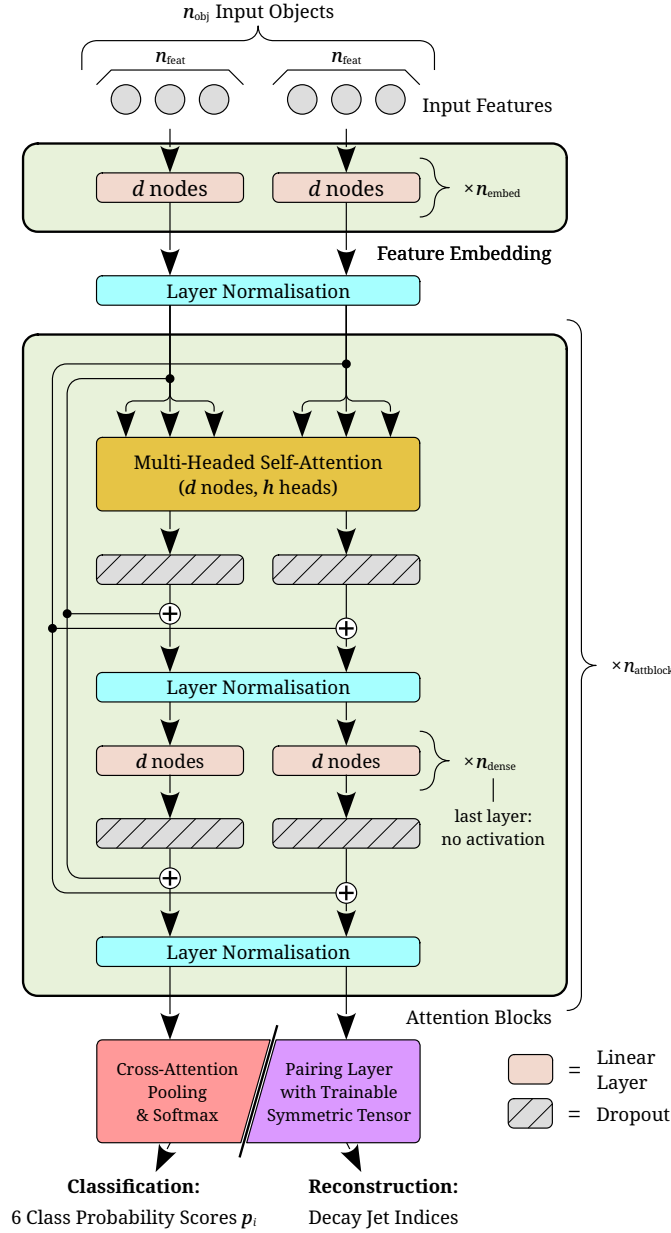


Figure 7.1: Combined schematic of the Transformer network architectures used in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis. Separate networks are trained for the event classification and p_{T}^H reconstruction task in each channel. The multi-headed self-attention layers in the attention blocks use the mechanism shown in Figure 5.3 and are followed by n_{dense} linear layers. Meanwhile, the final cross-attention pooling layer in the classification network follows the structure of Figure 5.4 with a trainable $6 \times d$ query tensor with linear regularisation loss. Afterwards, a linear layer without bias and activation performs a reduction over the d latent features to the six output nodes. The structure of the pairing layer in the reconstruction Transformer follows Figure 5.5. A linear regularisation loss is applied to the symmetric tensor in the pairing layer. Residual connections in the attention blocks are shown to the left of the layers contained in the blocks. The dropout layers are only used during training. Hyperparameter values found to yield optimal performance are specified in Table 7.2.

number of latent features, the embedding also combines and transforms the inputs in a fashion useful for the task at hand. Although it is possible to vary the number of latent features throughout the network after the initial coupling, it is held constant in the analysis Transformers. Following the embedding, *layer normalisation* [411] is performed across the latent features for additional regularisation: layer normalisation rescales all latent features to restore a normal distribution across each training batch, so that the inputs to the following network layers always lie in a similar range. This procedure stabilises the training when non-linear activations are involved.

After the embedding operation, the latent feature representation is fed into a series of n_{attblock} multi-headed attention blocks. Each attention block begins with a self-attention layer (c.f. Figure 5.3) with h heads, as described in Section 5.3, followed by a single linear layer without bias and a linear activation function. After the attention layer, a so-called *dropout* layer [412] is employed to reduce the network's reliance on individual latent features by randomly setting each feature to zero with a certain probability. Dropout effectively results in the training of a stochastic ensemble of similar models to avoid overfitting on specific patterns in the training data. No dropout is applied during model inference. Following the dropout, an additional MLP with n_{dense} linear layers and a final dropout layer further combines and transforms the latent features. The final linear layer in the block uses a linear activation. Additionally, input and output of each attention block are connected directly using *residual connections* [413], which directly feed the block input behind the attention layer and to the end of the attention block. Residual connections improve the gradient flow to layers deep in the network, which are difficult to optimise due to vanishing gradients from later layers otherwise. After each recombination of the block and the residual connection, layer normalisation is performed to keep the latent features around $O(1)$. Since the *swish* activation function [414], defined as

$$\text{swish}(x) = x \cdot \text{sigmoid}(x) = \frac{x}{1 + e^{-x}},$$

was found to perform better than the ReLU² function used in Reference [311] in initial tests, it is used for all linear layers where no linear activation is used.

7.1.3. Classification Layer

Event classification with a Transformer network presents the difficulty of turning a variable number of input objects into a fixed number of event classes. As a solution, each latent feature of the final attention block is pooled across all n_{obj} objects. The six classification scores for $t\bar{t}H$, $t\bar{t} + \geq 2b$, $t\bar{t} + 1b$, $t\bar{t} + 1B$, $t\bar{t} + \geq 1c$, and $t\bar{t} + \text{light}$ are derived from these pooled features. Finally, six class probabilities p_i for the classes i can be obtained by feeding the scores through a softmax function as the canonical link for multi-class classification. In the MK1 Transformer, the pooling was

² The *rectified linear unit* is an activation function defined as $\text{relu}(x) = \max(0, x)$.

performed via a simple max-pooling operation, in which the maximum value across all objects was chosen for each latent feature. For the Mk3 networks, the pooling operation was switched to a cross-attention layer (c.f. Figure 5.4) with slightly better performance. The resulting cross-attention classification layer uses a trainable $6 \times d$ query matrix in the cross-attention layer, followed by a single linear layer without bias terms using a softmax activation function.

7.1.4. Higgs Boson Transverse Momentum Reconstruction

The inference of the Higgs boson p_T was tested both via a direct regression using pooling with a single output node and an MSE loss and via indirect reconstruction by combining the four-vectors of the two jets most likely to be the Higgs decay products. The latter method employs a pairing layer introduced in Section 5.3 to extract the most likely jet combination based on the latent features of the jet objects.

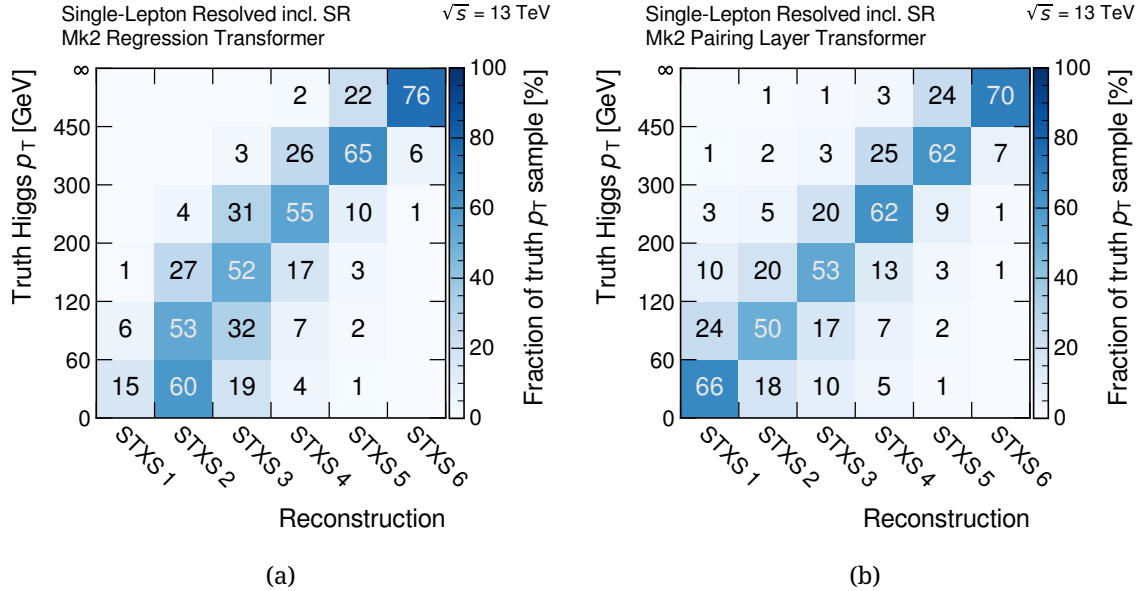


Figure 7.2: Comparison of STXS migration matrices in the SR of the single-lepton resolved channel for (a) the direct p_T^H regression and (b) the pairing layer p_T^H reconstruction methods outlined in Section 7.1.4. Both models were trained during the Mk2 retraining campaign. The values displayed in the matrices are normalised row-wise, so that each value represents the fraction of events in some truth STXS bin falling into the individual reco STXS bins.

In the Mk1 reco Transformers, the regression method was utilised. During the Mk2 campaign, a direct comparison of the regression and the pairing layer reconstruction performance was performed. For this validation, the reconstructed p_T^H was compared against its parton-level truth analogue in the $t\bar{t}H$ Mc samples using so-called *confusion matrices* as shown in Figure 7.2 for the SRs in the single-lepton resolved channel. The regression method was found to skew the p_T^H prediction to-

wards the second and third lowest STXS bins at $60 \text{ GeV} \leq p_T^H < 200 \text{ GeV}$ resulting in asymmetric off-diagonal elements for the outer STXS bins. Meanwhile, the p_T^H inference of the pairing layer model was more aligned with the main diagonal of the confusion matrix. Therefore, the indirect pairing layer inference method was chosen for the p_T^H reconstruction in the analysis. This reconstruction method additionally has the advantage of being more easily interpretable compared to the direct p_T regression, since it is fundamentally based on the physical objects in the event and allows additional validation of the jet matching as performed in Section 7.4 for the Mk3 models.

While the Mk2 pairing layer was based on a direct pairing layer described via Equations (5.12) and (5.13), the layer was refined by an additional trainable symmetric tensor for the Mk3 networks. The resulting tensor pairing layer (c.f. Figure 5.5) is very similar to the permutation-invariant tensor attention employed by SPANet [323] and was found to improve training performance with respect to direct pairing. Additionally, the reco-level STXS bins edges were optimised with respect to the amount of correctly assigned $t\bar{t}H$ events by designating the intersection points of the normalised reco p_T^H distributions for each truth-level STXS sample as the shifted bin edges. The corresponding values are listed in Table 7.3.

Table 7.3: Optimised reco-level STXS bin edges for the Mk3 reconstruction models compared to those of the stage 1.2 STXS formalism.

Channel		STXS Bin Edges [GeV]					
Stage 1.2 STXS	0	60	120	200	300	450	∞
Single Lepton Resolved Dilepton	0	60	114	192	282	408	∞
				186	270	402	

7.2. Network Training

7.2.1. General Training Setup

The networks are trained separately for the single-lepton and the dilepton channel on simulated events passing the single-lepton resolved or dilepton pre-selection using Keras [415] with the TensorFlow backend [416]. Model inference is performed using ONNXRuntime [417] in the TTHbbAnalysis framework. The specific $t\bar{t}(H \rightarrow b\bar{b})$ Transformer implementation is contained in a dedicated repository [418].

The classification networks are trained using $t\bar{t}H$, $t\bar{t} + \geq 1b$, and $t\bar{t}$ samples with training and validation batch sizes of 2048. A list of samples used in the training for each category, together with the total amount of Mc events available per sample, is provided in the appendix. In order to facilitate the training of batches with a variable

number of objects, the feature arrays are padded and empty jet objects masked during training to allow up to 11 (10) jets for the training in the single-lepton (dilepton) channel. For events with more jets, only the leading jets ordered by their PCBT bin and subsequently by their p_T are used during training³. No MC event weights are taken into account during the training⁴. As a result, the network training is not entirely optimal for data inference because of phase-space effects.

Events in the $t\bar{t}$ + jets samples are classified into five categories using the procedure described in Section 6.2.5 for the training of the event classification networks. The resulting six event categories to be predicted by the networks are balanced in each training and validation batch by drawing one sixth of every batch from each of the event categories⁵. Simultaneously, events are shuffled after each training iteration, while ensuring that each event is used an equal number of times. Meanwhile, the reconstruction networks are trained in batches of size 2048 on $t\bar{t}H$ events with two distinct reco-level jets truth-matched to the b -hadrons originating from the Higgs decay as determined from the parton history. The JetTruthMatchingTool [419] tool in the TTHbbAnalysis framework carries out the b -hadron parton assignment and performs the truth matching with a maximum separation of $\Delta R = 0.3$ between the hadron and the jet axis. Events with truth p_T^H falling into the six different STXS bins are balanced during the reconstruction Transformer training in an analogous way to the event classes in the classification Transformer.

Five-fold *cross-validation* is performed on the MC samples used for training by partitioning the samples into five parts based on their event number modulo five. During inference, events from data were equally split into five folds based on their event number modulo five, such that the inference of MC and data events is performed consistently in the analysis workflow using the network trained on the respective fold. To further enhance training statistics without altering kinematic properties of the physics processes, the MC events used for training are augmented by two transformations: the first of them consists of a random rotation of the whole event by an azimuthal angle ϕ around the beam pipe. The second augmentation consists of a beam inversion achieved by randomly mirroring the y - and z -coordinates simultaneously to preserve parity. During the Mk1 and Mk2 iterations, an additional random Lorentz boost along the beam was employed. The boost augmentation was removed for the Mk3 retraining, as it was found to produce highly unphysical results and could not be motivated by general phase-space considerations.

Both classification and Higgs boson p_T reconstruction use a categorical cross-entropy loss with *label smoothing* [420] of 10 % (meaning that 10 % of the truth label are distributed equally among the other labels to smooth the training process). Linear regularisation losses are included for the trainable tensors in the cross-attention pooling

³ While the Transformer is permutation invariant, the jet ordering has to be defined in the training because of the maximum jet cutoff used. No jet multiplicity cuts are employed during inference.

⁴ The usage of weights in the loss to more accurately reflect kinematic distributions in the analysis phase space was tested, but led to unstable training results.

⁵ As a result, events from categories with lower MC statistics are used more often during the training.

and the pairing layer of the classification and reconstruction models, respectively. To update model parameters during backpropagation, the AdamW optimiser [322, 421] with weight-decay $\lambda = 1 \times 10^{-6}$ and first and second moment decay rates of $\beta_1 = 0.9$ and $\beta_2 = 0.999$, respectively, is used following the notation of Reference [421]. Additionally, an *early stopping* procedure is adopted to terminate the network training and restore the best model after ten epochs⁶ of no improvement in the validation loss, thus avoiding model overtraining. The loss and accuracy curves of the Mk3 networks show no overtraining and are included in the appendix.

7.2.2. Hyperparameter Optimisation

An extensive programme to optimise the Transformer hyperparameters was carried out during the Mk3 retraining campaign. In a first step, the flat learning rate used for the Mk1 and Mk2 models was replaced by an annealing-like learning rate schedule to improve training performance [422, 423]. The schedule consists of a burn-in period at a low learning rate of 1×10^{-8} , a cosine ramp-up to a high learning rate plateau at 3.5×10^{-5} , and finally a slow ramp-down to a rate of 1×10^{-7} to allow the model to slowly settle in a loss minimum after exploring the parameter phase space during the plateau at a high rate. The burn-in period helps to settle down the Transformer parameters after their random initialisation, which leads to a more stable training.

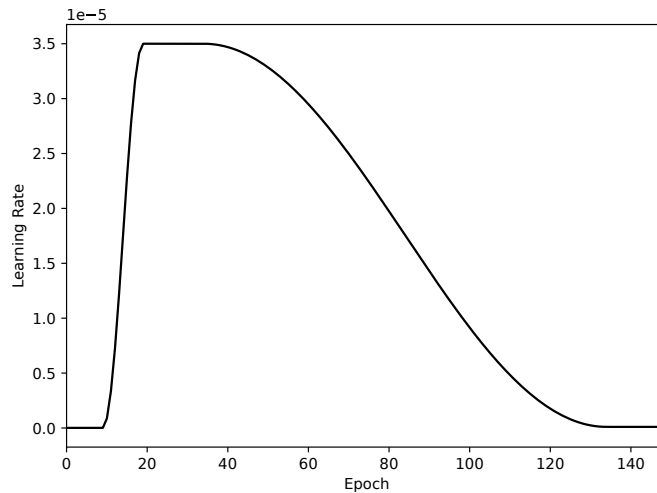


Figure 7.3: Profile of the learning rate annealing used for the Mk3 retraining campaign. Both ramp-up and ramp-down follow cosine curves.

The lengths of the annealing periods as well as the different learning rates were optimised based on training performance and speed, leading to the schedule shown in Figure 7.3. The precise hyperparameters are listed in the appendix. After the learning rate optimisation, other hyperparameters of the Mk3 models were optimised

⁶ An epoch marks the period during which each event of the event class with the highest Mc statistics is used once during the Transformer training.

separately in several grid searches via the final validation losses and under consideration of signs of overtraining or mode collapse⁷ in the training curves. The final Mk3 Transformers encompass between 9.8×10^5 trainable parameters for the dilepton reconstruction models and 6.7×10^6 trainable parameters for the single-lepton classification models.

7.3. Transformer Performance

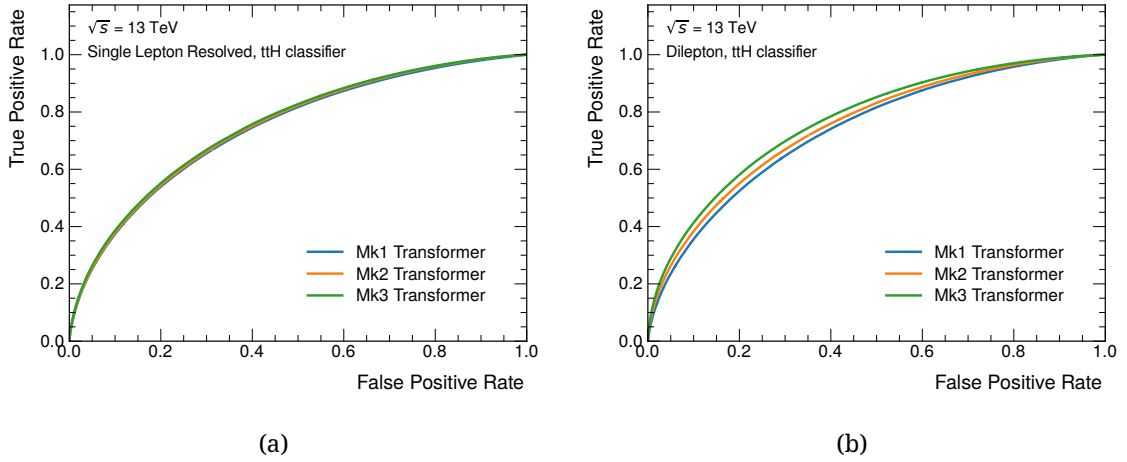


Figure 7.4: ROC curves of the event classification Transformer in the analysis (a) single-lepton resolved and (b) dilepton pre-selection regions for a comparison of $t\bar{t}H$ against $t\bar{t} + \geq 1b$ events. Signal classification is performed based on the fraction classifier $f_{t\bar{t}H}$, except for the Mk3 model, which uses the discriminant classifier $d_{t\bar{t}H}$. Performance gains with every retraining campaign are visible, especially in the dilepton channel.

The performance of the Transformer networks is evaluated on all $t\bar{t}H$ and $t\bar{t} + \text{jets}$ events fulfilling the kinematic pre-selection of the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis. For the classification models, Receiver-Operating Characteristic (ROC) curves⁸ are used as the basic tool for estimating the classification performance. Additionally, confusion and event yield matrices of the individual classes are used to judge the full classification performance based on the chosen classifier cuts. Figure 7.4 shows a comparison of the ROC curves for the model iterations when comparing the $t\bar{t}H$ classifier against the $t\bar{t} + \geq 1b$ background. Areas under the ROC curves (AUCs) of 0.753 in the single-lepton and 0.774 in the dilepton were achieved by the Mk3 models. Meanwhile, an overview of the Mk3 ROC curves for all target classes is shown in Figure 7.5. For the reconstruction models, the STXS confusion matrices, shown

⁷ This was a failure mode observed in some early dilepton reconstruction models with a non-optimal pairing layer, where the validation accuracy would suddenly decrease rapidly.

⁸ ROC curves plot the fraction of correctly assigned target events of a classifier, called the *true positive rate*, against the *false positive rate*, the fraction of other events erroneously assigned to the target class, in dependence on the classifier threshold. A high true positive rate at a low false positive rate is desired for good model performance.

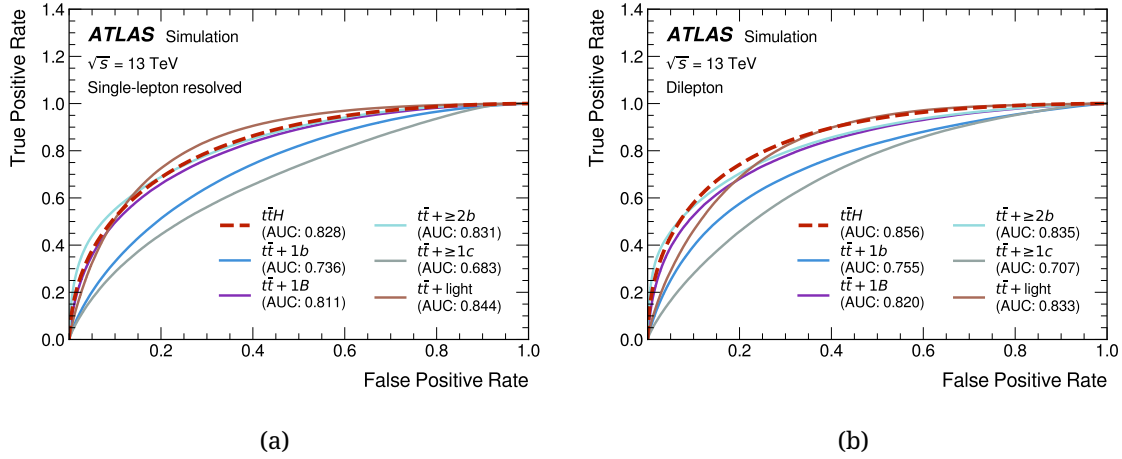


Figure 7.5: ROC curves of the event classification Mk3 Transformer in the analysis (a) single-lepton resolved and (b) dilepton pre-selection regions for all discriminants. The classification is performed using the discriminant classifiers d_i for every target class i , where the target class is evaluated against events of all other classes $j \neq i$.

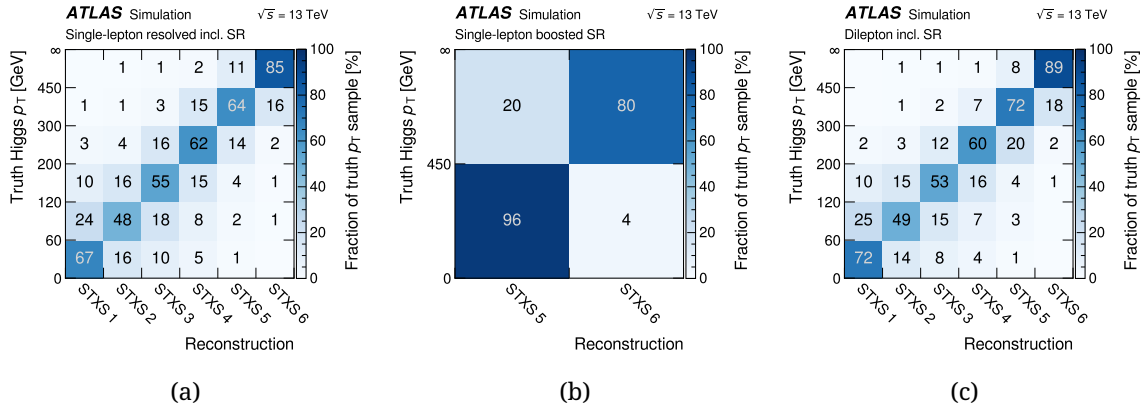


Figure 7.6: STXS confusion matrices for the Mk3 reconstruction models in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy (a) single-lepton resolved, (b) single-lepton boosted, and (c) dilepton SR. The bin edge optimisation introduced in Section 7.1.4 is applied.

in Figure 7.6 for the Mk3 models, form the primary performance metric. Additionally, the bin purities, defined as the number of correct STXS assignments divided by the total number of assignments to the same reco STXS bin, are listed for the Mk3 networks in Table 7.4.

7.3.1. Event Classifier & Signal Region Optimisation

The direct usage of the event classification probabilities p_i in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis is problematic because of significant yield differences between the $t\bar{t}H$ signal and various $t\bar{t}$ + jets background processes. To take into account these

Table 7.4: p_T^H reconstruction purity of the MK3 models in the analysis SRs. The first five STXS bins are taken together in the single-lepton boosted SR to be consistent with the analysis fit setup.

p_T^H Range [GeV]	Single Lepton		Dilepton
	Resolved	Boosted	
0 – 60	0.48		0.48
60 – 120	0.58		0.57
120 – 200	0.53	0.95	0.52
200 – 300	0.46		0.43
300 – 450	0.37		0.39
450 – ∞	0.32	0.80	0.39

yield differences and maximise the classifier sensitivity, two different modifications of p_i were tested.

The first modification, used for the MK1 and MK2 models, takes the predicted process yields⁹ N_i scaled to data into account as Bayesian priors for so-called *fraction* classifiers f_i as

$$f_i = \frac{p_i \cdot N_i}{\sum_j p_j \cdot N_j}. \quad (7.1)$$

The fractions retain the property $\sum_j f_j = 1$ of the class probabilities p_i , so that they can still be interpreted directly as probabilities. However, large imbalances in the fraction distributions with $f_{t\bar{t}H} \ll f_{i \neq t\bar{t}H}$ were observed as a result of the significant yield imbalance $N_{t\bar{t}H} \ll N_{i \neq t\bar{t}H}$, so that $t\bar{t}H$ and other small samples, such as $t\bar{t} + 1B$, are penalised by the fraction classifiers.

Therefore, discriminant classifiers d_i were included in the MK3 models as a second modification of p_i . Instead of the unnormalised yields used for the fractions, they make use of normalised yield fractions as follows:

$$d_i = \frac{p_i}{\sum_{j \neq i} p_j \cdot \hat{w}_{ij}}, \quad \text{where } \hat{w}_{ij} = \frac{N_j}{\sum_{k \neq i} N_k}. \quad (7.2)$$

As a result, the denominator of each discriminant d_i is a weighted average of all remaining class probabilities $p_{j \neq i}$, where the associated yield fractions $\hat{w}_{ij}$ exclude the yield of process i . Therefore, $\hat{w}_{ij}$ can be thought of as an effective event fraction of process j in the ‘background’ of process i . A similar discriminant is used in the b -tagging discriminants of the DL1r tagger [324]. Because of the property $\sum_{j \neq i} \hat{w}_{ij} = 1$ of the yield fractions, the balancing problems of f_i do not arise in d_i . However, $\sum_j d_j \neq 1$ means that the discriminants can no longer be identified directly as class probabilities.

⁹ Calculated from the predicted sample cross-sections, the luminosity, and the $t\bar{t}$ + jets scale factors derived in Section 6.5.1.

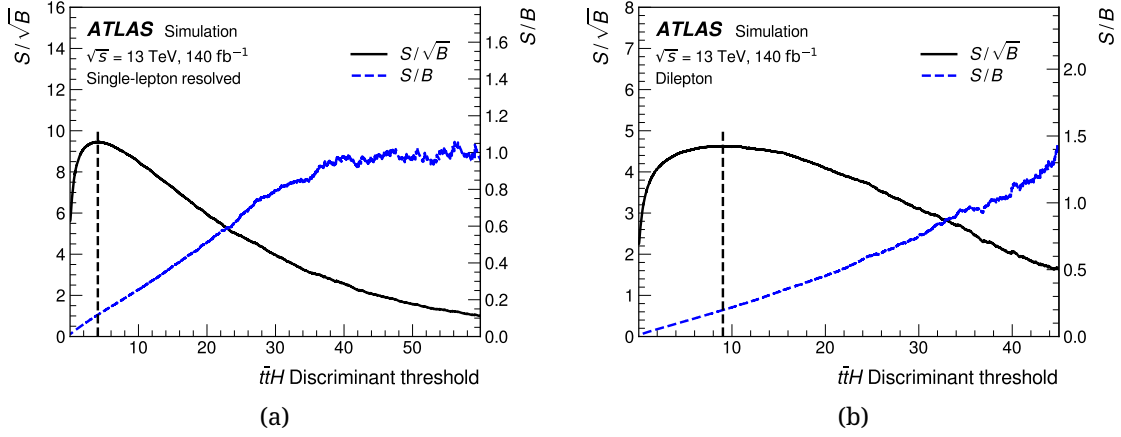


Figure 7.7: Dependence of the signal-over-background noise ratio $S/\sqrt{B}$ and of the signal-over-background ratio S/B on the $Mk3$ $t\bar{t}H$ discriminant classifier threshold in the (a) single-lepton resolved and (b) dilepton channel. Optimal thresholds were determined by finding the maxima of $S/\sqrt{B}$ in each case, while the distribution of S/B did not help to define usable thresholds. The optimal thresholds are plotted as dashed vertical lines in each plot.

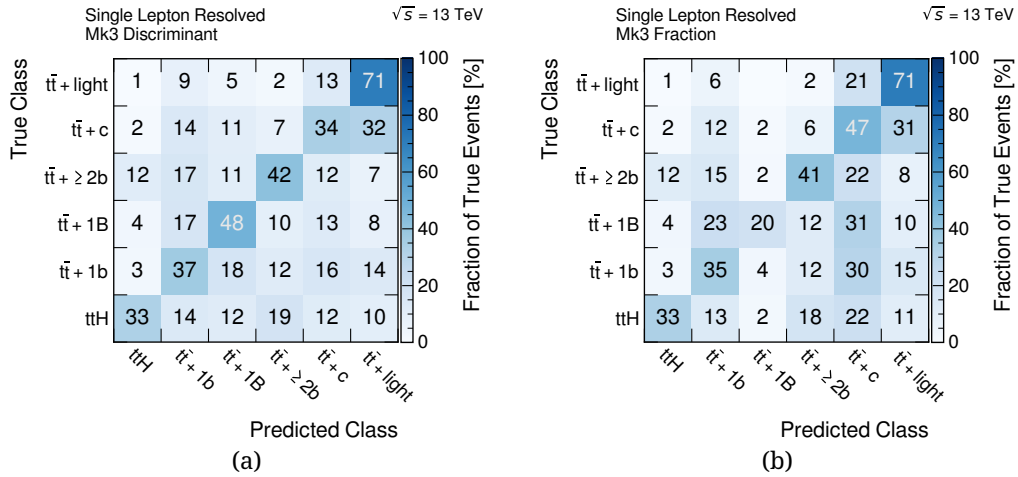


Figure 7.8: $Mk3$ classification confusion matrices at the optimised $t\bar{t}H$ (a) discriminant and (b) fraction classifier thresholds in the single-lepton resolved channel. The matrices are normalised per row. While the classification into the $t\bar{t}H$ region is very similar between the discriminant and the fraction approach, the background classification is significantly different. Analogous matrices for the dilepton channel are contained in the appendix.

In general, region assignments are done based on the maximum classifier. However, a minimum $t\bar{t}H$ classifier threshold $d_{t\bar{t}H}^{\min}$ was instead defined for the SRs to increase control over the SR composition. The optimal threshold is chosen to maximise the signal-over-background noise ratio $S/\sqrt{B}$ in the resulting signal region. Results from the discriminant classifier optimisation lead to thresholds of $d_{t\bar{t}H}^{\min} = 4.072$ in the single-lepton resolved and $d_{t\bar{t}H}^{\min} = 9.031$ in the dilepton channel. The corresponding threshold scans are plotted in Figure 7.7. Due to the small $t\bar{t}H$ yield, $f_{t\bar{t}H}$ and $d_{t\bar{t}H}$ perform very similar during the threshold optimisation¹⁰, while the CR composition shown in Figure 7.8 for the single-lepton resolved channel varies between the two classifier strategies. Therefore, dedicated fit comparisons with both classification strategies were carried out. The discriminant classifiers were eventually chosen for the analysis due to improved expected $t\bar{t}(H \rightarrow b\bar{b})$ discovery significance and a more balanced background event distribution, leading to more stability. In the fit setups presented in Chapter 8, the discriminants are rescaled to the interval $[0, 1)$ as d'_i according to logistic functions

$$d'_i = 1 - e^{-m_i d_i}, \quad (7.3)$$

where the parameters m_i are manually optimised. As the rescaling is performed after events are assigned to the analysis regions, it is purely visual.

7.4. Model Comparison

7.4.1. Comparison in the First Full Run 2 Analysis Signal Region

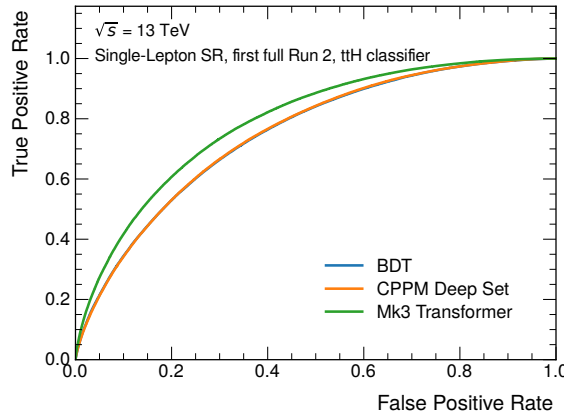


Figure 7.9: Comparison of the ROC curves of different network architectures in the single-lepton SR of the first full Run 2 $t\bar{t}(H \rightarrow b\bar{b})$ analysis. Only $t\bar{t} + \geq 1b$ events are considered for the background. The BDT and Deep Set curves lie on top of each other with an AUC of 0.751, while the Mk3 Transformer outperforms both with an AUC of 0.791.

¹⁰ $f_i \approx d_i \cdot N_i / \sum_{j \neq i} N_j$ if there is a yield imbalance with $N_i \ll \sum_{j \neq i} N_j$.

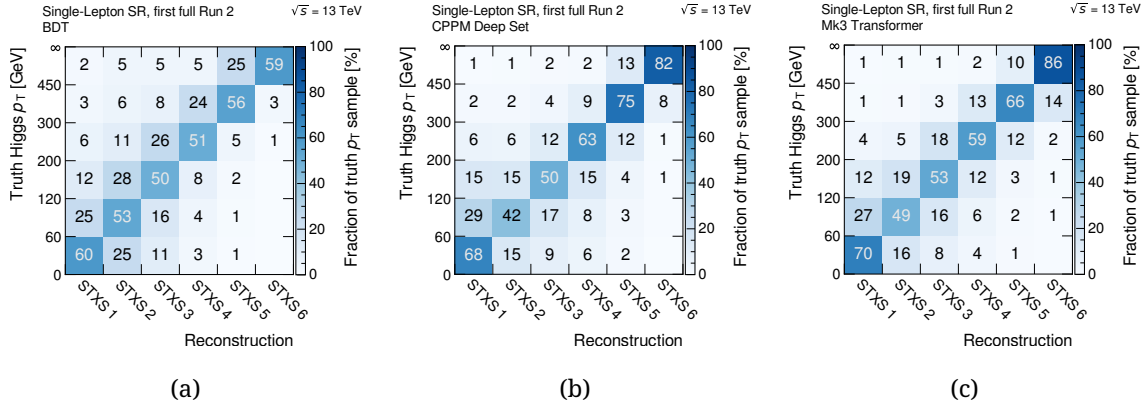


Figure 7.10: Comparison of the STXS confusion matrices for (a) the BDT of the previous analysis round, (b) the CPPM Deep Set, and (c) the MK3 Transformer in the single-lepton SR of the first full Run 2 $t\bar{t}(H \rightarrow b\bar{b})$ analysis.

In addition to the performance validation of the Transformers, the final Mk3 models used in the analysis can also be benchmarked against the BDT-based models used in the first full Run 2 $t\bar{t}(H \rightarrow b\bar{b})$ analysis [211] and the Deep Set developed at CPPM [410] on high-level event features as an alternative to the Transformers. The phase space of the single-lepton SR of the previous analysis was restricted to ≥ 6 jet and ≥ 4 jets b -tagged at the 70 % working point¹¹ and the CPPM DeepSet was also trained in this region. Therefore, the comparisons in this section use the same event selection.

In the ROC curve comparison of Figure 7.9, the BDT and Deep Set perform similarly, while the Mk3 classification Transformer clearly outperforms both models. Meanwhile, the STXS confusion matrices of the different reconstruction models are shown in Figure 7.10. Both BDT and Mk3 Transformer operate by determining the most likely Higgs boson decay products to reconstruct p_T^H . Meanwhile, the Deep Set directly classifies events into the six STXS bins in a multi-class approach. The Mk3 model and the Deep Set outperform the BDT. Overall, the Mk3 model also performs better than the Deep Set except in the region of $200 \text{ GeV} \leq p_T^H < 450 \text{ GeV}$.

7.4.2. Comparison with SPANet

The most recent SPANet iteration presented in Reference [424] can take lepton objects and global features, including E_T^{miss} , into account via heterogeneous embeddings. Therefore, the Higgs boson assignment accuracies, defined as the fraction of events with two correctly assigned Higgs decay b -jets, of SPANet and the Mk3 reconstruction Transformer in semileptonic $t\bar{t}(H \rightarrow b\bar{b})$ events are compared in Table 7.5. The reconstruction Transformer performance is evaluated in the SR of the first full Run 2 $t\bar{t}(H \rightarrow b\bar{b})$ analysis to approximate the SPANet event selection. Overall, the Mk3

¹¹ The DL1r b -tagger is used for the comparisons instead of the Mv2c10 b -tagger used in the first full Run 2 $t\bar{t}(H \rightarrow b\bar{b})$ analysis.

Transformer slightly outperforms SPANet in every jet multiplicity bin, although it should be noted that the underlying validation datasets are not entirely identical.

Table 7.5: Comparison of the Higgs boson reconstruction performance of the Mk3 model with SPANet in the single-lepton SR of the first full Run 2 $t\bar{t}(H \rightarrow b\bar{b})$ analysis split by jet multiplicity. Only events where both b -quarks coming from the Higgs decay can be truth-matched to reco-level jets are considered. For a successful match, both truth-matched jets have to be correctly identified. For the Mk3 model, the fraction of partially reconstructed events is provided in addition.

Model	Higgs Boson Assignment Accuracy [%]							
	6 Jets		7 Jets		≥ 8 Jets		Combined	
	Full	Partial	Full	Partial	Full	Partial	Full	Partial
SPANet	50	–	49	–	47	–	49	–
Mk3 Transformer	55	28	53	29	49	30	53	29

Using the analysis strategy outlined in Section 6.5, the Mc samples presented in Section 6.2 can be fitted to the collision data collected during Run 2 of the LHC by the ATLAS detector. To arrive at a meaningful statement about the degree of certainty on the analysis results, systematic uncertainties on the detector and the Mc modelling of signal and background processes have to be considered in addition to the statistical uncertainties of the collected data and the Mc samples. The systematics model employed in the profile likelihood fits, motivated by fit studies performed in the course of the analysis, will be introduced in Section 8.1. Since the fit optimisation was not the primary focus of this thesis, only a summary of the pre-fit modelling in the signal and control regions defined in Section 6.5 is shown in Section 8.2. Finally, the results of inclusive $t\bar{t}H$ cross-section and signal strength measurements, as well as a measurement of the cross-sections in the STXS bins, are presented in Section 8.3 and compared to previous results afterwards.

8.1. Systematic Uncertainties

In the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis, systematic uncertainties arise from the modelling of signal and background processes, and from detector effects. Additionally, systematic uncertainties can affect both the normalisation of process yields and the shapes of their variable distributions.

8.1.1. Experimental Systematics

The experimental systematics cover all of the objects reconstructed in the detector, as well as the luminosity and pile-up modelling. For the luminosity, a flat uncertainty of 0.83 % is employed following the calibration procedure described in Reference [277]. The uncertainty inherent in the pile-up modelling described in Section 6.2 is parametrised using a variation in the pile-up reweighting procedure corresponding to a per-event scale factor.

For electrons and muons, the uncertainties can be divided into components arising from the triggers, the lepton reconstruction, identification, isolation, and momentum scale and resolution. To account for differences between Mc simulation and data for leptons, scale factors are derived for the Mc events using tag-and-probe techniques on $Z \rightarrow \ell^+\ell^-$ and $J/\psi \rightarrow \ell^+\ell^-$ events [390, 391, 425]. The uncertainties in these scale factors are directly propagated to event weights, resulting in four independent components for the electrons and ten for the muons. Similarly, the lepton momentum scale and resolutions of simulated events are corrected to match those in data [426]. The associated uncertainties are evaluated by varying the scale within its uncertainty, while the resolution uncertainties are estimated by smearing the lepton energy and momentum. This procedure results in two independent components for electrons and four for muons. The muon identification uncertainties are also applied as independent systematics to the softmuons used in the muon-in-jet corrections, while the softmuon energy scale and resolution uncertainties are correlated with those of the muons.

Systematic uncertainties in the jet energy scale consist of contributions from jet-flavour composition, η -intercalibration, punch-through, single-particle response, the response of the calorimeters to jets of different flavours, pile-up effects, and effects from in-situ jet calibration techniques. Altogether, these give rise to 30 uncorrelated systematics, where the uncertainties originating from the in-situ calibration are combined into eight effective components using a Principal Component Analysis (PCA) in the so-called *category reduction* scheme [427]. Meanwhile, uncertainties in the jet energy resolution arise from differences between in-situ measurements of the jet energy resolution in data and Mc events. These differences are split into 13 uncorrelated components to smear Mc samples or data as appropriate. Additionally, the uncertainty associated with the jet vertex tagger discriminant is estimated via variations of the Mc efficiency between different Mc generators [387]. For the uncertainty treatment of the reclustered jets used for the boosted Higgs candidates, the systematic uncertainties of their sub-jets are propagated.

Uncertainties associated with the calibration of the DL1r b -tagger are derived on a jet-flavour basis for light, c -, and b -jets. They are divided into 20, 20, and 45 uncorrelated components, respectively, via a PCA [428–430]. Additionally, extrapolation uncertainties are derived for high- p_T jets not covered by the calibration range. This results in a further four (two) NPs each for b - and light jets (c -jets).

The uncertainties in E_T^{miss} are determined by propagating the energy scale and resolution uncertainties of all other detector objects. Three additional uncertainty components estimated in $Z \rightarrow \ell^+ \ell^-$ events arise from the modelling of the E_T^{miss} soft term [394].

8.1.2. Modelling Systematics

In principle, all of the uncertainties mentioned in Section 6.2 are considered for the modelling of signal and background processes. For the $t\bar{t}H$ signal, these include the cross-section uncertainties accounting for the QCD scale variations as well as those of the PDFs and α_S . Additionally, theory uncertainties originating from the migration of events between the different truth p_T^H STXS bins [431] are considered in the inclusive cross-section and signal strength measurement, but not in the differential cross-section measurement. A systematic uncertainty of 2.2 % is assigned for the $H \rightarrow b\bar{b}$ BR [168]. Additionally, the uncertainties due to varying μ_R and μ_F , ISR and FSR radiation from the VAR3c variations, and the alternative samples for the Ps and the NLO matching are considered for $t\bar{t}H$.

For the $t\bar{t}$ + jets components, all systematic uncertainties are assigned independent NPs in each component. No normalisation effects due to the modelling uncertainties are included since the background normalisation factors are free-floating in the likelihood fits. Therefore, all $t\bar{t}$ + jets systematic uncertainties are reweighted to the same yield as the nominal samples in the analysis pre-selection defined in Section 6.3.3. All modelling uncertainties introduced in Section 6.2.3 are used for the modelling of the $t\bar{t}$ + jets components. Uncertainties from the H_T reweighting of the $t\bar{t} + \geq 1c$ and $t\bar{t}$ + light background processes introduced in Section 6.5.1 are included independently per channel via alternative reweightings in which the $t\bar{t} + \geq 1c$ and $t\bar{t}$ + light normalisation factors are varied within their respective uncertainties. The $t\bar{t}$ + jets distributions without H_T reweighting for each jet multiplicity bin are used as additional uncertainties for the procedure.

Meanwhile, the modelling uncertainties of tW , $t\bar{t}W$, $t\bar{t}Z$ are based on those introduced in Section 6.2.4. The cross-section uncertainties of the background processes correspond to those listed in Table 6.3. Independent normalisation uncertainties of 50 % are considered for the single-lepton and dilepton fake estimates. Additionally, an alternative parametrisation of the single-lepton fake estimate is used to cover uncertainties related to the choice of fake factor parametrisation.

8.2. Pre-Fit Modelling

Prior to the fitting procedure, the pre-fit yields for the different MC processes were calculated. For this, the cross-sections determined theoretically under SM assump-

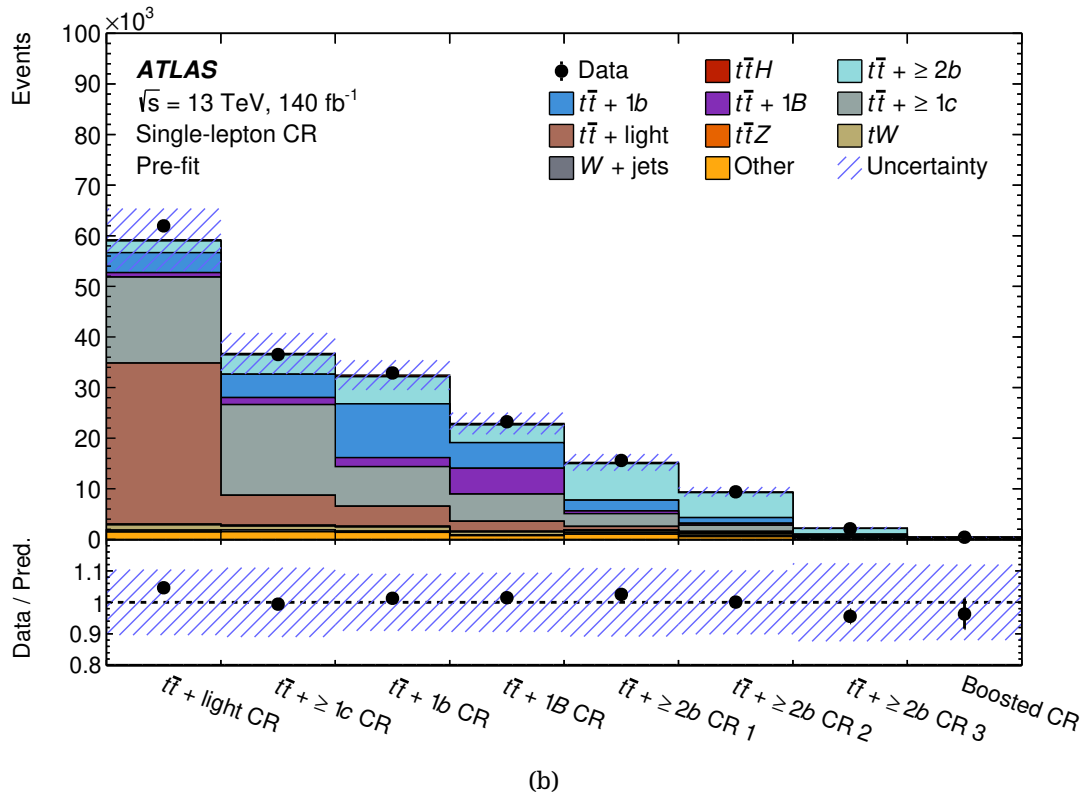
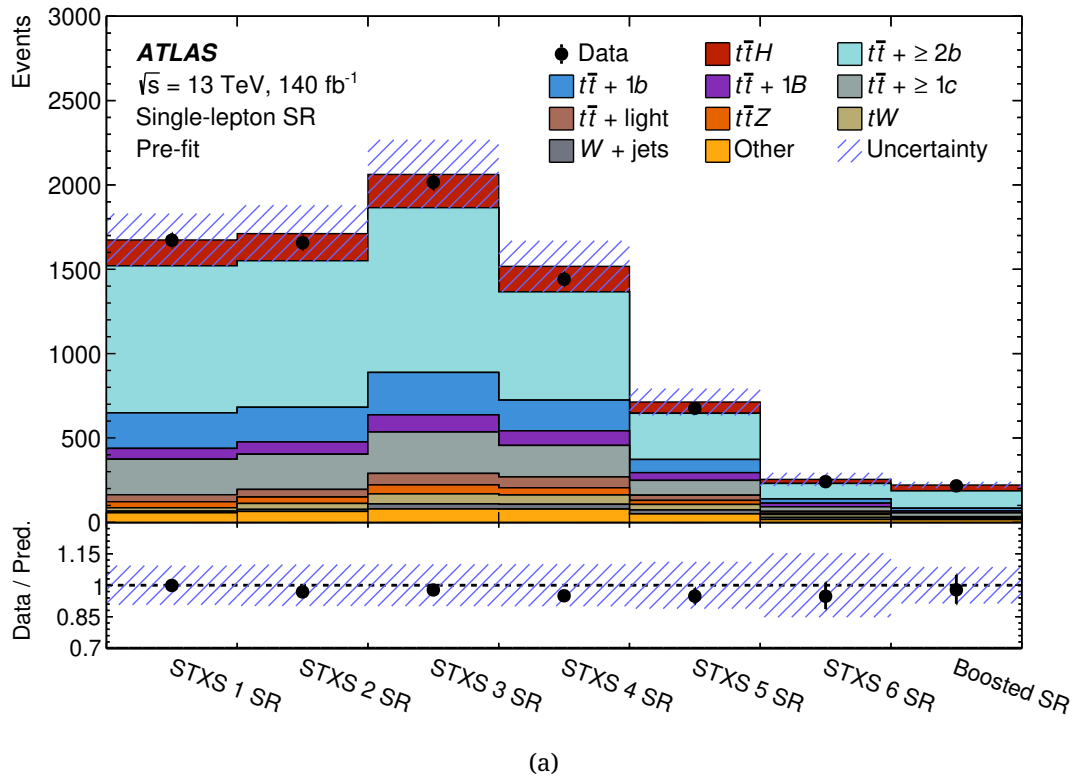
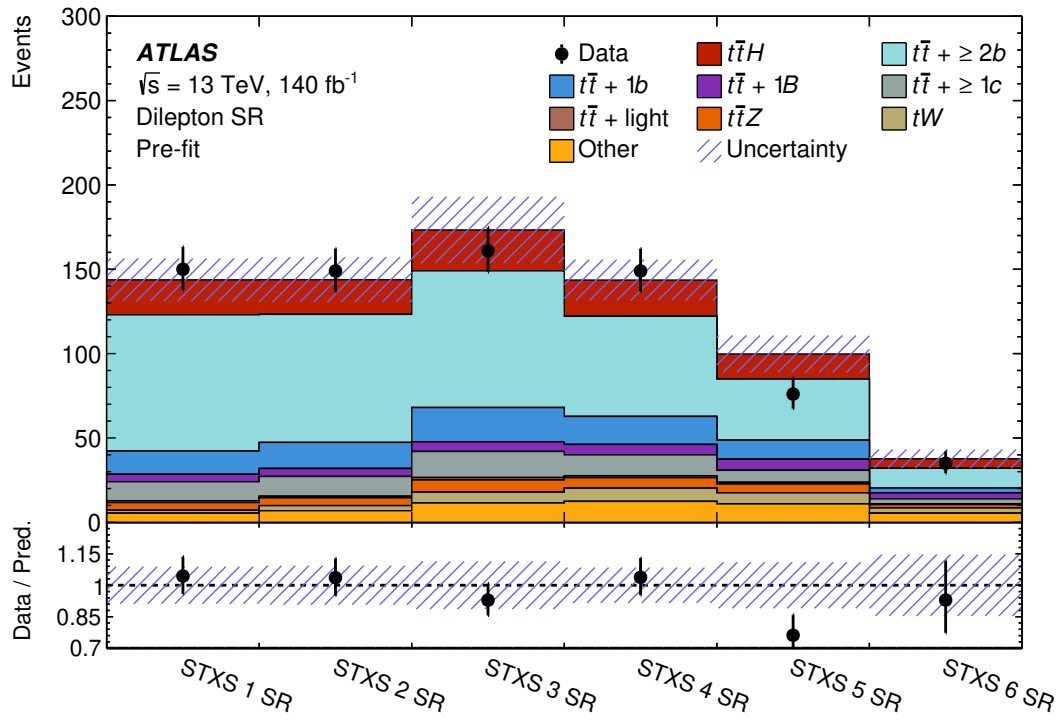
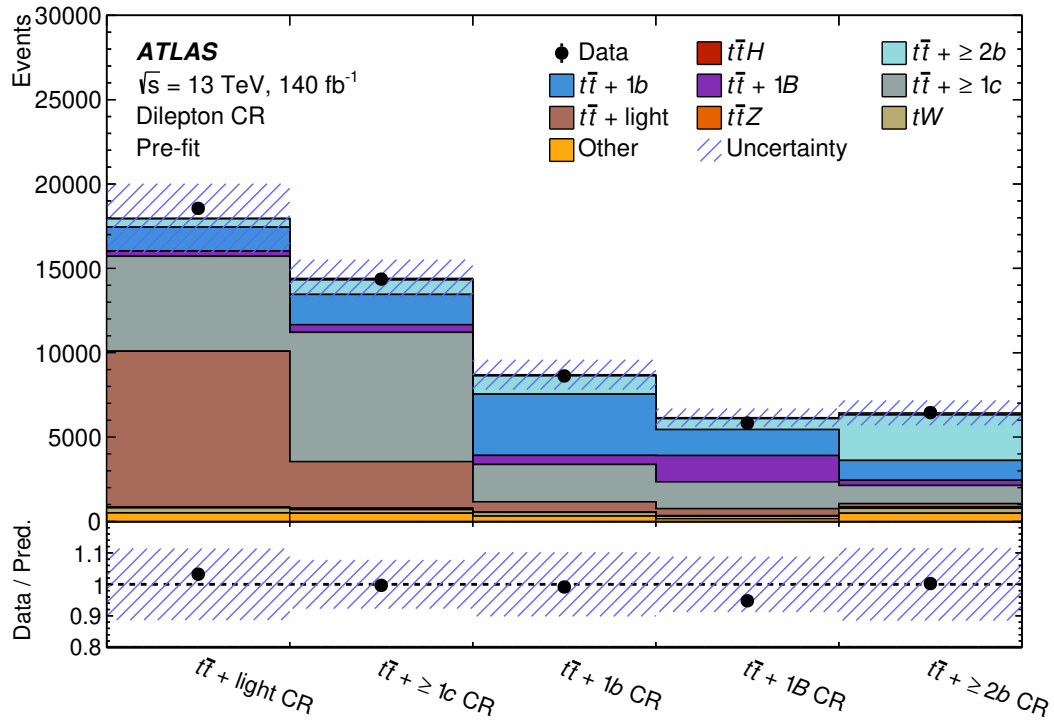


Figure 8.1: Log-scale pre-fit summary histograms of the event yields in (a) the SRs and (b) CRs of the single-lepton channel, where each region is aggregated into a single bin. The uncertainty band includes all uncertainties considered in the analysis.



(a)



(b)

Figure 8.2: Log-scale pre-fit summary histograms of the event yields in (a) the SRs and (b) the CRs of the dilepton channel, where each region is aggregated into a single bin. The uncertainty band includes all uncertainties considered in the analysis.

Chapter 8. Results

tions, as summarised in Section 6.2, and the Run 2 luminosity were used to normalise the yields to those expected in the data under consideration of reweighting factors originating from both the Mc modelling and the object reconstruction calibration. Additionally, the $t\bar{t}$ + jets backgrounds are rescaled by the normalisation factors introduced in Section 6.5.1. Summary plots of comparisons between data and the Mc processes considered in the analysis prior to the fit procedure are shown in Figures 8.1 and 8.2. Meanwhile, pre-fit yield comparisons between data and Mc processes with symmetrised uncertainties are listed in Tables 8.1 to 8.4. The full effect of all statistical and systematic uncertainties enters the overall uncertainty. The associated uncertainties on the $t\bar{t}$ + jets normalisation factors do not enter the uncertainty bands. Detailed post-fit plots are instead shown in Section 8.3.

Table 8.1: Pre-fit yields in the single-lepton SRs. The s- and t-channel single-top, tZq , and tWZ processes are combined under ‘Other’. The $t\bar{t}H$ samples are split by truth $p_{\perp}^H$ with a separate forward Higgs boson category. All uncertainties except those of the $t\bar{t}$ + jets normalisation factors are included. The yield uncertainties are symmetrised and rounded.

Process	STXS 1 SR	STXS 2 SR	STXS 3 SR	STXS 4 SR	STXS 5 SR	STXS 6 SR	Boosted SR
$t\bar{t}H, p_{\perp}^H \in [0, 60)$ GeV	77 ± 12	18.7 ± 3.2	11.1 ± 2.3	5.7 ± 1.0	1.66 ± 0.20	0.42 ± 0.09	0.58 ± 0.13
$t\bar{t}H, p_{\perp}^H \in [60, 120)$ GeV	49 ± 5	99 ± 13	37 ± 6	15.5 ± 2.2	4.4 ± 0.5	1.25 ± 0.19	1.54 ± 0.25
$t\bar{t}H, p_{\perp}^H \in [120, 200)$ GeV	22.7 ± 2.2	37 ± 4	125 ± 16	34 ± 5	8.1 ± 1.3	2.0 ± 0.4	2.31 ± 0.28
$t\bar{t}H, p_{\perp}^H \in [200, 300)$ GeV	4.0 ± 0.5	5.5 ± 0.7	22.5 ± 2.9	88 ± 12	19.4 ± 3.0	2.7 ± 0.4	3.3 ± 0.6
$t\bar{t}H, p_{\perp}^H \in [300, 450)$ GeV	0.39 ± 0.11	0.64 ± 0.16	1.4 ± 0.4	7.6 ± 1.3	32 ± 5	8.0 ± 1.3	20 ± 3
$t\bar{t}H, p_{\perp}^H \geq 450$ GeV	< 0.1	< 0.1	0.14 ± 0.06	0.21 ± 0.09	1.30 ± 0.31	9.8 ± 1.8	6.9 ± 1.3
$t\bar{t}H, y_H > 2.5$	0.16 ± 0.06	0.10 ± 0.04	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
$t\bar{t} + \geq 2b$	870 ± 80	870 ± 90	980 ± 100	640 ± 60	270 ± 40	92 ± 21	103 ± 8
$t\bar{t} + 1b$	210 ± 60	210 ± 50	250 ± 80	180 ± 60	78 ± 28	25 ± 9	17 ± 5
$t\bar{t} + 1B$	64 ± 19	73 ± 24	100 ± 40	87 ± 30	46 ± 19	22 ± 11	11 ± 4
$t\bar{t} + \geq 1c$	210 ± 50	210 ± 50	250 ± 50	190 ± 50	88 ± 24	27 ± 13	22 ± 8
$t\bar{t} + \text{light}$	42 ± 16	45 ± 12	68 ± 22	64 ± 19	31 ± 11	12 ± 11	3.2 ± 3.0
tW	20 ± 8	35 ± 12	60 ± 33	55 ± 32	33 ± 25	17 ± 15	5 ± 4
$t\bar{t}Z$	36 ± 6	38 ± 6	53 ± 8	43 ± 8	24 ± 5	7.2 ± 1.1	5.8 ± 1.2
$t\bar{t}W$	4.2 ± 0.5	5.9 ± 0.8	9.8 ± 1.6	8.9 ± 1.9	5.6 ± 0.7	3.2 ± 0.5	1.06 ± 0.26
$t\bar{t}t\bar{t}$	5.1 ± 2.2	6.9 ± 3.0	11 ± 5	9 ± 4	4.8 ± 2.1	2.5 ± 1.1	1.1 ± 0.5
$tHjb$	0.77 ± 0.22	1.06 ± 0.21	1.61 ± 0.32	1.58 ± 0.28	1.11 ± 0.30	0.26 ± 0.12	0.72 ± 0.14
tWH	1.68 ± 0.23	2.16 ± 0.27	3.8 ± 0.4	4.1 ± 0.5	2.22 ± 0.25	1.03 ± 0.12	1.01 ± 0.11
$W + \text{jets}$	10 ± 6	12 ± 8	30 ± 15	29 ± 14	23 ± 12	12 ± 6	3.7 ± 1.9
$Z + \text{jets}$	2.5 ± 1.1	5.1 ± 2.0	6.5 ± 2.6	7.4 ± 2.7	3.1 ± 1.1	1.7 ± 0.6	0.5 ± 0.2
Diboson	1.3 ± 0.8	2.5 ± 1.4	4.4 ± 2.3	4.7 ± 2.7	3.1 ± 1.6	1.4 ± 0.7	0.7 ± 0.4
Other	11.3 ± 1.2	11.0 ± 1.5	14.4 ± 2.0	11.9 ± 1.5	7.3 ± 1.1	2.43 ± 0.35	2.2 ± 0.6
Fakes	29 ± 12	29 ± 15	27 ± 12	31 ± 19	22 ± 13	3.8 ± 2.5	8 ± 5
Total Mc	1670 ± 160	1710 ± 170	2060 ± 210	1520 ± 150	710 ± 80	250 ± 40	221 ± 19
Data	1672	1657	2016	1441	676	241	216

Table 8.2: Pre-fit yields in the single-lepton CRs. The s - and t -channel single-top, tZq , and tWZ processes are combined under ‘Other’. The $t\bar{t}H$ samples are split by truth p_{T}^H with a separate forward Higgs boson category. All uncertainties except those of the $t\bar{t}$ + jets normalisation factors are included. The yield uncertainties are symmetrised and rounded.

Process	$t\bar{t}$ + light CR	$t\bar{t}$ + $\geq 1c$ CR	$t\bar{t}$ + $1b$ CR	$t\bar{t}$ + $1B$ CR	$t\bar{t}$ + $\geq 2b$ CR 1	$t\bar{t}$ + $\geq 2b$ CR 2	$t\bar{t}$ + $\geq 2b$ CR 3	Boosted CR
$t\bar{t}H$, $p_{\text{T}}^H \in [0, 60)$ GeV	56 $\pm$ 6	59 $\pm$ 7	89 $\pm$ 11	46 $\pm$ 6	80 $\pm$ 10	12.5 $\pm$ 1.7	2.09 $\pm$ 0.26	0.54 $\pm$ 0.14
$t\bar{t}H$, $p_{\text{T}}^H \in [60, 120)$ GeV	86 $\pm$ 5	98 $\pm$ 8	133 $\pm$ 12	92 $\pm$ 13	112 $\pm$ 11	33 $\pm$ 5	4.5 $\pm$ 0.9	1.15 $\pm$ 0.23
$t\bar{t}H$, $p_{\text{T}}^H \in [120, 200)$ GeV	59 $\pm$ 4	76 $\pm$ 7	76 $\pm$ 8	88 $\pm$ 11	41 $\pm$ 4	68 $\pm$ 8	5.4 $\pm$ 1.1	1.36 $\pm$ 0.28
$t\bar{t}H$, $p_{\text{T}}^H \in [200, 300)$ GeV	23.0 $\pm$ 2.5	31.9 $\pm$ 3.5	21.2 $\pm$ 3.1	41 $\pm$ 5	7.3 $\pm$ 1.1	34 $\pm$ 4	6.7 $\pm$ 1.1	1.42 $\pm$ 0.32
$t\bar{t}H$, $p_{\text{T}}^H \in [300, 450)$ GeV	7.7 $\pm$ 1.0	10.0 $\pm$ 1.5	4.7 $\pm$ 0.8	14.3 $\pm$ 2.0	1.12 $\pm$ 0.27	4.4 $\pm$ 0.8	9.8 $\pm$ 1.5	4.8 $\pm$ 0.7
$t\bar{t}H$, $p_{\text{T}}^H \geq 450$ GeV	1.67 $\pm$ 0.27	2.4 $\pm$ 0.5	0.78 $\pm$ 0.19	4.1 $\pm$ 0.7	0.18 $\pm$ 0.05	0.35 $\pm$ 0.12	3.0 $\pm$ 0.7	1.75 $\pm$ 0.32
$t\bar{t}H$, $ y_H > 2.5$	0.98 $\pm$ 0.08	0.72 $\pm$ 0.17	1.38 $\pm$ 0.22	0.57 $\pm$ 0.15	0.37 $\pm$ 0.05	0.188 $\pm$ 0.029	< 0.1	< 0.1
$t\bar{t}$ + $\geq 2b$	2320 $\pm$ 280	3800 $\pm$ 400	5400 $\pm$ 600	3500 $\pm$ 400	7200 $\pm$ 600	4900 $\pm$ 500	1140 $\pm$ 170	141 $\pm$ 10
$t\bar{t}$ + $1b$	3900 $\pm$ 400	4600 $\pm$ 700	10600 $\pm$ 1700	5100 $\pm$ 600	2100 $\pm$ 500	1120 $\pm$ 320	230 $\pm$ 100	54 $\pm$ 19
$t\bar{t}$ + $1B$	890 $\pm$ 80	1390 $\pm$ 270	1780 $\pm$ 270	5100 $\pm$ 400	530 $\pm$ 190	380 $\pm$ 140	100 $\pm$ 50	32 $\pm$ 7
$t\bar{t}$ + $\geq 1c$	17000 $\pm$ 3300	17900 $\pm$ 2600	7800 $\pm$ 1000	5400 $\pm$ 1300	2500 $\pm$ 500	1220 $\pm$ 270	240 $\pm$ 70	133 $\pm$ 35
$t\bar{t}$ + light	31800 $\pm$ 3500	5900 $\pm$ 1100	4000 $\pm$ 600	2000 $\pm$ 500	700 $\pm$ 180	320 $\pm$ 110	49 $\pm$ 26	34 $\pm$ 8
tW	1040 $\pm$ 200	750 $\pm$ 220	820 $\pm$ 240	560 $\pm$ 190	350 $\pm$ 140	360 $\pm$ 150	130 $\pm$ 90	10 $\pm$ 6
$t\bar{t}Z$	152 $\pm$ 20	171 $\pm$ 25	151 $\pm$ 21	146 $\pm$ 22	122 $\pm$ 18	117 $\pm$ 15	36 $\pm$ 5	5.7 $\pm$ 0.9
$t\bar{t}W$	100 $\pm$ 12	114 $\pm$ 17	43 $\pm$ 7	44 $\pm$ 5	25.6 $\pm$ 3.5	26 $\pm$ 5	10.5 $\pm$ 1.6	1.97 $\pm$ 0.33
$t\bar{t}t\bar{t}$	2.9 $\pm$ 1.2	24 $\pm$ 10	7.1 $\pm$ 3.0	19 $\pm$ 8	21 $\pm$ 9	33 $\pm$ 14	17 $\pm$ 7	1.4 $\pm$ 0.6
$tHjb$	8.2 $\pm$ 1.4	5.4 $\pm$ 0.9	8.1 $\pm$ 1.6	3.5 $\pm$ 0.6	5.4 $\pm$ 1.1	3.6 $\pm$ 0.7	0.53 $\pm$ 0.15	0.65 $\pm$ 0.20
tWH	7.1 $\pm$ 0.7	6.7 $\pm$ 0.7	9.6 $\pm$ 1.0	5.7 $\pm$ 0.6	5.4 $\pm$ 0.6	5.8 $\pm$ 0.7	1.78 $\pm$ 0.20	0.44 $\pm$ 0.07
W + jets	370 $\pm$ 180	370 $\pm$ 180	230 $\pm$ 120	170 $\pm$ 80	370 $\pm$ 190	250 $\pm$ 130	90 $\pm$ 40	5.6 $\pm$ 2.9
Z + jets	110 $\pm$ 40	110 $\pm$ 40	110 $\pm$ 40	64 $\pm$ 23	100 $\pm$ 40	55 $\pm$ 20	16 $\pm$ 6	0.9 $\pm$ 0.4
Diboson	37 $\pm$ 19	39 $\pm$ 20	23 $\pm$ 12	21 $\pm$ 11	24 $\pm$ 13	23 $\pm$ 12	10 $\pm$ 5	1.1 $\pm$ 0.7
Other	370 $\pm$ 40	370 $\pm$ 40	303 $\pm$ 35	228 $\pm$ 22	179 $\pm$ 22	111 $\pm$ 12	27.4 $\pm$ 2.8	3.9 $\pm$ 0.9
Fakes	810 $\pm$ 300	800 $\pm$ 400	880 $\pm$ 350	340 $\pm$ 140	660 $\pm$ 300	280 $\pm$ 130	57 $\pm$ 32	6.3 $\pm$ 3.3
Total Mc	59000 $\pm$ 6000	37000 $\pm$ 4000	32500 $\pm$ 2900	22900 $\pm$ 2200	15200 $\pm$ 1700	9400 $\pm$ 1000	2190 $\pm$ 270	440 $\pm$ 50
Data	61954	36528	32887	23245	15595	9397	2097	426

Table 8.3: Pre-fit yields in the dilepton SRs. The tZq , tWZ , and diboson processes are combined under ‘Other’. The $t\bar{t}H$ samples are split by truth p_{T}^H with a separate forward Higgs boson category. All uncertainties except those of the $t\bar{t}$ + jets normalisation factors are included. The yield uncertainties are symmetrised and rounded.

Process	STXS 1 SR	STXS 2 SR	STXS 3 SR	STXS 4 SR	STXS 5 SR	STXS 6 SR
$t\bar{t}H$, $p_{\text{T}}^H \in [0, 60)$ GeV	10.6 $\pm$ 1.5	2.03 $\pm$ 0.31	1.14 $\pm$ 0.17	0.63 $\pm$ 0.11	0.19 $\pm$ 0.07	0.046 $\pm$ 0.023
$t\bar{t}H$, $p_{\text{T}}^H \in [60, 120)$ GeV	6.5 $\pm$ 0.7	13.0 $\pm$ 1.8	4.1 $\pm$ 0.7	1.9 $\pm$ 0.4	0.70 $\pm$ 0.11	0.13 $\pm$ 0.07
$t\bar{t}H$, $p_{\text{T}}^H \in [120, 200)$ GeV	3.06 $\pm$ 0.34	4.7 $\pm$ 0.5	16.3 $\pm$ 2.0	5.0 $\pm$ 0.7	1.36 $\pm$ 0.19	0.24 $\pm$ 0.07
$t\bar{t}H$, $p_{\text{T}}^H \in [200, 300)$ GeV	0.50 $\pm$ 0.07	0.72 $\pm$ 0.08	2.50 $\pm$ 0.29	13.0 $\pm$ 1.7	4.4 $\pm$ 0.6	0.41 $\pm$ 0.06
$t\bar{t}H$, $p_{\text{T}}^H \in [300, 450)$ GeV	< 0.1	< 0.1	0.197 $\pm$ 0.033	0.82 $\pm$ 0.18	8.0 $\pm$ 1.1	1.98 $\pm$ 0.32
$t\bar{t}H$, $p_{\text{T}}^H \geq 450$ GeV	< 0.1	< 0.1	< 0.1	< 0.1	0.26 $\pm$ 0.08	2.8 $\pm$ 0.5
$t\bar{t}H$, $ y_H > 2.5$	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1	< 0.1
$t\bar{t}$ + $\geq 2b$	81 $\pm$ 8	76 $\pm$ 7	81 $\pm$ 13	59 $\pm$ 4	36.2 $\pm$ 3.0	11.7 $\pm$ 2.1
$t\bar{t}$ + $1b$	14 $\pm$ 5	15 $\pm$ 6	20 $\pm$ 9	17 $\pm$ 5	11 $\pm$ 6	2.9 $\pm$ 1.9
$t\bar{t}$ + $1B$	4.4 $\pm$ 2.6	4.8 $\pm$ 3.2	5.7 $\pm$ 2.3	6.2 $\pm$ 1.9	6.6 $\pm$ 2.9	3.5 $\pm$ 2.3
$t\bar{t}$ + $\geq 1c$	11.3 $\pm$ 2.4	11.7 $\pm$ 2.6	15.4 $\pm$ 3.4	12.5 $\pm$ 2.6	7.1 $\pm$ 2.6	3.0 $\pm$ 1.7
$t\bar{t}$ + light	1.1 $\pm$ 0.7	0.9 $\pm$ 0.5	1.4 $\pm$ 0.8	1.0 $\pm$ 0.5	1.0 $\pm$ 0.4	0.37 $\pm$ 0.30
tW	1.7 $\pm$ 0.8	3.0 $\pm$ 1.8	6 $\pm$ 4	8 $\pm$ 5	7 $\pm$ 5	3.1 $\pm$ 2.9
$t\bar{t}Z$	4.6 $\pm$ 0.7	4.6 $\pm$ 1.1	7.2 $\pm$ 1.1	6.2 $\pm$ 1.3	5.3 $\pm$ 1.1	2.0 $\pm$ 0.6
$t\bar{t}W$	0.65 $\pm$ 0.11	1.18 $\pm$ 0.15	2.10 $\pm$ 0.23	2.64 $\pm$ 0.30	2.7 $\pm$ 0.4	1.2 $\pm$ 0.5
$t\bar{t}t\bar{t}$	0.9 $\pm$ 0.4	1.1 $\pm$ 0.5	1.8 $\pm$ 0.7	1.7 $\pm$ 0.7	1.3 $\pm$ 0.6	0.65 $\pm$ 0.28
tWH	0.23 $\pm$ 0.04	0.31 $\pm$ 0.04	0.57 $\pm$ 0.07	0.70 $\pm$ 0.08	0.66 $\pm$ 0.07	0.243 $\pm$ 0.030
Z + jets	1.8 $\pm$ 0.9	1.9 $\pm$ 0.9	3.4 $\pm$ 1.3	3.8 $\pm$ 1.4	3.3 $\pm$ 1.2	1.8 $\pm$ 0.7
Other	< 0.1	< 0.1	0.28 $\pm$ 0.14	0.42 $\pm$ 0.19	0.23 $\pm$ 0.08	0.29 $\pm$ 0.13
Fakes	1.7 $\pm$ 0.9	2.2 $\pm$ 1.1	3.2 $\pm$ 1.6	2.9 $\pm$ 1.5	2.6 $\pm$ 1.3	1.1 $\pm$ 0.6
Total Mc	144 $\pm$ 13	144 $\pm$ 13	173 $\pm$ 20	143 $\pm$ 12	100 $\pm$ 11	38 $\pm$ 6
Data	150	149	161	149	76	35

Table 8.4: Pre-fit yields in the dilepton CRs. The tZq , tWZ , and diboson processes are combined under ‘Other’. The $t\bar{t}H$ samples are split by truth p_T^H with a separate forward Higgs boson category. All uncertainties except those of the $t\bar{t}$ + jets normalisation factors are included. The yield uncertainties are symmetrised and rounded.

Process	$t\bar{t}$ + light CR	$t\bar{t}$ + $\geq 1c$ CR	$t\bar{t}$ + $1b$ CR	$t\bar{t}$ + $1B$ CR	$t\bar{t}$ + $\geq 2b$ CR
$t\bar{t}H$, $p_T^H \in [0, 60)$ GeV	11.9 ± 1.4	19.7 ± 2.3	19.8 ± 2.3	9.9 ± 1.4	26.0 ± 3.0
$t\bar{t}H$, $p_T^H \in [60, 120)$ GeV	17.0 ± 1.3	34.5 ± 2.9	29.5 ± 3.0	20.4 ± 2.5	42 ± 4
$t\bar{t}H$, $p_T^H \in [120, 200)$ GeV	10.0 ± 0.9	27.1 ± 2.7	17.0 ± 2.1	19.2 ± 2.4	34.0 ± 3.5
$t\bar{t}H$, $p_T^H \in [200, 300)$ GeV	3.10 ± 0.35	11.1 ± 1.3	4.6 ± 0.7	9.2 ± 1.2	14.9 ± 1.9
$t\bar{t}H$, $p_T^H \in [300, 450)$ GeV	0.94 ± 0.15	3.5 ± 0.5	0.94 ± 0.19	3.3 ± 0.5	5.4 ± 0.8
$t\bar{t}H$, $p_T^H \geq 450$ GeV	0.33 ± 0.07	0.65 ± 0.12	0.116 ± 0.033	0.97 ± 0.15	1.33 ± 0.26
$t\bar{t}H$, $ y_H > 2.5$	0.31 ± 0.05	0.25 ± 0.04	0.40 ± 0.05	0.127 ± 0.028	0.166 ± 0.018
$t\bar{t}$ + $\geq 2b$	495 ± 35	860 ± 60	1080 ± 200	650 ± 100	2700 ± 200
$t\bar{t}$ + $1b$	1410 ± 130	1790 ± 230	3600 ± 700	1540 ± 180	1200 ± 400
$t\bar{t}$ + $1B$	310 ± 40	460 ± 40	540 ± 210	1560 ± 230	310 ± 110
$t\bar{t}$ + $\geq 1c$	5600 ± 500	7700 ± 600	2210 ± 280	1590 ± 330	1070 ± 170
$t\bar{t}$ + light	9200 ± 1500	2800 ± 400	590 ± 130	410 ± 90	190 ± 70
tW	300 ± 60	210 ± 50	220 ± 70	150 ± 50	300 ± 140
$t\bar{t}Z$	43 ± 6	81 ± 12	35 ± 6	36 ± 5	92 ± 12
$t\bar{t}W$	37 ± 5	72 ± 7	12.1 ± 1.3	14.3 ± 3.3	29.8 ± 3.5
$t\bar{t}t\bar{t}$	1.6 ± 0.7	10 ± 4	1.0 ± 0.4	4.8 ± 2.0	28 ± 12
tWH	0.98 ± 0.11	2.24 ± 0.22	2.69 ± 0.29	1.42 ± 0.14	4.0 ± 0.4
Z + jets	340 ± 130	260 ± 100	250 ± 90	93 ± 34	320 ± 120
Other	5.6 ± 2.4	8.1 ± 3.1	4.4 ± 2.0	2.1 ± 0.7	12 ± 5
Fakes	110 ± 50	120 ± 60	42 ± 21	41 ± 21	90 ± 40
Total Mc	18000 ± 2000	14400 ± 1100	8700 ± 900	6200 ± 500	6400 ± 700
Data	18557	14361	8624	5830	6448

8.3. $t\bar{t}H$ Measurement Results

Three different measurements of the $t\bar{t}H$ signal process were conducted with the regions outlined in Section 6.5: a measurement of the inclusive $t\bar{t}H$ signal strength $\mu_{t\bar{t}H}$, a measurement of the inclusive $t\bar{t}H$ cross-section $\sigma_{t\bar{t}H}$, and a differential measurement of the $t\bar{t}H$ cross-section in the STXS stage 1.2 formalism. Simultaneous profile likelihood fits to all SRs and CRs of the single-lepton and the dilepton channel are used for all measurements. Since the Transformer discriminants retain a significant separation after events are assigned to the analysis regions, the discriminant corresponding to the target process of each region is used as the fit variable. Exceptions are the boosted SR and CR, where the reconstructed Higgs boson p_T in bins corresponding to the STXS ranges is used as a fit variable.

The signal strength measurement uses a single POI for the normalisation of the $t\bar{t}H$ signal process in the full analysis phase space. As mentioned in Section 6.4, free-floating normalisation factors for each of the five $t\bar{t}$ + jets components are determined simultaneously with the signal strength. A best-fit signal strength of

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

was measured in this setup. Herein, the total statistical uncertainty is defined as

the uncertainty arising in a constrained fit, in which all systematic NPs are fixed to their best-fit values. Meanwhile, the total systematic uncertainty is defined via the difference in quadrature between the total and the statistical uncertainty. The systematic uncertainty on $\mu_{t\bar{t}H}$ includes the $t\bar{t}H$ cross-section uncertainties defined in the previous section. These are dropped for the two cross-section measurements. By using a profile likelihood ratio between a background-only hypothesis of $\mu_{t\bar{t}H} = 0$ and the best-fit results, the observed excess of events has a significance of 4.6σ compared to an expected significance of 5.4σ determined using an Asimov dataset. Post-fit normalisation factors of the $t\bar{t} + \text{jets}$ components are shown with their uncertainties in Table 8.5. Separate signal strengths of the individual channels were determined in single-channel fits as

$$\mu_{t\bar{t}H}^{\text{single-lep}} = 0.72 \pm 0.12 \text{ (stat.) } {}^{+0.21}_{-0.17} \text{ (syst.)}$$

in the single-lepton and

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

in the dilepton channel, showing good compatibility between the single-channel and the combined measurements. The overall sensitivity is driven by the single-lepton channel, where higher event statistics lead to a systematically dominated measurement.

Table 8.5: Best-fit values of the $t\bar{t} + \text{jets}$ normalisation factors as determined in the inclusive $t\bar{t}H$ signal strength measurement. The values are consistent with those from the inclusive $t\bar{t}H$ cross-section measurement. Independent single-channel normalisation factors are used for $t\bar{t} + \text{light}$, $t\bar{t} + \geq 1c$, and $t\bar{t} + 1b$.

Channel	$t\bar{t} + \text{light}$	$t\bar{t} + \geq 1c$	$t\bar{t} + 1b$	$t\bar{t} + 1B$	$t\bar{t} + \geq 2b$
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	$0.85 {}^{+0.11}_{-0.10}$	$1.36 {}^{+0.10}_{-0.10}$	$1.24 {}^{+0.09}_{-0.09}$		

As opposed to the $t\bar{t}H$ signal strength measurement, no theory uncertainties for the $t\bar{t}H$ cross-section enter the result in the inclusive cross-section measurement. The inclusive $t\bar{t}H$ cross-section is measured to be

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

using a profile likelihood fit. The statistical and systematic uncertainties are determined analogously to those of the signal strength measurement. Overall, the measured $t\bar{t}H$ cross-section is compatible with the SM prediction of $\sigma^{\text{SM}} = 506 {}^{+35}_{-50} \text{ fb}$. A ranking of the dominant systematic uncertainties impacting the inclusive signal strength measurement is shown in Figure 8.3. Additionally, the impact of different groups of systematics was determined by sequentially repeating the fit with each

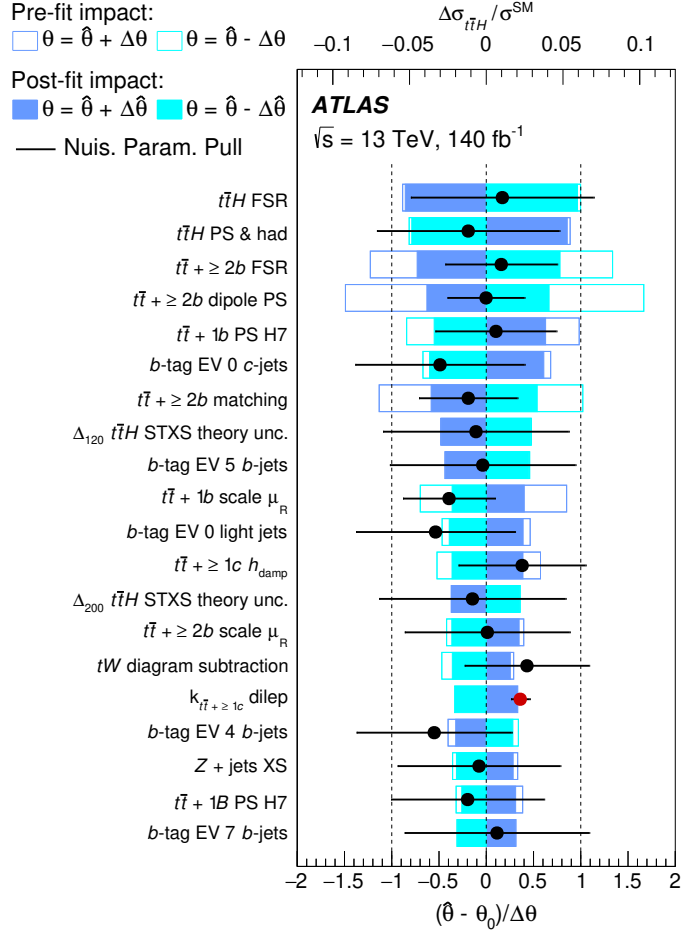


Figure 8.3: Ranking of the 20 dominant systematic uncertainties of the inclusive $t\bar{t}H$ cross-section measurement. The empty and filled rectangles represent the pre- and post-fit impact of the systematic on the $t\bar{t}H$ cross-section, respectively, as determined by profiling the respective NP of the uncertainty. Meanwhile, the markers show the NP pulls and constraints, which reflect the shift of the best-fit value of each NP with respect to its pre-fit value.

systematics category fixed to the best-fit values of the corresponding NPs. The group uncertainty was then determined as the difference in quadrature between the total uncertainty and the uncertainty of the constrained fit. The resulting grouped impacts are presented in Table 8.6. An overview of the full analysis phase space, where the individual bins entering the fit are grouped by their respective signal-over-background ratios S/B , is plotted in Figure 8.4.

The fit setup for the differential $t\bar{t}H$ cross-section measurement is identical to that of the inclusive cross-section measurement, but uses a total of six POIs for the $t\bar{t}H$ process yields in the different Higgs boson p_T ranges. For this purpose, the $t\bar{t}H$ signal sample is split into six parts based on the truth Higgs boson p_T in the Mc events. Simultaneously, no $t\bar{t}H$ migration uncertainties are considered for the differential cross-section measurement. Additionally, Mc events with an absolute Higgs boson

Table 8.6: List of the absolute uncertainties of the inclusive $t\bar{t}H$ cross-section grouped into categories of systematic uncertainties. The procedure used to arrive at the individual uncertainties is outlined in the text. Correlations between the groups of uncertainties are not taken into account. ISR, FSR, and QCD scales are included in the $t\bar{t}H$ and $t\bar{t} + \geq 1b$ radiation categories, while the $t\bar{t}H$ theory category includes Higgs boson BR uncertainties and the STXS migration uncertainties. Uncertainties in the background normalisation factors are included with the statistical uncertainties.

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

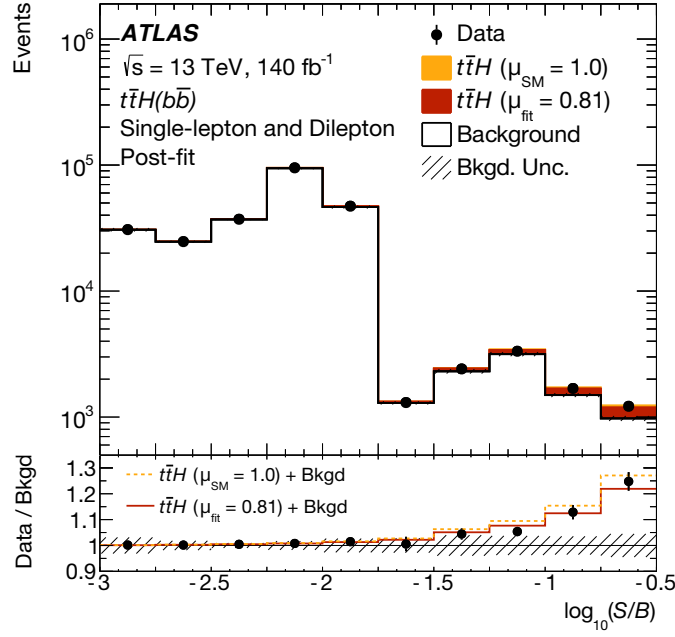


Figure 8.4: Logarithmic overview plot over all bins in the analysis phase space grouped by their respective signal-over-background ratios S/B . Both the observed and the expected event yields are shown on top of the background yields and compared to the measured yields in data. The bottom plot shows the ratio of data, observed, and expected yields with respect to the background yields. The shaded band represents the uncertainty in the background prediction.

rapidity of $|y_H| > 2.5$ are not considered for the PoIs but instead held constant at the SM prediction following the STXS recommendations. The cross-section result of this differential cross-section measurement is shown in Figure 8.5, while the post-fit correlations between the measured per-bin $t\bar{t}H$ cross-section PoIs are shown in Figure 8.6.

Post-fit plots of the inclusive cross-section fit are shown in Figures 8.7 to 8.12. For readability, their captions are kept short and all common information is provided in the following: all post-fit comparisons of data and MC predictions stem from a simultaneous profile likelihood fit of the inclusive $t\bar{t}H$ cross-section to all regions in the single-lepton and dilepton channels. The uncertainty band includes all uncertainties from the corresponding fit setup with their correlations. The single-lepton resolved and dilepton SRs are split by p_T^H as reconstructed by the Transformer networks as introduced in Section 6.5.4. Similarly, the single-lepton $t\bar{t} + \geq 2b$ CR as determined by the classification Transformer is further split into three regions corresponding to the STXS bins 1+2, 3+4, and 5+6 by the reconstruction Transformer to improve the constraints of the associated background after mis-modelling was observed during the analysis. The discriminant classifiers lie in the range between zero and one following the rescaling introduced by Equation (7.3).

8.3. $t\bar{t}H$ Measurement Results

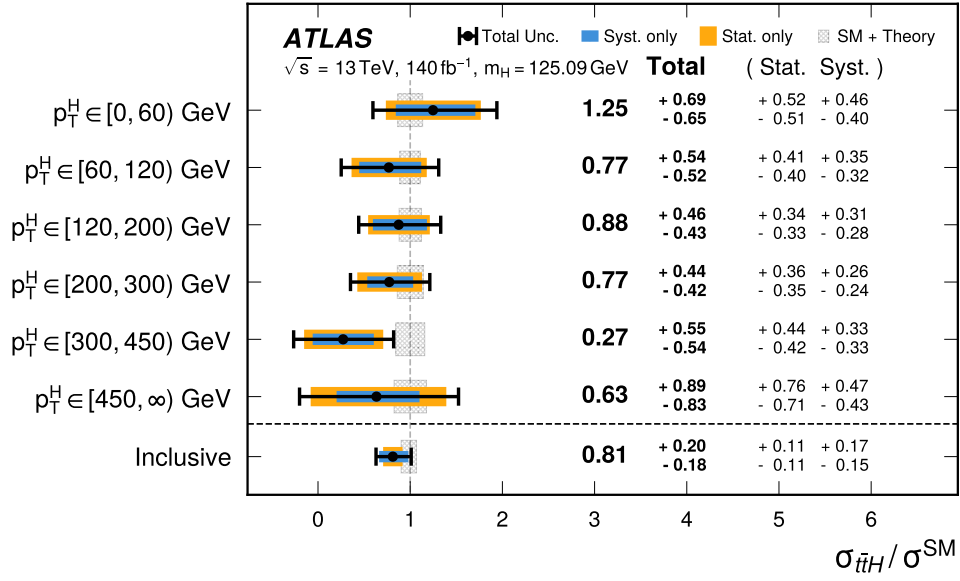


Figure 8.5: Results of the differential $t\bar{t}H$ cross-section measurement in bins of truth p_T^H for $|y_H| \leq 2.5$ as obtained from a combined profile likelihood fit with six POIs. The p_T^H ranges correspond to those of the STXS stage 1.2 formalism.

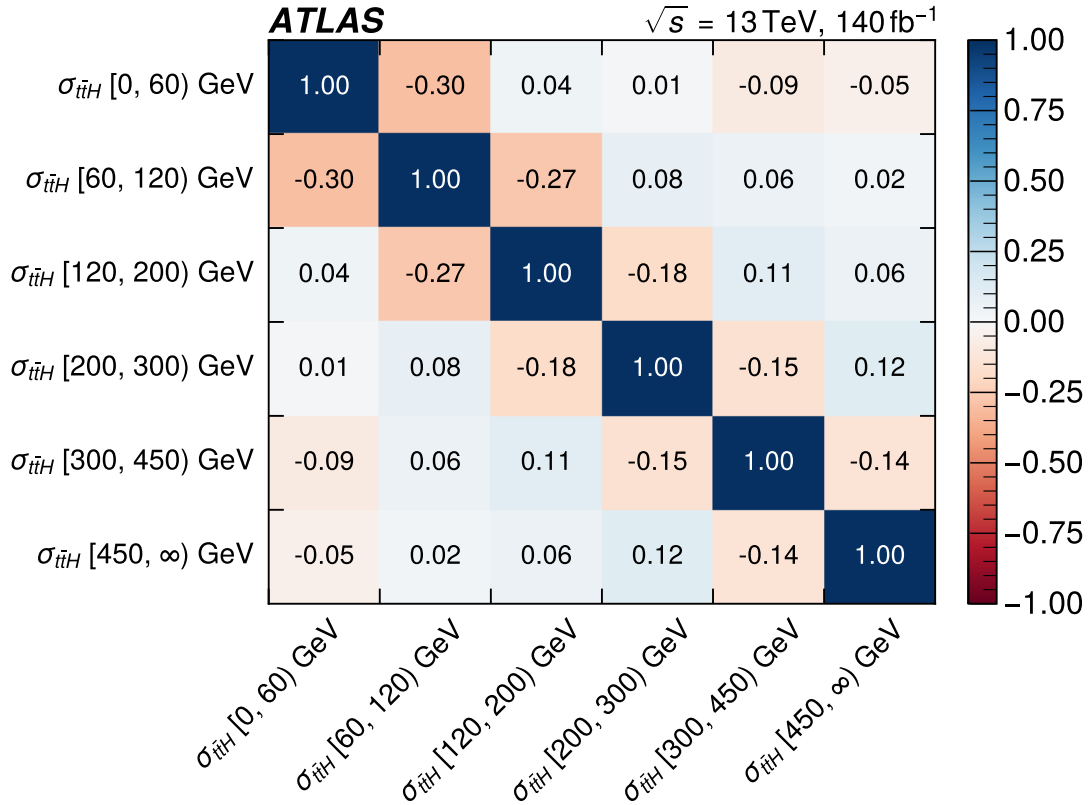


Figure 8.6: Post-fit correlations between the six $t\bar{t}H$ cross-section POIs measured differentially in bins of truth p_T^H for $|y_H| \leq 2.5$ as obtained from a combined profile likelihood fit.

Chapter 8. Results

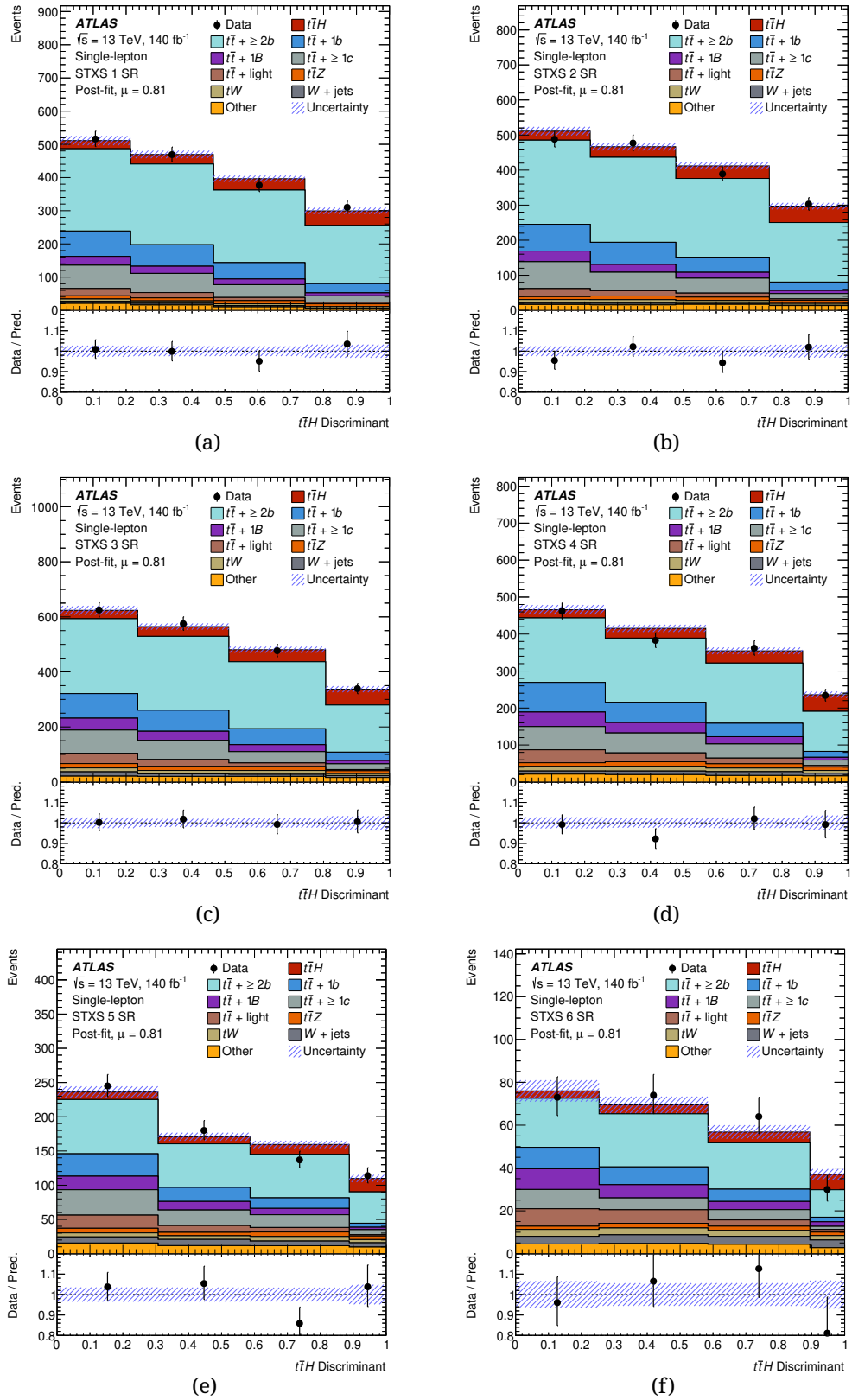


Figure 8.7: Post-fit comparison of data and Mc predictions in the single-lepton resolved SRS from a combined profile likelihood fit. Additional details are contained in the text.

8.3. $t\bar{t}H$ Measurement Results

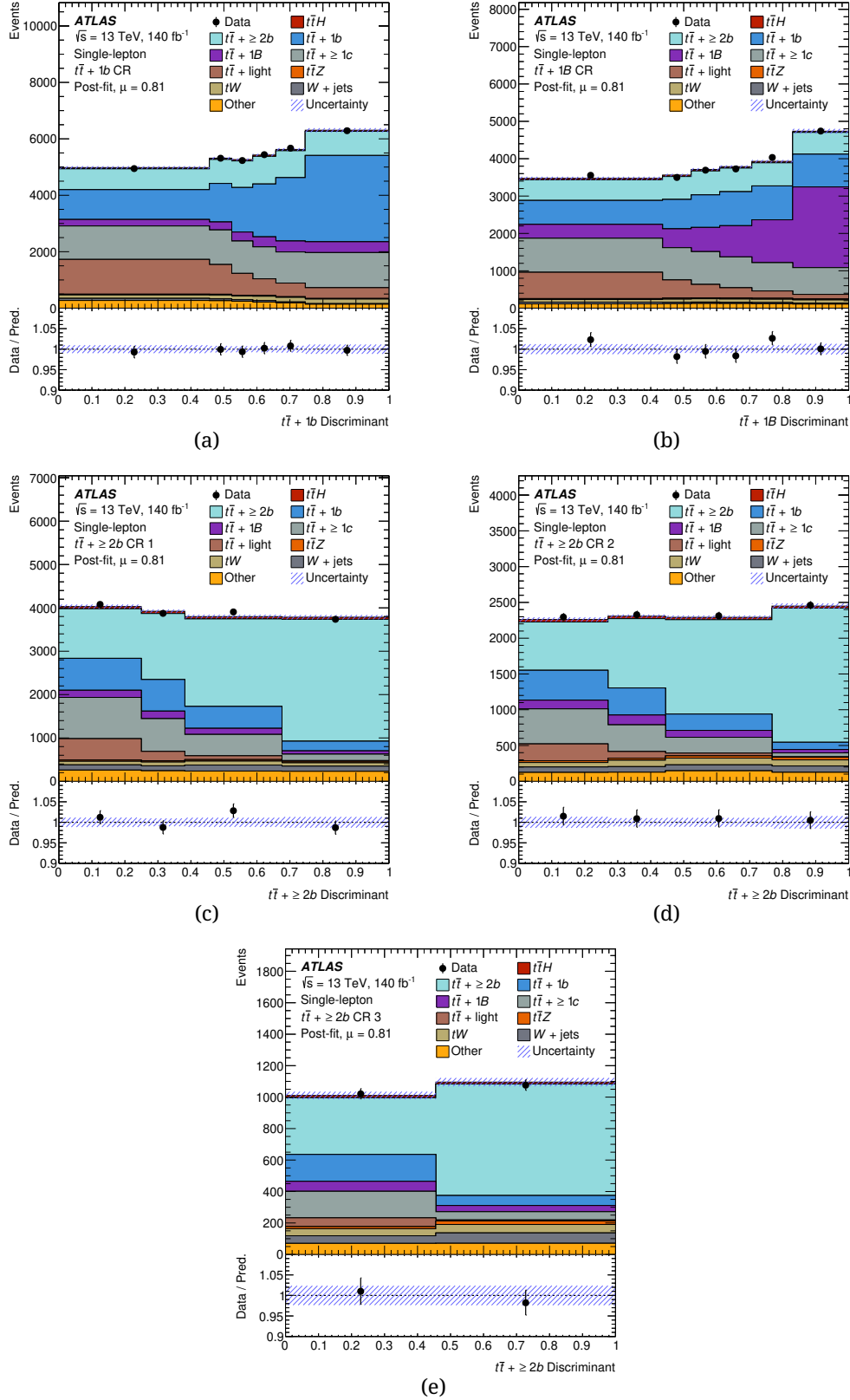


Figure 8.8: Post-fit comparison of data and MC predictions in the single-lepton resolved $t\bar{t} + \geq 1b$ CRs from a combined profile likelihood fit. Additional details are contained in the text.

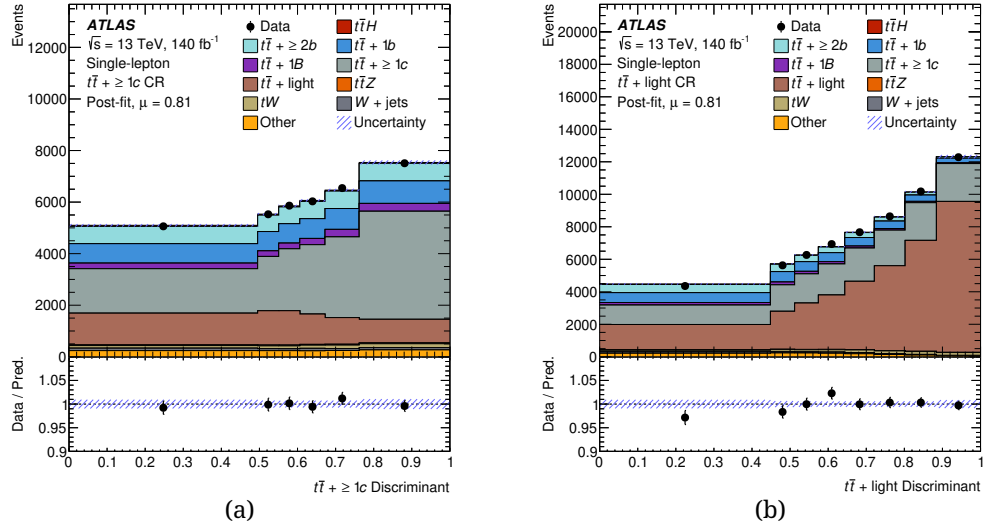


Figure 8.9: Post-fit comparison of data and Mc predictions in the single-lepton resolved $t\bar{t} + \geq 1c$ and $t\bar{t} + \text{light}$ CRs from a combined profile likelihood fit. Additional details are contained in the text.

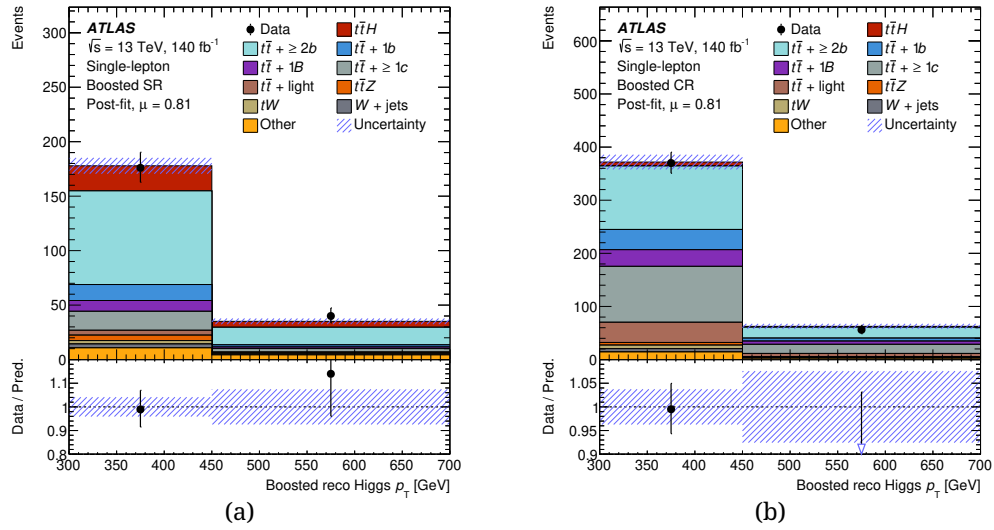


Figure 8.10: Post-fit comparison of data and Mc predictions in the single-lepton boosted SR and CR from a combined profile likelihood fit. Additional details are contained in the text. Underflow and overflow bins of the p_T^H distributions are included in the fit in the first and last bin, respectively.

8.3. $t\bar{t}H$ Measurement Results

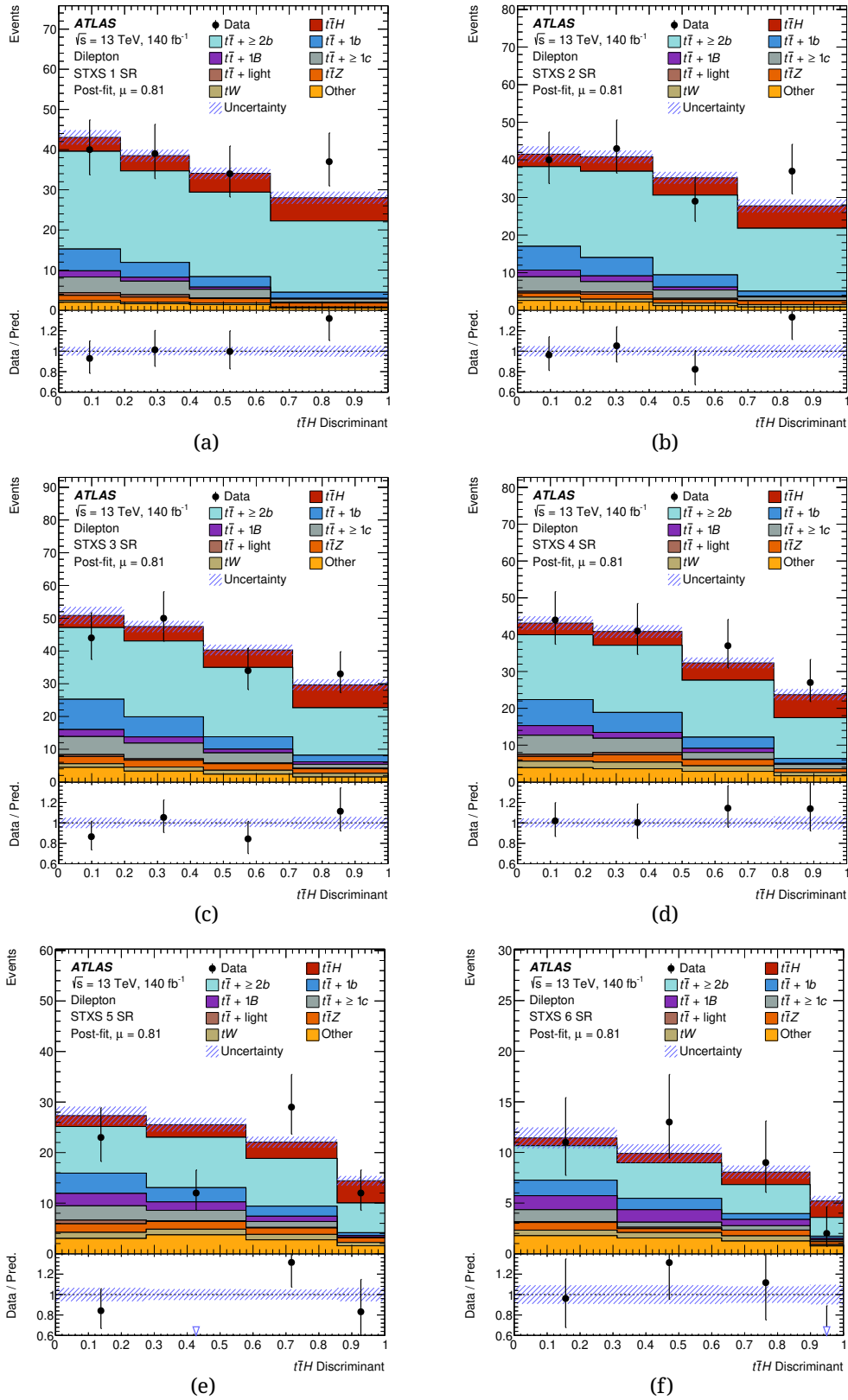


Figure 8.11: Post-fit comparison of data and Mc predictions in the dilepton SRS from a combined profile likelihood fit. Additional details are contained in the text.

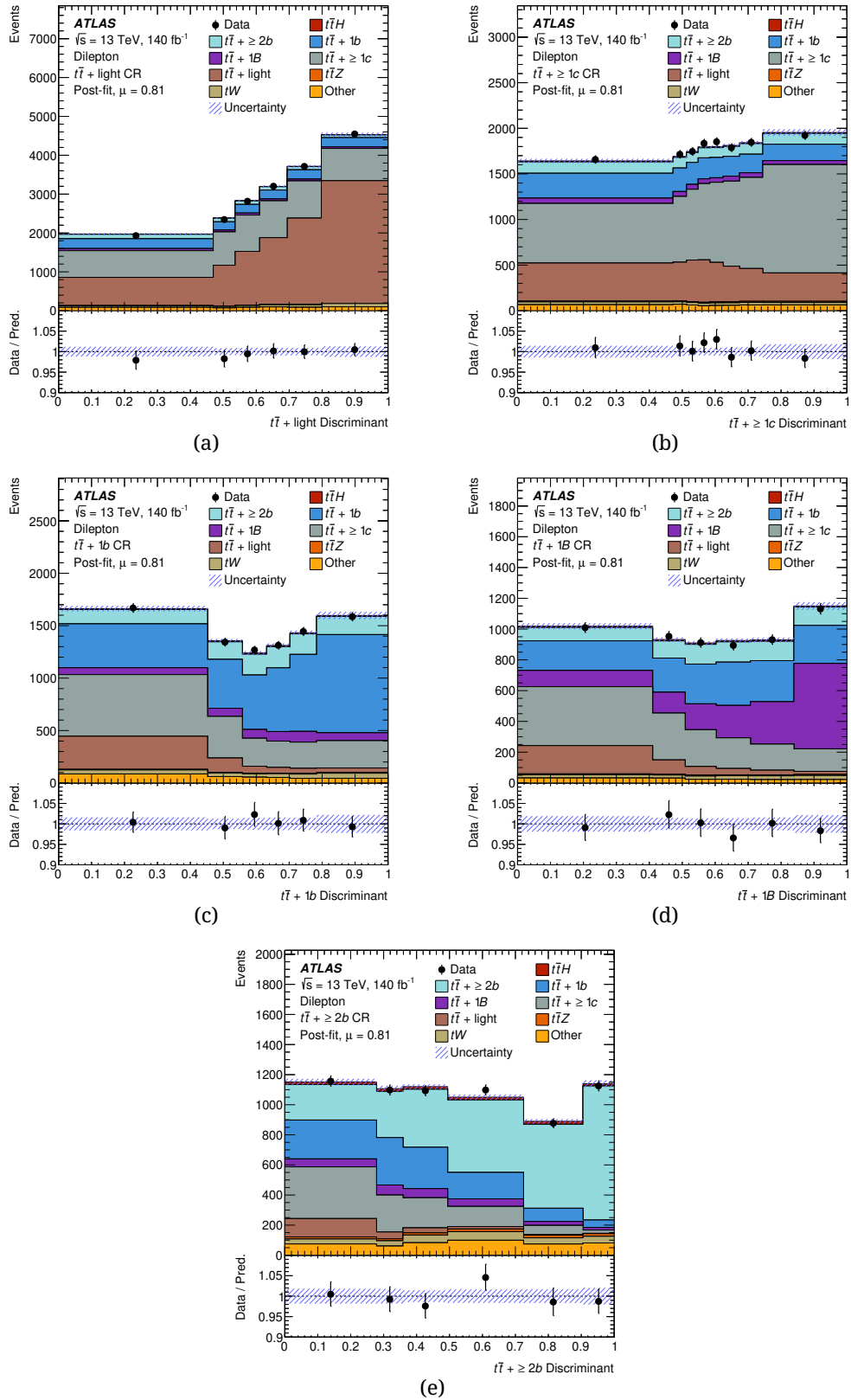


Figure 8.12: Post-fit comparison of data and Mc predictions in the dilepton CRs from a combined profile likelihood fit. Additional details are contained in the text.

8.3.1. Comparison to Previous Results

As mentioned in Section 1.2.3, both the previous full Run 2 ATLAS analysis published as Reference [211] and the full Run 2 CMS analysis contained in Reference [208] measured low $t\bar{t}H$ signal strengths of $\mu_{t\bar{t}H} = 0.35^{+0.36}_{-0.34}$ and $\mu_{t\bar{t}H} = 0.33 \pm 0.26$, respectively. Consequently, their compatibility with the SM expectation is 8.5 % and 2 %, respectively. Meanwhile, the measurements of the ATLAS $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis do not appear to be in tension with the SM.

Since the first full Run 2 ATLAS analysis uses the same dataset as the legacy analysis presented in this thesis – albeit with a different event selection and obsoleted object recommendations – a compatibility study of the two analyses was performed. Overall, 64 % (29 %) of events in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy single-lepton (dilepton) channels were not selected by the previous analysis. In a bootstrapping method assigning each data event a Poisson distribution with a mean value of unity, the statistical correlation between the analyses was estimated to be 19 % when assuming entirely uncorrelated systematics between both analyses. The overhauled $t\bar{t} + \geq 1b$ systematics model and the updated experimental uncertainties may justify this approximation. Based on these correlations, the probability of compatible results between the two analyses is estimated at 21 %.

Comparing dominant systematic uncertainties of the previous analysis, shown in Figure 8.13, with those of the legacy analysis in Figure 8.3 reveals that $t\bar{t} + \geq 1b$ systematic uncertainties completely dominated the previous analysis, while the legacy analysis is dominated by systematics related to the $t\bar{t}H$ signal modelling, while $t\bar{t} + \geq 1b$ systematics follow at lower ranks. Therefore, it seems as though the improved background model developed in Reference [219] together with the higher constraining power of the auxiliary $t\bar{t} + \text{jets}$ measurements afforded by the state-of-the-art Transformer models were able to reduce the impact of the uncertainties in the dominant $t\bar{t} + \geq 1b$ background. This fact can also be seen when comparing the expected significance of 2.7σ in the first full Run 2 ATLAS analysis with that of 5.4σ expected by the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis. Additionally, the improvements in the analysis strategy have afforded the legacy analysis a measurement in all six bins of the STXS formalism, while the previous ATLAS analysis had to combine the two lowest bins. The CMS analysis, meanwhile, combined the two highest STXS bins. Overall, the inclusive cross-section measurement of the ATLAS $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis constitutes the most precise cross-section measurement of the $t\bar{t}H$ process in a single Higgs boson decay channel to date.

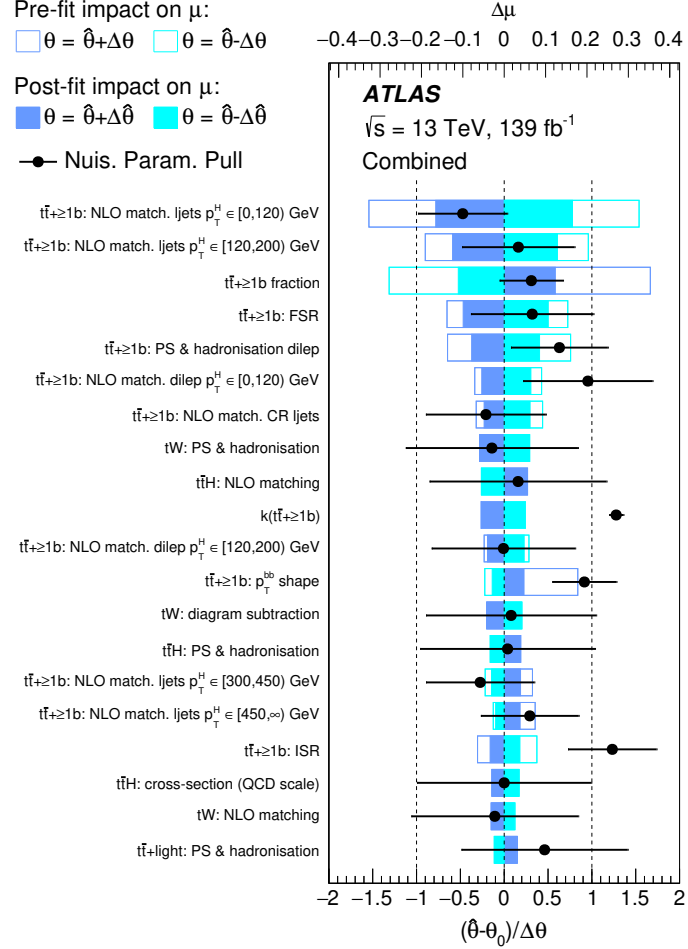


Figure 8.13: Ranking of the 20 dominant systematic uncertainties of the inclusive $t\bar{t}H$ cross-section measurement in the first full Run 2 ATLAS analysis. The plot has the same format as that shown for the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis in Figure 8.3. However, the ranking of the first full Run 2 analysis is entirely dominated by $t\bar{t} + \geq 1b$ systematics, while signal modelling uncertainties dominate the ranking of the legacy analysis. The plot is taken from Reference [211].

Summary and Conclusion

Since the discovery of the Higgs boson as the last missing piece of the Standard Model, its properties have been extensively probed in high-energy collisions by the ATLAS and CMS collaborations. The Yukawa coupling of the heavy top quark to the Higgs boson has a special place in these investigations. As the largest Higgs coupling with a value close to unity, it is a crucial ingredient to the shape of the Higgs potential due to higher-order corrections. In 2018, the observation of the $t\bar{t}H$ production process opened the door for direct measurements of the top Yukawa coupling.

In this thesis, a measurement of $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel and targeting the leptonic top quark decays was presented. The analysis uses the full ATLAS Run 2 dataset collected in pp collisions at 13 TeV. It represents a re-analysis incorporating various improvements from the modelling of the dominant $t\bar{t} + \geq 1b$ background process, new object definitions, and a loosening of event selection requirements. Additionally, the analysis utilises state-of-the-art Transformer neural networks, leading to significant improvements in event classification and Higgs boson p_T reconstruction performance. The development and validation of these machine learning techniques was one of the main topics discussed in this thesis.

As a result of these improvements, the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis can claim strong evidence of the production process with an observed event excess of 4.6σ , while an excess of 5.4σ was expected. The inclusive cross-section measurement of

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

represents the most precise $t\bar{t}H$ cross-section measurement in a single Higgs boson decay channel to date. Additionally, a measurement of the differential cross-section in the full STXS stage 1.2 formalism could be realised by the analysis. The improved $t\bar{t} + \text{jets}$ process modelling and usage of control regions for auxiliary measurements of each $t\bar{t} + \text{jets}$ component led to a shift of the dominant systematic uncertainty from the modelling of the $t\bar{t} + \geq 1b$ background to the modelling of the $t\bar{t}H$ signal process.

The performance gains achieved by the usage of the Transformer networks foreshadow the shape of things to come for MVA methods in future particle physics analyses: machine learning techniques are developing at lightning speed, exemplified by the fact that many of the methods used in the Transformers trained as a part of this thesis have arrived on the scene in the past five to ten years. Nevertheless, this still represents an eternity in machine learning research. Given the length of an average particle physics analysis in ATLAS, many measurements utilising modern machine learning techniques are likely to be published in the upcoming years. At the same time, there are still many improvements to be made to the MVA techniques used in the $t\bar{t}(H \rightarrow b\bar{b})$ legacy analysis: a combined model encompassing both channels, as well as the event classification and Higgs boson p_T reconstruction tasks, might benefit from larger training statistics and mutual data patterns used in both classification and reconstruction. More broadly, the pooling of computing resources for the development of truly general models in event reconstruction and data analysis might yield significant improvements for future analyses. Larger and more diverse training sets have the potential to build generally applicable models. Simultaneously, the usage of the same model by multiple analyses might simplify combination efforts and streamline systematic uncertainties arising from the usage of machine learning techniques.

In addition to the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy measurement, this thesis also presented the development of automated system tests for the ATLAS Pixel detector. The Pixel detector is crucial for track reconstruction and transverse momentum measurements. By identifying secondary displaced decay vertices and tracks with large impact parameters, the Pixel detector is also a critical component in the identification of b -jets, which represent an important event signature for the $t\bar{t}H$ process, but also many physics processes in general. Since the high level of pile-up and the continued high radiation environment during Run 3 present unique challenges to the Pixel detector, several mitigation strategies have been developed and deployed since Long Shutdown 2. Changes to the firmware and software of the off-detector electronics are the only viable option to retain detector operability in this context, as the on-detector components are not programmable and cannot be accessed physically during the multi-year runs of the LHC. By reducing the time required for testing and enhancing the validation of the off-detector read-out system following any changes, the automated tests to validate detector firmware and software play an important role in ensuring the continued peak performance of the Pixel detector until the end of Run 3.

In the future, the increased event statistics of Run 2 and Run 3 in combination with further developments in process modelling, object definitions, and analysis techniques should enable even more precise measurements to probe the top Yukawa coupling. These analysis efforts will also provide a testbed to develop suitable strategies for other challenging processes suffering from similar difficulties as $t\bar{t}(H \rightarrow b\bar{b})$ production, such as tH or the elusive di-Higgs boson production. Modern machine learning techniques will play a crucial part in these efforts going forward.

Bibliography

- [1] ATLAS Collaboration, *Measurement of the associated production of a top-anti-top-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*, *Eur. Phys. J. C* **85** (2025) 210
- [2] C. Scheulen, *QT Final Report: Automated Testing of Pixel Calibration, Final Qualification Task Presentation at ATLAS Pixel General Meeting* (Oct. 18, 2021), URL: <https://indico.cern.ch/event/1087955/#171-qt-final-report-automated>
- [3] A. Ghosal, *Final QT Report: Automated DAQ Testing Framework, Final Qualification Task Presentation at ATLAS Pixel General Meeting* (Oct. 25, 2021), URL: <https://indico.cern.ch/event/1090185/#178-final-qt-report-automated>
- [4] S. L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) 579
- [5] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264
- [6] A. Salam, *Weak and Electromagnetic Interactions, Proceedings of the 8th Nobel Symposium* (Lerum, Sweden, May 19–25, 1968), *Conf. Proc.* **C680519** (1968) 367
- [7] G. 't Hooft, *Renormalization of Massless Yang-Mills Fields*, *Nucl. Phys. B* **33** (1971) 173
- [8] G. 't Hooft, *Renormalizable Lagrangians for Massive Yang-Mills Fields*, *Nucl. Phys. B* **35** (1971) 167
- [9] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, *Nucl. Phys. B* **44** (1972) 189
- [10] G. 't Hooft and M. Veltman, *Combinatorics of gauge fields*, *Nucl. Phys. B* **50** (1972) 318
- [11] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, *Phys. Rev. D* **2** (1970) 1285

Bibliography

- [12] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343
- [13] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346
- [14] S. Tomonaga, *On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields*, Riken Iho (Bull. I. P. C. R.) **22** (1943) 545, [*Prog. Theor. Phys.* **1** (1946) 27], first of a series of five papers
- [15] S. Tomonaga and J. R. Oppenheimer, *On Infinite Field Reactions in Quantum Field Theory*, *Phys. Rev.* **74** (1948) 224
- [16] J. Schwinger, *On Quantum-Electrodynamics and the Magnetic Moment of the Electron*, *Phys. Rev.* **73** (1948) 416
- [17] J. Schwinger, *Quantum Electrodynamics. I. A Covariant Formulation*, *Phys. Rev.* **74** (1948) 1439, first of a series of three papers
- [18] R. P. Feynman, *The Theory of Positrons*, *Phys. Rev.* **76** (1949) 749
- [19] R. P. Feynman, *Space-Time Approach to Quantum Electrodynamics*, *Phys. Rev.* **76** (1949) 769
- [20] R. P. Feynman, *Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction*, *Phys. Rev.* **80** (1950) 440
- [21] F. J. Dyson, *The Radiation Theories of Tomonaga, Schwinger, and Feynman*, *Phys. Rev.* **75** (1949) 486
- [22] F. J. Dyson, *The S Matrix in Quantum Electrodynamics*, *Phys. Rev.* **75** (1949) 1736
- [23] H. Yukawa, *On the Interaction of Elementary Particles. I.* *Proc. Phys.-Math. Soc. Jpn.* **17** (1935) 48
- [24] E. Fermi, *Versuch einer Theorie der β -Strahlen*, *Z. Phys.* **88** (1934) 161, [*Am. J. Phys.* **36** (1968) 1150]
- [25] W. Heisenberg, *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*, *Z. Phys.* **33** (1925) 879
- [26] M. Born and P. Jordan, *Zur Quantenmechanik*, *Z. Phys.* **34** (1925) 858
- [27] M. Born, W. Heisenberg, and P. Jordan, *Zur Quantenmechanik. II.* *Z. Phys.* **35** (1926) 557
- [28] P. A. M. Dirac, *The fundamental equations of quantum mechanics*, *Proc. R. Soc. Lond. A* **109** (1925) 642
- [29] E. Schrödinger, *Quantisierung als Eigenwertproblem*, *Ann. Phys.* **386** (1926) 361, first of a series of four papers
- [30] E. Schrödinger, *Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinen*, *Ann. Phys.* **386** (1926) 734
- [31] A. Einstein, *Zur Elektrodynamik bewegter Körper*, *Ann. Phys.* **322** (1905) 26

- [32] A. Einstein, *Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?*, *Ann. Phys.* **322** (1905) 639
- [33] W. Gerlach and O. Stern, *Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld*, *Z. Phys.* **9** (1922) 349
- [34] C. D. Anderson, *The Positive Electron*, *Phys. Rev.* **43** (1933) 491
- [35] C. T. R. Wilson, *On a method of making visible the paths of ionising particles through a gas*, *Proc. R. Soc. Lond. A* **85** (1911) 285
- [36] E. O. Lawrence and N. E. Edlefsen, *On the production of high speed protons*, *Science* **72** (1930) 372
- [37] S. Navas et al. (Particle Data Group), *Review of Particle Physics*, *Phys. Rev. D* **110** (2024) 030001
- [38] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132
- [39] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321
- [40] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508
- [41] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (1964) 585
- [42] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, *Phys. Rev.* **145** (1966) 1156
- [43] T. W. B. Kibble, *Symmetry Breaking in Non-Abelian Gauge Theories*, *Phys. Rev.* **155** (1967) 1554
- [44] E. Noether, *Invariante Variationsprobleme*, *Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse* (1918) 235, [*Transp. Theor. Statist. Phys.* **1** (1971) 186]
- [45] W. Heisenberg, *Über den Bau der Atomkerne. I.* *Z. Phys.* **77** (1932) 1, first of a series of three papers
- [46] B. Cassen and E. U. Condon, *On Nuclear Forces*, *Phys. Rev.* **50** (1936) 846
- [47] E. Wigner, *On the Consequences of the Symmetry of the Nuclear Hamiltonian on the Spectroscopy of Nuclei*, *Phys. Rev.* **51** (1937) 106
- [48] M. Gell-Mann, *Isotopic Spin and New Unstable Particles*, *Phys. Rev.* **92** (1953) 833
- [49] T. Nakano and K. Nishijima, *Charge Independence for V-particles*, *Prog. Theor. Phys.* **10** (1953) 581
- [50] K. Nishijima, *Charge Independence Theory of V Particles*, *Prog. Theor. Phys.* **13** (1955) 285
- [51] M. Gell-Mann, *The Interpretation of the New Particles as Displaced Charge Multiplets*, *Nuovo Cim.* **4** (1956) 848

Bibliography

- [52] M. Gell-Mann, *The Eightfold Way: A Theory of Strong Interaction Symmetry*, **CTSL-20** (Caltech Synchrotron Lab, Pasadena, 1961)
- [53] Y. Ne'eman, *Derivation of strong interactions from a gauge invariance*, **Nucl. Phys. 26** (1961) 222
- [54] E. D. Bloom et al., *High-Energy Inelastic $e-p$ Scattering at 6° and 10°* , **Phys. Rev. Lett. 23** (1969) 930
- [55] M. Breidenbach et al., *Observed Behavior of Highly Inelastic Electron-Proton Scattering*, **Phys. Rev. Lett. 23** (1969) 935
- [56] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, **Phys. Lett. 8** (1964) 214
- [57] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking*, **CERN-TH-412** (CERN, Geneva, 1964), version 2
- [58] M.-Y. Han and Y. Nambu, *Three-Triplet Model with Double $SU(3)$ Symmetry*, **Phys. Rev. 139** (1965) B1006
- [59] W. A. Bardeen, H. Fritzsch, and M. Gell-Mann, *Light-Cone current Algebra, π^0 Decay and e^+e^- Annihilation*, **CERN-TH-1538** (CERN, Geneva, 1972)
- [60] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, **Phys. Lett. B 47** (1973) 365
- [61] S. Weinberg, *Non-Abelian Gauge Theories of the Strong Interactions*, **Phys. Rev. Lett. 31** (1973) 494
- [62] D. J. Gross and F. Wilczek, *Asymptotically Free Gauge Theories. I*, **Phys. Rev. D 8** (1973) 3633, first of a series of two papers
- [63] TASSO Collaboration, *Evidence for planar events in e^+e^- annihilation at high energies*, **Phys. Lett. B 86** (1979) 243
- [64] MARK-J Collaboration, *Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA*, **Phys. Rev. Lett. 43** (1979) 830
- [65] PLUTO Collaboration, *Evidence for gluon bremsstrahlung in e^+e^- annihilations at high energies*, **Phys. Lett. B 86** (1979) 418
- [66] JADE Collaboration, *Observation of planar three-jet events in e^+e^- annihilation and evidence for gluon bremsstrahlung*, **Phys. Lett. B 91** (1980) 142
- [67] C. N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, **Phys. Rev. 96** (1954) 191
- [68] B. W. Lee and J. Zinn-Justin, *Spontaneously Broken Gauge Symmetries*, **Phys. Rev. D 5** (1972) 3121, first of a series of three papers
- [69] NNPDF Collaboration, *Parton distributions from high-precision collider data*, **Eur. Phys. J. C 77** (2017) 663
- [70] R. P. Feynman, *Very high-energy collisions of hadrons*, **Phys. Rev. Lett. 23** (1969) 1415

- [71] R. P. Feynman, *The behavior of hadron collisions at extreme energies*, *Proceedings of 3rd International Conference on High Energy Collisions* (Stony Book, New York, USA, Sept. 5–6, 1969), 237
- [72] J. D. Bjorken, *Asymptotic Sum Rules at Infinite Momentum*, **Phys. Rev.** **179** (1969) 1547
- [73] R. K. Ellis et al., *Perturbation Theory and the Parton Model in QCD*, **Nucl. Phys. B** **152** (1979) 285
- [74] J. C. Collins, D. E. Soper, and G. F. Sterman, *Does the Drell-Yan Cross-section Factorize?*, **Phys. Lett. B** **109** (1982) 388
- [75] J. C. Collins, D. E. Soper, and G. F. Sterman, *Relation of Parton Distribution Functions in Drell-Yan Process to Deeply Inelastic Scattering*, **Phys. Lett. B** **126** (1983) 275
- [76] G. T. Bodwin, *Factorization of the Drell-Yan Cross-Section in Perturbation Theory*, **Phys. Rev. D** **31** (1985) 2616, [Erratum: **Phys. Rev. D** **34** (1986) 3932]
- [77] V. N. Gribov and L. N. Lipatov, *Deep inelastic $e-p$ scattering in perturbation theory*, *Yad. Fiz.* **15** (1972) 781, [*Sov. J. Nucl. Phys.* **15** (1972) 438]
- [78] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics*. *Zh. Eksp. Teor. Fiz.* **73** (1977) 1216, [*Sov. Phys. J. Exp. Theor. Phys.* **46** (1977) 641]
- [79] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, **Nucl. Phys. B** **126** (1977) 298
- [80] S. D. Drell and T.-M. Yan, *Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies*, **Phys. Rev. Lett.** **25** (1970) 316, [Erratum: **Phys. Rev. Lett.** **25** (1970) 902]
- [81] T. D. Lee and C. N. Yang, *Question of Parity Conservation in Weak Interactions*, **Phys. Rev.** **104** (1956) 254
- [82] C. S. Wu et al., *Experimental Test of Parity Conservation in Beta Decay*, **Phys. Rev.** **105** (1957) 1413
- [83] E. C. G. Sudarshan and R. E. Marshak, *Chirality Invariance and the Universal Fermi Interaction*, **Phys. Rev.** **109** (1958) 1860
- [84] R. P. Feynman and M. Gell-Mann, *Theory of the Fermi Interaction*, **Phys. Rev.** **109** (1958) 193
- [85] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, **Phys. Rev. Lett.** **10** (1963) 531
- [86] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, **Prog. Theor. Phys.** **49** (1973) 652
- [87] J. H. Christenson et al., *Evidence for the 2π Decay of the K_2^0 Meson*, **Phys. Rev. Lett.** **13** (1964) 138

Bibliography

- [88] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, **Phys. Rev. Lett.** **53** (1984) 1802
- [89] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions*, **Phys. Rev. Lett.** **39** (1977) 252
- [90] M. L. Perl et al., *Evidence for Anomalous Lepton Production in e^+e^- Annihilation*, **Phys. Rev. Lett.** **35** (1975) 1489
- [91] J. Goldstone, *Field Theories with Superconductor Solutions*, **Nuovo Cim.** **19** (1961) 154
- [92] J. Goldstone, A. Salam, and S. Weinberg, *Broken Symmetries*, **Phys. Rev.** **127** (1962) 965
- [93] P. W. Anderson, *Plasmons, Gauge Invariance, and Mass*, **Phys. Rev.** **130** (1963) 439
- [94] S. Weinberg, *General Theory of Broken Local Symmetries*, **Phys. Rev. D** **7** (1973) 1068
- [95] J. J. Aubert et al., *Experimental Observation of a Heavy Particle J*, **Phys. Rev. Lett.** **33** (1974) 1404
- [96] J.-E. Augustin et al., *Discovery of a Narrow Resonance in e^+e^- Annihilation*, **Phys. Rev. Lett.** **33** (1974) 1406
- [97] F. J. Hasert et al., *Search for elastic muon-neutrino electron scattering*, **Phys. Lett. B** **46** (1973) 121
- [98] F. J. Hasert et al., *Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment*, **Phys. Lett. B** **46** (1973) 138
- [99] UA1 Collaboration, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, **Phys. Lett. B** **122** (1983) 103
- [100] UA2 Collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN $\bar{p}p$ Collider*, **Phys. Lett. B** **122** (1983) 476
- [101] UA1 Collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around $95 \text{ GeV}/c^2$ at the CERN SPS Collider*, **Phys. Lett. B** **126** (1983) 398
- [102] UA2 Collaboration, *Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ Collider*, **Phys. Lett. B** **129** (1983) 130
- [103] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, **Phys. Lett. B** **716** (2012) 1
- [104] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, **Phys. Lett. B** **716** (2012) 30
- [105] G. Steigman, *Observational Tests of Antimatter Cosmologies*, **Annu. Rev. Astron. Astrophys.** **14** (1976) 339

- [106] A. G. Cohen, A. De Rújula, and S. L. Glashow, *A Matter-Antimatter Universe?*, *Astrophys. J.* **495** (1998) 539
- [107] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, *Pis'ma Zh. Eksp. Teor. Fiz.* **5** (1967) 32, [*J. Exp. Theor. Phys. Lett.* **5** (1967) 24, *Usp. Fiz. Nauk* **161** (1991) 61, *Sov. Phys. Usp.* **34** (1991) 393]
- [108] G. 't Hooft, *Symmetry Breaking through Bell-Jackiw Anomalies*, *Phys. Rev. Lett.* **37** (1976) 8
- [109] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov, *On Anomalous Electroweak Baryon-Number Non-Conservation in the Early Universe*, *Phys. Lett. B* **155** (1985) 36
- [110] M. Fukugita and T. Yanagida, *Baryogenesis without Grand Unification*, *Phys. Lett. B* **174** (1986) 45
- [111] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, *Helv. Phys. Acta* **6** (1933) 110, [*Gen. Relativ. Gravit.* **41** (2009) 207]
- [112] J. Einasto, A. Kaasik, and E. Saar, *Dynamic evidence on massive coronas of galaxies*, *Nature* **250** (1974) 309, [Erratum: *Nature* **250** (1974) 790]
- [113] J. P. Ostriker, P. J. E. Peebles, and A. Yahil, *The Size and Mass of Galaxies, and the Mass of the Universe*, *Astrophys. J.* **193** (1974) L1
- [114] D. Clowe et al., *A Direct Empirical Proof of the Existence of Dark Matter*, *Astrophys. J.* **648** (2006) L109
- [115] P. J. E. Peebles, *Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations*, *Astrophys. J.* **263** (1982) L1
- [116] Planck Collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6, [Erratum: *Astron. Astrophys.* **652** (2021) C4]
- [117] É. B. Gliner, *Algebraic Properties of the Energy-Momentum Tensor and Vacuum-like States of Matter*, *Zh. Eksp. Teor. Fiz.* **49** (1966) 542, [*Sov. Phys. J. Exp. Theor. Phys.* **22** (1966) 378]
- [118] Y. B. Zel'dovich, *The cosmological constant and the theory of elementary particles*, *Usp. Fiz. Nauk* **95** (1968) 209, [*Sov. Phys. Usp.* **11** (1968) 381, *Gen. Rel. Grav.* **40** (2008) 1557]
- [119] B. Pontecorvo, *Inverse Beta Processes and Nonconservation of Lepton Charge*, *Zh. Eksp. Teor. Fiz.* **34** (1957) 247, [*Sov. Phys. J. Exp. Theor. Phys.* **7** (1958) 172]
- [120] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, *Prog. Theor. Phys.* **28** (1962) 870
- [121] SNO Collaboration, *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **89** (2002) 011301
- [122] J. M. Pendlebury et al., *Revised experimental upper limit on the electric dipole moment of the neutron*, *Phys. Rev. D* **92** (2015) 092003

Bibliography

- [123] B. Graner et al., *Reduced Limit on the Permanent Electric Dipole Moment of ^{199}Hg* , *Phys. Rev. Lett.* **116** (2016) 161601, [Erratum: *Phys. Rev. Lett.* **119** (2017) 119901]
- [124] K. G. Wilson, *The Renormalization Group and Strong Interactions*, *Phys. Rev. D* **3** (1971) 1818
- [125] L. Susskind, *Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory*, *Phys. Rev. D* **20** (1979) 2619
- [126] M. J. G. Veltman, *The Infrared-Ultraviolet Connection*, *Acta Phys. Polon. B* **12** (1981) 437
- [127] S. Dimopoulos and H. Georgi, *Softly broken supersymmetry and $SU(5)$* , *Nucl. Phys. B* **193** (1981) 150
- [128] G. Steigman and M. S. Turner, *Cosmological constraints on the properties of weakly interacting massive particles*, *Nucl. Phys. B* **253** (1985) 375
- [129] H. Pagels and J. R. Primack, *Supersymmetry, Cosmology, and New Physics at Teraelectronvolt Energies*, *Phys. Rev. Lett.* **48** (1982) 223
- [130] J. Wess and B. Zumino, *Supergauge Transformations in Four Dimensions*, *Nucl. Phys. B* **70** (1974) 39
- [131] H. E. Haber, G. L. Kane, and T. Sterling, *The Fermion Mass Scale and Possible Effects of Higgs Bosons on Experimental Observables*, *Nucl. Phys. B* **161** (1979) 493
- [132] J. F. Donoghue and L.-F. Li, *Properties of charged Higgs bosons*, *Phys. Rev. D* **19** (1979) 945
- [133] J. F. Gunion and H. E. Haber, *Higgs bosons in supersymmetric models (I)*, *Nucl. Phys. B* **272** (1986) 1, [Erratum: *Nucl. Phys. B* **402** (1993) 567], first of a series of three papers
- [134] D. López-Val, T. Plehn, and M. Rauch, *Measuring extended Higgs sectors as a consistent free couplings model*, *JHEP* **10** (2013) 134
- [135] C. Englert et al., *Precision measurements of Higgs couplings: implications for new physics scales*, *J. Phys. G* **41** (2014) 113001
- [136] CDF Collaboration, *Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab*, *Phys. Rev. Lett.* **74** (1995) 2626
- [137] DØ Collaboration, *Observation of the Top Quark*, *Phys. Rev. Lett.* **74** (1995) 2632
- [138] CLEO Collaboration, *Upper Limit on Flavor-Changing Neutral-Current Decays of the b Quark*, *Phys. Rev. Lett.* **53** (1984) 1309
- [139] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, *Comput. Phys. Commun.* **185** (2014) 2930
- [140] A. D. Martin et al., *Uncertainties on α_s in global PDF analyses and implications for predicted hadronic cross sections*, *Eur. Phys. J. C* **64** (2009) 653

- [141] J. Gao et al., *CT10 next-to-next-to-leading order global analysis of QCD*, *Phys. Rev. D* **89** (2014) 033009
- [142] NNPDF Collaboration, *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244
- [143] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, 2011, arXiv: 1101.0538 [hep-ph]
- [144] LHC Top Working Group, *LHCTopWG Summary Plots*, URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots> (visited on Aug. 21, 2024)
- [145] M. Aliev et al., *HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR*, *Comput. Phys. Commun.* **182** (2011) 1034
- [146] 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
- [147] J. Campbell, T. Neumann, and Z. Sullivan, *Single-top-quark production in the t -channel at NNLO*, *JHEP* **02** (2021) 040
- [148] N. Kidonakis and N. Yamanaka, *Higher-order corrections for tW production at high-energy hadron colliders*, *JHEP* **05** (2021) 278
- [149] PDF4LHC Working Group, *The PDF4LHC21 combination of global PDF fits for the LHC Run III*, *J. Phys. G* **49** (2022) 080501
- [150] S. Frixione et al., *Single-top hadroproduction in association with a W boson*, *JHEP* **07** (2008) 029
- [151] L.-B. Chen et al., *Analytic result for the top-quark width at next-to-next-to-leading order in QCD*, *Phys. Rev. D* **108** (2023) 054003
- [152] ATLAS Collaboration, *Measurement of the top-quark decay width in top-quark pair events in the dilepton channel at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *ATLAS-CONF-2019-038* (CERN, Geneva, 2019)
- [153] CMS Collaboration, *Measurement of the ratio $\mathcal{B}(t \rightarrow Wb)/\mathcal{B}(t \rightarrow Wq)$ in pp collisions at $\sqrt{s} = 8$ TeV*, *Phys. Lett. B* **736** (2014) 33
- [154] I. I. Y. Bigi et al., *Production and Decay Properties of Ultra-Heavy Quarks*, *Phys. Lett. B* **181** (1986) 157
- [155] A. F. Falk and M. E. Peskin, *Production, decay, and polarization of excited heavy hadrons*, *Phys. Rev. D* **49** (1994) 3320
- [156] J. S. Bell, *On the Einstein Podolsky Rosen paradox*, *Phys. Phys. Fiz.* **1** (1964) 195
- [157] Y. Afik and J. R. M. de Nova, *Entanglement and quantum tomography with top quarks at the LHC*, *Eur. Phys. J. Plus* **136** (2021) 907
- [158] M. Fabbrichesi, R. Floreanini, and G. Panizzo, *Testing Bell Inequalities at the LHC with Top-Quark Pairs*, *Phys. Rev. Lett.* **127** (2021) 161801

Bibliography

- [159] ATLAS Collaboration, *Observation of quantum entanglement with top quarks at the ATLAS detector*, **Nature** **633** (2024) 542
- [160] CMS Collaboration, *Observation of quantum entanglement in top quark pair production in proton–proton collisions at $\sqrt{s} = 13$ TeV*, **Rep. Prog. Phys.** **87** (2024) 117801
- [161] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, **Phys. Rev. Lett.** **114** (2015) 191803
- [162] ATLAS Collaboration, *Combined Measurement of the Higgs Boson Mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ Decay Channels with the ATLAS Detector Using $\sqrt{s} = 7, 8$, and 13 TeV pp Collision Data*, **Phys. Rev. Lett.** **131** (2023) 251802
- [163] CMS Collaboration, *A measurement of the Higgs boson mass in the diphoton decay channel*, **Phys. Lett. B** (2020) 135425
- [164] ATLAS Collaboration, *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, **Eur. Phys. J. C** **75** (2015) 476, [Erratum: **Eur. Phys. J. C** **76** (2016) 152]
- [165] ATLAS Collaboration, *Test of CP-invariance of the Higgs boson in vector-boson fusion production and in its decay into four leptons*, **JHEP** **05** (2024) 105
- [166] CMS Collaboration, *Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV*, **Phys. Rev. D** **92** (2015) 012004
- [167] CMS Collaboration, *Constraints on anomalous Higgs boson couplings to vector bosons and fermions in its production and decay using the four-lepton final state*, **Phys. Rev. D** **104** (2021) 052004
- [168] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, CERN Yellow Reports: Monographs, **CERN-2017-002-M** (CERN, Geneva, 2017)
- [169] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, **JHEP** **05** (2016) 058
- [170] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC*, **Phys. Lett. B** **726** (2013) 88, [Erratum: **Phys. Lett. B** **734** (2014) 406]
- [171] 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
- [172] CMS Collaboration, *Observation of the Diphoton Decay of the Higgs Boson and Measurement of Its Properties*, **Eur. Phys. J. C** **74** (2014) 3076
- [173] CMS Collaboration, *Measurement of the Properties of a Higgs Boson in the Four-Lepton Final State*, **Phys. Rev. D** **89** (2014) 092007

- [174] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, CERN Yellow Reports: Monographs, **CERN-2012-002** (CERN, Geneva, 2012)
- [175] E. W. N. Glover and J. J. van der Bij, *Higgs Boson Pair Production via Gluon Fusion*, **Nucl. Phys. B** **309** (1988) 282
- [176] D. A. Dicus, C. Kao, and S. S. Willenbrock, *Higgs Boson Pair Production from Gluon Fusion*, **Phys. Lett. B** **203** (1988) 457
- [177] ATLAS Collaboration, *A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector*, **Phys. Lett. B** **812** (2021) 135980
- [178] CMS Collaboration, *Evidence for Higgs boson decay to a pair of muons*, **JHEP** **01** (2021) 148
- [179] ATLAS Collaboration, *Direct constraint on the Higgs-charm coupling from a search for Higgs boson decays into charm quarks with the ATLAS detector*, **Eur. Phys. J. C** **82** (2022) 717
- [180] CMS Collaboration, *Search for Higgs Boson Decay to a Charm Quark-Antiquark Pair in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV*, **Phys. Rev. Lett.** **131** (2023) 061801
- [181] CMS Collaboration, *Search for Higgs Boson and Observation of Z Boson through their Decay into a Charm Quark-Antiquark Pair in Boosted Topologies in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV*, **Phys. Rev. Lett.** **131** (2023) 041801
- [182] ATLAS Collaboration, *A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery*, **Nature** **607** (2022) 52, [Erratum: **Nature** **612** (2022) E24]
- [183] CMS Collaboration, *A portrait of the Higgs boson by the CMS experiment ten years after the discovery*, **Nature** **607** (2022) 60, [Erratum: **Nature** **623** (2023) E4]
- [184] A. Azatov and J. Galloway, *Light custodians and Higgs physics in composite models*, **Phys. Rev. D** **85** (2012) 055013
- [185] S. Alekhin, A. Djouadi, and S. Moch, *The top quark and Higgs boson masses and the stability of the electroweak vacuum*, **Phys. Lett. B** **716** (2012) 214
- [186] S. Frixione and A. Mitov, *Determination of the top quark mass from leptonic observables*, **JHEP** **09** (2014) 012
- [187] F. Demartin et al., *tWH associated production at the LHC*, **Eur. Phys. J. C** **77** (2017) 34
- [188] F. Demartin et al., *Higgs production in association with a single top quark at the LHC*, **Eur. Phys. J. C** **75** (2015) 267
- [189] F. Maltoni, G. Ridolfi, and M. Ubiali, *b-initiated processes at the LHC: a reappraisal*, **JHEP** **07** (2012) 022, [Erratum: **JHEP** **04** (2013) 095]
- [190] J. Ellis et al., *Disentangling Higgs-Top Couplings in Associated Production*, **JHEP** **04** (2014) 004

Bibliography

- [191] F. Demartin et al., *Higgs characterisation at NLO in QCD: CP properties of the top-quark Yukawa interaction*, *Eur. Phys. J. C* **74** (2014) 3065
- [192] F. Boudjema et al., *Laboratory-frame observables for probing the top-Higgs boson interaction*, *Phys. Rev. D* **92** (2015) 015019
- [193] G. Degrossi et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*, *JHEP* **08** (2012) 098
- [194] D. Buttazzo et al., *Investigating the near-criticality of the Higgs boson*, *JHEP* **12** (2013) 089
- [195] A. Andreassen, W. Frost, and M. D. Schwartz, *Scale-invariant instantons and the complete lifetime of the standard model*, *Phys. Rev. D* **97** (2018) 056006
- [196] F. Bezrukov and M. Shaposhnikov, *Why Should We Care about the Top Quark Yukawa Coupling?*, *J. Exp. Theor. Phys.* **120** (2015) 335
- [197] ATLAS Collaboration, *Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **05** (2016) 160
- [198] CMS Collaboration, *Search for the associated production of the Higgs boson with a top-quark pair*, *JHEP* **09** (2014) 087, [Erratum: *JHEP* **10** (2014) 106]
- [199] CMS Collaboration, *Search for a standard model Higgs boson produced in association with a top-quark pair and decaying to bottom quarks using a matrix element method*, *Eur. Phys. J. C* **75** (2015) 251
- [200] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, *Phys. Lett. B* **784** (2018) 173
- [201] CMS Collaboration, *Observation of $t\bar{t}H$ production*, *Phys. Rev. Lett.* **120** (2018) 231801
- [202] CMS Collaboration, *Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **81** (2021) 378
- [203] ATLAS Collaboration, *CP Properties of Higgs Boson Interactions with Top Quarks in the $t\bar{t}H$ and tH Processes Using $H \rightarrow \gamma\gamma$ with the ATLAS Detector*, *Phys. Rev. Lett.* **125** (2020) 061802
- [204] CMS Collaboration, *Measurements of $t\bar{t}H$ Production and the CP Structure of the Yukawa Interaction between the Higgs Boson and Top Quark in the Diphoton Decay Channel*, *Phys. Rev. Lett.* **125** (2020) 061801
- [205] ATLAS Collaboration, *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, *Phys. Rev. D* **97** (2018) 072003
- [206] CMS Collaboration, *Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s} = 13$ TeV*, *JHEP* **08** (2018) 066

- [207] ATLAS Collaboration, *Probing the CP nature of the top–Higgs Yukawa coupling in $t\bar{t}H$ and tH events with $H \rightarrow b\bar{b}$ decays using the ATLAS detector at the LHC*, *Phys. Lett. B* **849** (2024) 138469
- [208] CMS Collaboration, *Measurement of the $t\bar{t}H$ and tH production rates in the $H \rightarrow b\bar{b}$ decay channel using proton–proton collision data at $\sqrt{s} = 13$ TeV*, *JHEP* **02** (2025) 097
- [209] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **75** (2015) 349
- [210] ATLAS Collaboration, *Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **97** (2018) 072016
- [211] 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
- [212] CMS Collaboration, *Search for $t\bar{t}H$ production in the all-jet final state in proton–proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **06** (2018) 101
- [213] CMS Collaboration, *Search for $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton–proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **03** (2019) 026
- [214] J. M. Butterworth et al., *Jet Substructure as a New Higgs Search Channel at the Large Hadron Collider*, *Phys. Rev. Lett.* **100** (2008) 242001
- [215] T. Plehn, G. P. Salam, and M. Spannowsky, *Fat Jets for a Light Higgs Boson*, *Phys. Rev. Lett.* **104** (2010) 111801
- [216] M. R. Buckley and D. Gonçalves, *Boosting the Direct CP Measurement of the Higgs-Top Coupling*, *Phys. Rev. Lett.* **116** (2016) 091801
- [217] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, CERN Yellow Reports: Monographs, **CERN-2013-004** (CERN, Geneva, 2013)
- [218] LHC Higgs Cross Section Working Group, *Simplified Template Cross Sections – Stage 1.1*, **LHCHSWG-2019-003** (CERN, Geneva, 2019)
- [219] ATLAS Collaboration, *Studies of Monte Carlo predictions for the $t\bar{t}b\bar{b}$ process*, **ATL-PHYS-PUB-2022-006** (CERN, Geneva, 2022)
- [220] LHC Higgs Working Group, *Study of $t\bar{t}b\bar{b}$ and $t\bar{t}W$ background modelling for $t\bar{t}H$ analyses*, **LHCHWG-2022-003**, 2022
- [221] S. Höche, J. Krause, and F. Siegert, *Multijet merging in a variable flavor number scheme*, *Phys. Rev. D* **100** (2019) 014011
- [222] L. Ferencz et al., *$t\bar{t}b\bar{b}$ at NLO precision in a variable flavor number scheme*, *JHEP* **07** (2024) 026

Bibliography

- [223] ATLAS Collaboration, *Measurement of $t\bar{t}$ production in association with additional b -jets in the $e\mu$ final state in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, **JHEP** **01** (2025) 068
- [224] CMS Collaboration, *Inclusive and differential cross section measurements of $t\bar{t}b\bar{b}$ production in the lepton+jets channel at $\sqrt{s} = 13$ TeV*, **JHEP** **05** (2024) 042
- [225] L. Evans and P. Bryant (eds.), *LHC Machine*, **JINST** **3** (2008) S08001
- [226] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, **JINST** **3** (2008) S08003
- [227] CMS Collaboration, *The CMS Experiment at the CERN LHC*, **JINST** **3** (2008) S08004
- [228] LHCb Collaboration, *The LHCb Detector at the LHC*, **JINST** **3** (2008) S08005
- [229] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, **JINST** **3** (2008) S08002
- [230] LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, **JINST** **3** (2008) S08006
- [231] TOTEM Collaboration, *The TOTEM Experiment at the CERN Large Hadron Collider*, **JINST** **3** (2008) S08007
- [232] MoEDAL Collaboration, *Technical Design Report of the MoEDAL Experiment*, **CERN-LHCC-2009-006** (CERN, Geneva, 2009)
- [233] FASER Collaboration, *The FASER detector*, **JINST** **19** (2024) P05066
- [234] SND@LHC Collaboration, *SND@LHC: the scattering and neutrino detector at the LHC*, **JINST** **19** (2024) P05067
- [235] S. Myers and E. Picasso (eds.), *The design, construction and commissioning of the CERN large Electron-Positron collider*, **Contemp. Phys.** **31** (1990) 387
- [236] O. S. Brüning et al. (eds.), *LHC Design Report Vol. 1: The LHC Main Ring*, **CERN-2004-003-V-1** (CERN, Geneva, 2004)
- [237] R. Alemany-Fernandez et al., *Operation and Configuration of the LHC in Run 1*, **CERN-ACC-NOTE-2013-0041** (CERN, Geneva, 2013)
- [238] J. Wenninger, *Operation and Configuration of the LHC in Run 2*, **CERN-ACC-NOTE-2019-0007** (CERN, Geneva, 2019)
- [239] T. Baer et al., *UFOs in the LHC, Proceedings of the 2nd International Particle Accelerator Conference* (San Sebastián, Spain, Sept. 4–9, 2011), **CERN-ATS-2011-051**
- [240] G. Papotti et al., *Macroparticle-Induced Losses During 6.5 TeV LHC Operation, Proceedings of the 7th International Particle Accelerator Conference* (Busan, Korea, May 8–13, 2016), **CERN-ACC-2016-249**
- [241] M. Benedikt et al. (eds.), *LHC Design Report Vol. 3: The LHC Injector Chain*, **CERN-2004-003-V-3** (CERN, Geneva, 2004)

- [242] E. Boltezar et al., *The new CERN 50 MeV Linac*, **CERN-PS-LR-79-28** (CERN, Geneva, 1979)
- [243] M. Vretenar (ed.), *Linac4 Design Report*, CERN Yellow Reports: Monographs, **CERN-2020-006** (CERN, Geneva, 2020)
- [244] ATLAS Collaboration, *Public ATLAS Luminosity Results for Run-2 of the LHC*, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2> (visited on Aug. 1, 2024)
- [245] ATLAS Collaboration, *Public ATLAS Luminosity Results for Run-3 of the LHC*, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun3> (visited on Nov. 27, 2024)
- [246] ATLAS Collaboration, *LuminosityPublicResultsRun1*, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun1> (visited on Aug. 6, 2024)
- [247] F. Alessio, F. Moortgat, and C. Young, *LHC experiments: main goals for the rest of Run 3, Chamonix Workshop 2024* (Chamonix, France, Jan. 29–Feb. 1, 2024), **LHC Programme Coordination report**
- [248] ATLAS Collaboration, *New Small Wheel Technical Design Report*, **CERN-LHCC-2013-006** (CERN, Geneva, 2013)
- [249] ATLAS Collaboration, *ATLAS Inner Detector: Technical Design Report, Volume I*, **CERN-LHCC-97-016** (CERN, Geneva, 1997)
- [250] ATLAS Collaboration, *ATLAS Inner Detector: Technical Design Report, Volume II*, **CERN-LHCC-97-017** (CERN, Geneva, 1997)
- [251] D. Froidevaux and M. A. Parker, *The Performance Specifications of the ATLAS Inner Detector*, **ATL-INDET-94-046** (CERN, Geneva, 1994)
- [252] ATLAS Collaboration, *ATLAS Pixel Detector: Technical Design Report*, **CERN-LHCC-98-013** (CERN, Geneva, 1998)
- [253] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, **CERN-LHCC-2010-013** (CERN, Geneva, 2010), [Addendum: **CERN-LHCC-2012-009**]
- [254] G. Aad et al., *ATLAS pixel detector electronics and sensors*, **JINST 3 (2008) P07007**
- [255] ATLAS IBL Collaboration, *Production and Integration of the ATLAS Insertable B-Layer*, **JINST 13 (2018) T05008**
- [256] A. Ahmad et al., *The Silicon Microstrip Sensors of the ATLAS Semiconductor Tracker*, **Nucl. Instrum. Meth. A 578 (2007) 98**
- [257] ATLAS TRT Collaboration, *The ATLAS TRT Barrel Detector*, **JINST 3 (2008) P02014**
- [258] ATLAS TRT Collaboration, *The ATLAS TRT end-cap detectors*, **JINST 3 (2008) P10003**

Bibliography

- [259] ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance*, [JINST 3 \(2008\) P02013](#)
- [260] K. Mistry (on behalf of the ATLAS Collaboration), *ATLAS Transition Radiation Tracker (TRT) Electronics Operation Experience at High Rates*, Poster at the 2015 Topical Workshop on Electronics for Particle Physics (Lisbon, Portugal, Sept. 28–Oct. 2, 2015), [ATL-INDET-SLIDE-2015-712](#)
- [261] G. M. Garibyan, *Transition Radiation Effects in Particle Energy Losses*, Zh. Eksp. Teor. Fiz. **37** (1959) 527, [Sov. Phys. J. Exp. Theor. Phys. **37** (1960) 372]
- [262] ATLAS Collaboration, *ATLAS Liquid Argon Calorimeter: Technical Design Report*, [CERN-LHCC-96-041](#) (CERN, Geneva, 1996)
- [263] ATLAS Collaboration, *ATLAS Tile Calorimeter: Technical Design Report*, [CERN-LHCC-96-042](#) (CERN, Geneva, 1996)
- [264] ATLAS Collaboration, *ATLAS Calorimeter Performance: Technical Design Report*, [CERN-LHCC-96-040](#) (CERN, Geneva, 1996)
- [265] ATLAS Collaboration, *ATLAS Muon Spectrometer: Technical Design Report*, [CERN-LHCC-97-022](#) (CERN, Geneva, 1997)
- [266] ATLAS Collaboration, *ATLAS Magnet System: Technical Design Report, Volume 1*, [CERN-LHCC-97-018](#) (CERN, Geneva, 1997)
- [267] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, [JINST 15 \(2020\) P10004](#)
- [268] ATLAS Collaboration, *ATLAS Level-1 Trigger: Technical Design Report*, [CERN-LHCC-98-014](#) (CERN, Geneva, 1998)
- [269] ATLAS Collaboration, *ATLAS High-Level Trigger, Data Acquisition and Controls: Technical Design Report*, [CERN-LHCC-2003-022](#) (CERN, Geneva, 2003)
- [270] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, [CERN-LHCC-2013-018](#) (CERN, Geneva, 2013)
- [271] ATLAS Collaboration, *ATLAS Fast TracKer (FTK) Technical Design Report*, [CERN-LHCC-2013-007](#) (CERN, Geneva, 2013)
- [272] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider: A Description of the Detector Configuration for Run 3*, [JINST 19 \(2024\) P05063](#)
- [273] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, [Eur. Phys. J. C 77 \(2017\) 317](#)
- [274] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, [Eur. Phys. J. C 76 \(2016\) 653](#)
- [275] V. Cindro et al., *The ATLAS Beam Conditions Monitor*, [JINST 3 \(2008\) P02004](#)
- [276] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, [JINST 13 \(2018\) P07017](#)

- [277] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, *Eur. Phys. J. C* **83** (2023) 982
- [278] S. van der Meer, *Calibration of the effective beam height in the ISR*, *CERN-ISR-PO-68-31* (CERN, Geneva, 1968)
- [279] A. Polini et al., *Design of the ATLAS IBL Readout System*, *Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics* (Chicago, Illinois, USA, June 9–14, 2011), *Phys. Procedia* **37** (2012) 1948
- [280] I. Perić et al., *The FEI3 readout chip for the ATLAS pixel detector*, *Proceedings of the 3rd International Workshop on Semiconductor Pixel Detectors for Particles and Imaging* (Bonn, Germany, Sept. 5–8, 2005), *Nucl. Instrum. Meth. A* **565** (2006) 178
- [281] M. Garcia-Sciveres et al., *The FE-I4 pixel readout integrated circuit*, *Proceedings of the 7th International ‘Hiroshima’ Symposium on the Development and Application of Semiconductor Tracking Detectors* (Hiroshima, Japan, Aug. 29–Sept. 1, 2009), *Nucl. Instrum. Meth. A* **636** (2011) S155
- [282] V. Zivkovic et al., *The FE-I4 pixel readout system-on-chip resubmission for the insertable B-Layer project*, *Proceedings of the 2011 Topical Workshop on Electronics for Particle Physics* (Vienna, Austria, Sept. 26–30, 2011), *JINST* **7** (2012) C02050
- [283] R. Beccherle et al., *MCC: the Module Controller Chip for the ATLAS Pixel Detector*, *Nucl. Instrum. Meth. A* **492** (2002) 117
- [284] S. Parker, C. Kenney, and J. Segal, *3D – A proposed new architecture for solid-state radiation detectors*, *Nucl. Instrum. Meth. A* **395** (1997) 328
- [285] A. Gabrielli et al., *Firmware development and testing of the ATLAS Pixel Detector / IBL ROD card*, *Proceedings of the 7th International Workshop on Semiconductor Pixel Detectors for Particles and Imaging* (Niagara Falls, Canada, Sept. 1–5, 2014), *JINST* **10** (2015) C03037
- [286] M. Wensing (on behalf of the ATLAS Collaboration), *Firmware development and testing of the ATLAS IBL Back-Of-Crate card*, *Proceedings of the 3rd International Conference on Technology and Instrumentation in Particle Physics* (Amsterdam, Netherlands, June 2–6, 2014), *PoS TIPP2014* (2014) 216
- [287] G. Balbi et al., *Commissioning of the read-out driver (ROD) card for the ATLAS IBL detector and upgrade studies for the pixel Layers 1 and 2*, *Proceedings of the 9th International ‘Hiroshima’ Symposium on the Development and Application of Semiconductor Tracking Detectors* (Hiroshima, Japan, Sept. 2–5, 2013), *Nucl. Instrum. Meth. A* **765** (2014) 232
- [288] A. Gabrielli (on behalf of the ATLAS Pixel Detector Collaboration), *Commissioning of ROD boards for the entire ATLAS Pixel Detector*, *JINST* **13** (2018) T09009

Bibliography

- [289] G. Balbi et al., *A PowerPC-based control system for the Read-Out-Driver module of the ATLAS IBL*, *Proceedings of the 2011 Topical Workshop on Electronics for Particle Physics* (Vienna, Austria, Sept. 26–30, 2011), **JINST 7 (2012) C02016**
- [290] ATLAS Collaboration, *Threshold Tuning of the ATLAS Pixel Detector*, **ATL-INDET-PUB-2010-001** (CERN, Geneva, 2010)
- [291] M. Keil (on behalf of the ATLAS Collaboration), *Online Calibration and Performance of the ATLAS Pixel Detector*, *Proceedings of the 5th International Workshop on Semiconductor Pixel Detectors for Particles and Imaging* (Grindelwald, Switzerland, Sept. 6–10, 2010), **Nucl. Instrum. Meth. A 650 (2011) 9**
- [292] J. Große-Knetter, *Vertex Measurement at a Hadron Collider – The ATLAS Pixel Detector*, Habilitation, **BONN-IR-2008-04**, Bonn University, 2008
- [293] M. Karagounis et al., *Development of the ATLAS FE-I4 pixel readout IC for b-layer Upgrade and Super-LHC*, *Proceedings of the 2008 Topical Workshop on Electronics for Particle Physics* (Naxos, Greece, Sept. 15–19, 2008), **CERN-2008-008, 70**
- [294] R. Ushioda, *Analysis of Timing Scans for 2023*, *Talk at the May 2023 ATLAS Pixel Week* (CERN, Geneva, Switzerland, May 2–4, 2023), URL: <https://indico.cern.ch/event/1276979/#189-analysis-of-pixel-timing-s>
- [295] J. P. Iddon (on behalf of the ATLAS Collaboration), *Operational experience and performance with the ATLAS Pixel detector at the Large Hadron Collider at CERN*, **Nucl. Instrum. Meth. A 1063 (2024) 169312**
- [296] ATLAS Collaboration, *Pixel ROD-level desynchronization in 2023*, **PIX-2023-004** (CERN, Geneva, 2023)
- [297] N. Giangiacomi, *ATLAS Pixel Detector and readout upgrades for the improved LHC performance*, PhD thesis, **CERN-THESIS-2018-437**, University of Bologna, 2018
- [298] S. Roychoudhury, *FPGA Design Upgrades for the ATLAS Pixel Readout System in the Large Hadron Collider*, MSc thesis, **CERN-THESIS-2023-060**, University of Washington, 2023
- [299] M. Menouni et al., *SEU tolerant memory design for the ATLAS pixel readout chip*, *Proceedings of the 2012 Topical Workshop on Electronics for Particle Physics* (Oxford, United Kingdom, Sept. 17–21, 2012), **JINST 8 (2013) C02026**
- [300] G. Balbi et al., *Measurements of Single Event Upset in ATLAS IBL*, **JINST 15 (2020) P06023**
- [301] M. Bindi, *2024 Fw status (plans)*, *Talk at the March 2024 ATLAS Pixel Week* (CERN, Geneva, Switzerland, Mar. 19–21, 2024), URL: <https://indico.cern.ch/event/1378115/#226-rod-boc-fw>
- [302] ATLAS Collaboration, *IBL ToT and threshold evolution in 2017 and 2018*, **PIX-2018-009** (CERN, Geneva, 2018)

- [303] A. Sciandra, *Recap of 2023 Pixel-Detector Operating Conditions, Talk at the May 2023 ATLAS Pixel Week* (CERN, Geneva, Switzerland, May 2–4, 2023), URL: <https://indico.cern.ch/event/1276979/#214-recap-of-2023-pixel-detect>
- [304] ATLAS Collaboration, *Pixel Hit-on-Track Efficiency in 2023*, **PIX-2023-003** (CERN, Geneva, 2023)
- [305] ATLAS Pixel Operations Team, *atlaspixeldaQ*, GitLab repository, URL: <https://gitlab.cern.ch/atlas-pixel/daq/atlaspixeldaQ>
- [306] P. Buncic et al., *CernVM – a virtual software appliance for LHC applications*, **J. Phys. Conf. Ser.** **219** (2010) 042003
- [307] ATLAS Pixel Operations Team, *pixelDAQtest*, GitLab repository, URL: <https://gitlab.cern.ch/atlas-pixel/daq/pixeldaqtest>
- [308] G. Avolio, M. Caprini, and G. Lehmann Miotto, *The integrated graphical user interface of the trigger and data acquisition system of the ATLAS experiment at the LHC, Proceedings of the 18th International Conference on Computing in High Energy and Nuclear Physics* (Taipei, Taiwan, Oct. 18–22, 2010), **J. Phys. Conf. Ser.** **331** (2011) 022032
- [309] G. Lehmann Miotto et al., *Configuration and control of the ATLAS trigger and data acquisition, Proceedings of the 1st International Conference on Technology and Instrumentation in Particle Physics* (Tsukuba, Japan, Mar. 12–17, 2009), **Nucl. Instrum. Meth. A** **623** (2010) 549
- [310] ATLAS Pixel Operations Team, *PixelIBL Testing Note*, 2019
- [311] A. Vaswani et al., *Attention Is All You Need*, (2017), arXiv: [1706.03762](https://arxiv.org/abs/1706.03762) [[cs.CL](#)]
- [312] W. S. McCulloch and W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, **Bull. Math. Biol.** **5** (1943) 115
- [313] F. Rosenblatt, *The perceptron: a probabilistic model for information storage and organization in the brain*, **Psychol. Rev.** **65** (1958) 386
- [314] M. Minsky and S. A. Papert, *Perceptrons: An Introduction to Computational Geometry*, Cambridge, United States: MIT Press, 1969
- [315] K.-I. Funahashi, *On the approximate realization of continuous mappings by neural networks*, **Neural Netw.** **2** (1989) 183
- [316] D. Rumelhart, G. Hinton, and R. Williams, *Learning representations by back-propagating errors*, **Nature** **323** (1986) 533
- [317] B. T. Polyak, *Some methods of speeding up the convergence of iteration methods*, **Zh. Vychisl. Mat. Mat. Fiz.** **4** (1964) 791, [**USSR Comput. Math. Math. Phys.** **4** (1964) 1]
- [318] Y. E. Nesterov, *A method of solving a convex programming problem with convergence rate $O(1/k^2)$* , **Dokl. Akad. Nauk SSSR** **269** (1983) 542, [**Sov. Math., Dokl.** **27** (1983) 372]

Bibliography

- [319] I. Sutskever et al., *On the importance of initialization and momentum in deep learning*, *Proceedings of the 30th International Conference on Machine Learning* (Atlanta, Georgia, USA, June 17–19, 2013), *Proc. Mach. Learn. Res.* **28** (2013) 1139
- [320] J. Duchi, E. Hazan, and Y. Singer, *Adaptive Subgradient Methods for Online Learning and Stochastic Optimization*, *J. Mach. Learn. Res.* **12** (2011) 2121
- [321] T. Tieleman and G. Hinton, *rmsprop: Divide the gradient by a running average of its recent magnitude*, Lecture 6e in *Neural Networks for Machine Learning*, [video lecture on Coursera](#), 2012
- [322] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, *Proceedings of the 3rd International Conference on Learning Representations* (San Diego, California, USA, May 7–9, 2015), arXiv: 1412.6980 [cs.LG]
- [323] M. J. Fenton et al., *Permutationless many-jet event reconstruction with symmetry preserving attention networks*, *Phys. Rev. D* **105** (2022) 112008
- [324] ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, *Eur. Phys. J. C* **83** (2023) 681
- [325] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, *Nucl. Phys. B* **238** (1984) 492
- [326] G. Marchesini and B. R. Webber, *Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation*, *Nucl. Phys. B* **310** (1988) 461
- [327] B. Andersson et al., *Parton Fragmentation and String Dynamics*, *Phys. Rep.* **97** (1983) 31
- [328] T. Sjöstrand, *Jet Fragmentation of Nearby Partons*, *Nucl. Phys. B* **248** (1984) 469
- [329] ATLAS Collaboration, *Measurement of the underlying event in jet events from 7 TeV proton–proton collisions with the ATLAS detector*, *Eur. Phys. J. C* **74** (2014) 2965
- [330] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823
- [331] GEANT4 Collaboration, *Geant4 – a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250
- [332] K. Edmonds et al., *The fast ATLAS track simulation (FATRAS)*, [ATL-SOFT-PUB-2008-001](#) (CERN, Geneva, 2008)
- [333] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, [ATL-PHYS-PUB-2010-013](#) (CERN, Geneva, 2010)
- [334] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159
- [335] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*, [ATL-PHYS-PUB-2016-017](#) (CERN, Geneva, 2016)

- [336] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152
- [337] M. Bähr et al., *Herwig++ physics and manual*, *Eur. Phys. J. C* **58** (2008) 639
- [338] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196
- [339] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, *ATL-PHYS-PUB-2014-021* (CERN, Geneva, 2014)
- [340] L. A. Harland-Lang et al., *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J. C* **75** (2015) 204
- [341] 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
- [342] S. Frixione et al., *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, *JHEP* **04** (2007) 081
- [343] P. Artoisenet et al., *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, *JHEP* **03** (2013) 015
- [344] R. Brun and F. Rademakers, *ROOT – An object oriented data analysis framework*, *Nucl. Instrum. Meth. A* **389** (1997) 81
- [345] ATLAS $t\bar{t}(H \rightarrow b\bar{b})$ Analysis Team, *TTHbbAnalysis*, GitLab repository, URL: https://gitlab.cern.ch/atlasHTop/TTHbbAnalysis/-/tree/TTHbb21.2.238_Run2_Legacy_v2.5.1
- [346] ATLAS Top Reconstruction Sub-Group, *Top Reco Documentation: AnalysisTop-21.2*, URL: <https://top-reco.docs.cern.ch/AnalysisTop/AnalysisTop-21-2-X/> (visited on Sep. 2, 2024)
- [347] ATLAS Top Reconstruction Sub-Group, *AnalysisTop21*, release notes, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AnalysisTop21> (visited on Sep. 2, 2024)
- [348] I. Bird et al. (eds.), *LHC computing Grid. Technical design report*, *CERN-LHCC-2005-024* (CERN, Geneva, 2005)
- [349] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, *JINST* **15** (2020) P04003
- [350] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040
- [351] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070
- [352] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043
- [353] S. Frixione, P. Nason, and G. Ridolfi, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, *JHEP* **09** (2007) 126

Bibliography

- [354] H. B. Hartanto et al., *Higgs boson production in association with top quarks in the POWHEG BOX*, *Phys. Rev. D* **91** (2015) 094003
- [355] NNPDF Collaboration, *Parton distributions for the LHC run II*, *JHEP* **04** (2015) 040
- [356] S. Höche et al., *A study of QCD radiation in VBF Higgs production with VINCIA and PYTHIA*, *SciPost Phys.* **12** (2022) 010
- [357] T. Ježo and P. Nason, *On the treatment of resonances in next-to-leading order calculations matched to a parton shower*, *JHEP* **12** (2015) 065
- [358] T. Ježo et al., *New NLOPS predictions for $t\bar{t} + b$ -jet production at the LHC*, *Eur. Phys. J. C* **78** (2018) 502
- [359] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601
- [360] A. Denner, S. Dittmaier, and L. Hofer, *COLLIER: a fortran-based complex one-loop library in extended regularizations*, *Comput. Phys. Commun.* **212** (2017) 220
- [361] F. Buccioni et al., *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) 866
- [362] R. Frederix, E. Re, and P. Torrielli, *Single-top t -channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*, *JHEP* **09** (2012) 130
- [363] S. Alioli et al., *NLO single-top production matched with shower in POWHEG: s - and t -channel contributions*, *JHEP* **09** (2009) 111, [Erratum: *JHEP* **02** (2010) 011]
- [364] E. Re, *Single-top Wt -channel production matched with parton showers using the POWHEG method*, *Eur. Phys. J. C* **71** (2011) 1547
- [365] ATLAS Collaboration, *Studies of $t\bar{t}/tW$ interference effects in $b\bar{b}\ell^+\ell'^-\nu\bar{\nu}'$ final states with Powheg and MadGraph5_aMC@NLO setups*, *ATL-PHYS-PUB-2021-042* (CERN, Geneva, 2021)
- [366] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, *ATL-PHYS-PUB-2016-020* (CERN, Geneva, 2016)
- [367] *Wrong (no) top quark lineshape in Powheg ST_wtch_DR*, JIRA ticket, URL: <https://its.cern.ch/jira/browse/AGENE-2210> (visited on Sep. 3, 2024)
- [368] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007
- [369] E. Bothmann et al., *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034
- [370] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039
- [371] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, *JHEP* **03** (2008) 038
- [372] S. Catani et al., *QCD Matrix Elements + Parton Showers*, *JHEP* **11** (2001) 063

- [373] S. Höche et al., *QCD matrix elements and truncated showers*, **JHEP** **05** (2009) 053
- [374] S. Höche et al., *QCD matrix elements + parton showers. The NLO case*, **JHEP** **04** (2013) 027
- [375] L. Buonocore et al., *Precise Predictions for the Associated Production of a W Boson with a Top-Antitop Quark Pair at the LHC*, **Phys. Rev. Lett.** **131** (2023) 231901
- [376] M. van Beekveld, A. Kulesza, and L. M. Valero, *Threshold Resummation for the Production of Four Top Quarks at the LHC*, **Phys. Rev. Lett.** **131** (2023) 211901
- [377] C. Anastasiou et al., *High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order*, **Phys. Rev. D** **69** (2004) 094008
- [378] ATLAS Collaboration, *Multi-Boson Simulation for 13 TeV ATLAS Analyses*, **ATL-PHYS-PUB-2016-002** (CERN, Geneva, 2016)
- [379] ATLAS Collaboration, *Observation of four-top-quark production in the multilepton final state with the ATLAS detector*, **Eur. Phys. J. C** **83** (2023) 496, [Erratum: **Eur. Phys. J. C** **84** (2024) 156]
- [380] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k_T jet clustering algorithm*, **JHEP** **04** (2008) 063
- [381] *HadronOriginClassifier*, Athena GitLab repository, URL: <https://gitlab.cern.ch/atlas/athena/-/blob/0d4c2a/PhysicsAnalysis/DerivationFramework/DerivationFrameworkMCTruth/src/HadronOriginClassifier.cxx> (commit hash 0d4c2aebc2b5d11b0deffe2b5a35e083d359337f)
- [382] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, **Eur. Phys. J. C** **80** (2020) 47
- [383] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, **JINST** **15** (2020) P09015
- [384] ATLAS Collaboration, *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*, **ATL-DAQ-PUB-2016-001** (CERN, Geneva, 2016)
- [385] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, **Eur. Phys. J. C** **77** (2017) 466
- [386] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, **Eur. Phys. J. C** **72** (2012) 1896
- [387] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, **Eur. Phys. J. C** **76** (2016) 581
- [388] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV*, **ATL-PHYS-PUB-2015-026** (CERN, Geneva, 2015)

Bibliography

- [389] ATLAS Collaboration, *Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector*, [ATLAS-CONF-2015-029](#) (CERN, Geneva, 2015)
- [390] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, [JINST 14 \(2019\) P12006](#)
- [391] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, [Eur. Phys. J. C 81 \(2021\) 578](#)
- [392] ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, [ATL-PHYS-PUB-2015-045](#) (CERN, Geneva, 2015)
- [393] ATLAS Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, [ATL-PHYS-PUB-2019-033](#) (CERN, Geneva, 2019)
- [394] ATLAS Collaboration, *The performance of missing transverse momentum reconstruction and its significance with the ATLAS detector using 140 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions*, [Eur. Phys. J. C 85 \(2025\) 606](#)
- [395] *Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014*, [ATL-PHYS-INT-2014-018](#) (CERN, Geneva, 2014), ATLAS internal note
- [396] B. Nachman et al., *Jets from jets: re-clustering as a tool for large radius jet reconstruction and grooming at the LHC*, [JHEP 02 \(2015\) 075](#)
- [397] A. G. Borbély, *Measurement of the associated production of a top quark pair and a Higgs boson ($t\bar{t}H$) with boosted topologies*, PhD thesis, [glathesis:2023-83710](#), University of Glasgow, 2023
- [398] R. Barlow, *Extended maximum likelihood*, [Nucl. Instrum. Meth. A 297 \(1990\) 496](#)
- [399] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, [Ann. Math. Stat. 9 \(1938\) 60](#)
- [400] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, [Eur. Phys. J. C 71 \(2011\) 1554](#), [Erratum: [Eur. Phys. J. C 73 \(2013\) 2501](#)]
- [401] TRExFitter Developer Team, *Documentation for TRExFitter*, URL: <https://trefitter-docs.web.cern.ch/trefitter-docs/> (visited on Sep. 9, 2024)
- [402] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, *Proceedings of the 13th International Conference on Computing in High-Energy and Nuclear Physics* (San Diego, California, USA, Mar. 24–28, 2003), [eConf C0303241 \(2003\) MOLT007](#)

- [403] L. Moneta et al., *The RooStats Project, Proceedings of the 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research* (Jaipur, India, Feb. 22–27, 2010), **PoS ACAT2010 (2010) 057**
- [404] K. Cranmer et al. (ROOT Collaboration), *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, **CERN-OPEN-2012-016**, 2012
- [405] F. James and M. Winkler, *Minuit2 User's Guide*, URL: <https://root.cern/root/html/doc/guides/minuit2/Minuit2.html> (visited on Sep. 9, 2024)
- [406] ATLAS Collaboration, *Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, **JINST 18 (2023) T11004**
- [407] ATLAS $t\bar{t}H(b\bar{b})$ Leptonic Analysis Team, *Measurement of the $t\bar{t}H$ production cross-section in the $H \rightarrow b\bar{b}$ decay channel using 140 fb^{-1} of pp collision data at $\sqrt{s} = 13\text{ TeV}$* , **ATL-COM-PHYS-2021-985** (CERN, Geneva, 2024), ATLAS internal note
- [408] ATLAS Collaboration, *Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated $(W/Z)H$ production with the ATLAS detector*, **JHEP 01 (2015) 069**
- [409] *Decoupled b-jet-energy corrections (muon-in-jet, neutrino-in-jet, and transverse momentum) for $H \rightarrow b\bar{b}$ searches for ATLAS in early Run II*, **ATL-COM-PHYS-2016-035** (CERN, Geneva, 2016), ATLAS internal note
- [410] N. Kumari, *Unravelling the top-Higgs coupling with the ATLAS experiment at LHC*, PhD thesis, **CERN-THESIS-2022-279**, Aix-Marseille University, 2022
- [411] J. L. Ba, J. R. Kiros, and G. E. Hinton, *Layer Normalization*, 2016, arXiv: **1607.06450 [stat.ML]**
- [412] G. E. Hinton et al., *Improving neural networks by preventing co-adaptation of feature detectors*, 2012, arXiv: **1207.0580 [cs.NE]**
- [413] K. He et al., *Deep Residual Learning for Image Recognition*, *Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, Nevada, USA, June 27–30, 2016), arXiv: **1512.03385 [cs.CV]**
- [414] P. Ramachandran, B. Zoph, and Q. V. Le, *Searching for Activation Functions*, 2017, arXiv: **1710.05941 [cs.NE]**
- [415] F. Chollet et al., *Keras*, URL: <https://keras.io>, 2015
- [416] M. Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation* (Savannah, Georgia, USA, Nov. 2–4, 2016), URL: <https://tensorflow.org>
- [417] ONNX Runtime developers, *ONNX Runtime*, URL: <https://onnxruntime.ai/> (visited on Apr. 29, 2025)
- [418] *Neural Network for $t\bar{t}Hbb$* , GitLab repository, URL: <https://gitlab.cern.ch/atlasHTop/ANA-HIGG-2020-24/neural-network-for-tthbb/-/tree/f538a9>

Bibliography

- [419] *JetTruthMatchingTool*, TTHbbAnalysis GitLab repository, URL: <https://gitlab.cern.ch/atlasHTop/TTHbbAnalysis/-/blob/TTHbb212238/TTHbbLeptonic/Root/JetTruthMatchingTool.cxx>
- [420] C. Szegedy et al., *Rethinking the Inception Architecture for Computer Vision*, *Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, Nevada, USA, June 27–30, 2016), arXiv: [1512.00567](#) [[cs.CV](#)]
- [421] I. Loshchilov and F. Hutter, *Decoupled Weight Decay Regularization*, *Proceedings of the 7th International Conference on Learning Representations* (New Orleans, Louisiana, USA, May 6–9, 2019), arXiv: [1711.05101](#) [[cs.LG](#)]
- [422] P. Goyal et al., *Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour*, 2018, arXiv: [1706.02677](#) [[cs.CV](#)]
- [423] L. N. Smith and N. Topin, *Super-Convergence: Very Fast Training of Neural Networks Using Large Learning Rates*, 2018, arXiv: [1708.07120](#) [[cs.LG](#)]
- [424] M. J. Fenton et al., *Reconstruction of unstable heavy particles using deep symmetry-preserving attention networks*, [Commun. Phys. 7 \(2024\) 139](#)
- [425] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV*, [Eur. Phys. J. C 79 \(2019\) 639](#)
- [426] ATLAS Collaboration, *Studies of the muon momentum calibration and performance of the ATLAS detector with pp collisions at $\sqrt{s} = 13$ TeV*, [Eur. Phys. J. C 83 \(2023\) 686](#)
- [427] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [Eur. Phys. J. C 81 \(2021\) 689](#)
- [428] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, [Eur. Phys. J. C 79 \(2019\) 970](#)
- [429] ATLAS Collaboration, *Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector*, [Eur. Phys. J. C 82 \(2022\) 95](#)
- [430] ATLAS Collaboration, *Calibration of the light-flavour jet mistagging efficiency of the b-tagging algorithms with Z + jets events using 139 fb^{-1} of ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*, [Eur. Phys. J. C 83 \(2023\) 728](#)
- [431] ATLAS Collaboration, *Evaluation of QCD uncertainties for Higgs boson production through gluon fusion and in association with two top quarks for simplified template cross-section measurements*, [ATL-PHYS-PUB-2023-031](#) (CERN, Geneva, 2023)

Appendices

Additional Mva Material

A.1. Network Training

A.1.1. Training Curves

In Figures A.1 to A.4, the training curves of the final Transformer networks used in the $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis are shown. To conserve space, no subfigure captions are included. In each plot, training dataset curves are plotted in black solid lines and validation dataset curves are plotted in blue dashed lines. Loss curves are plotted on the left, while accuracy curves are plotted on the right. The different network folds are plotted with increasing index starting with fold zero on the top row of each figure. The point of early stopping after ten training epochs without any decrease in the validation loss is shown as two intersecting red dashed lines in each plot.

The validation dataset loss curves for the reconstruction Transformer networks are translated along the y-axis so that the barycentres of training and validation loss curves agree on the y-axis. The respective translated y-axis is displayed to the right of the loss plots.

Appendix A. Additional MVA Material

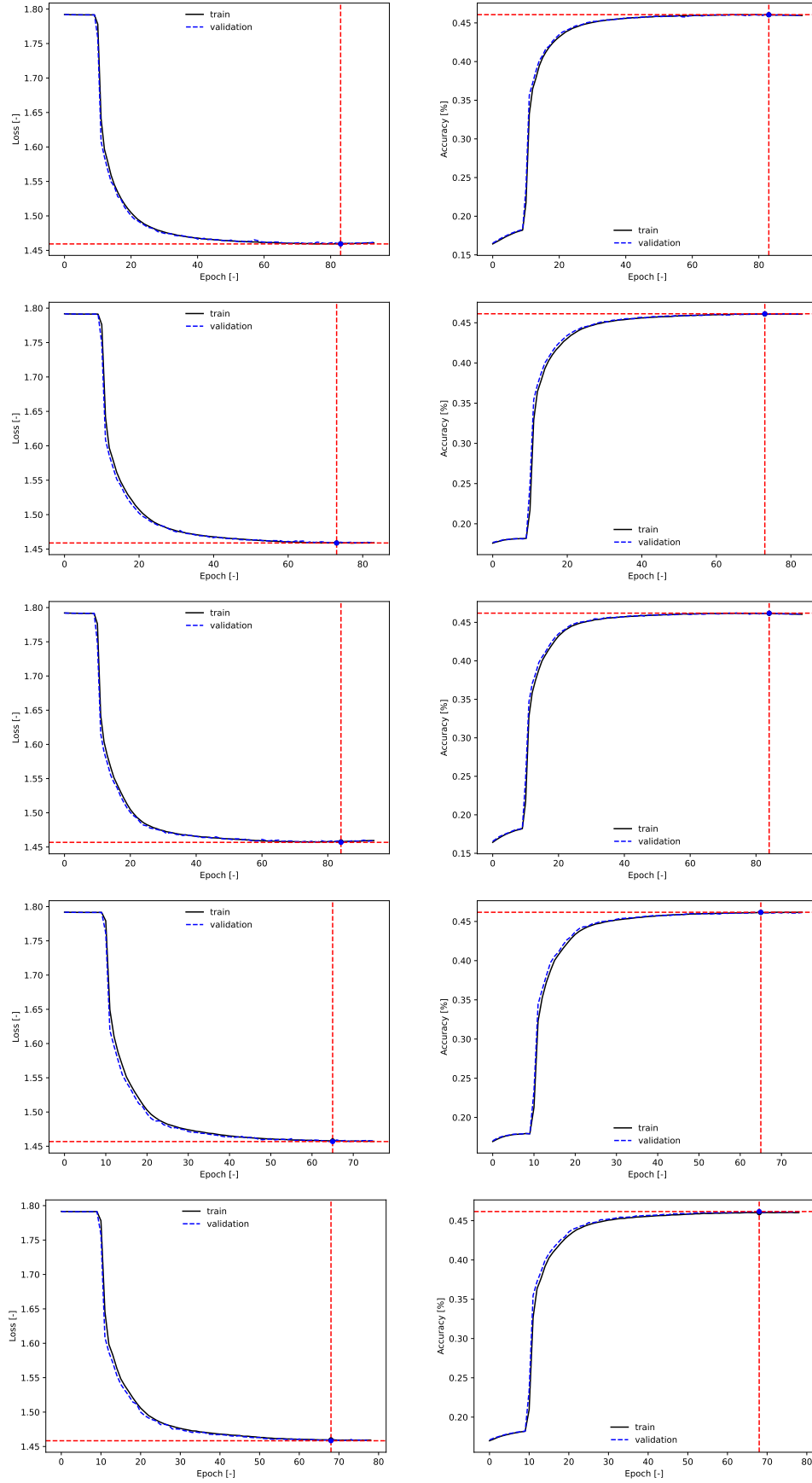


Figure A.1: Single Lepton Classification Network.

A.1. Network Training

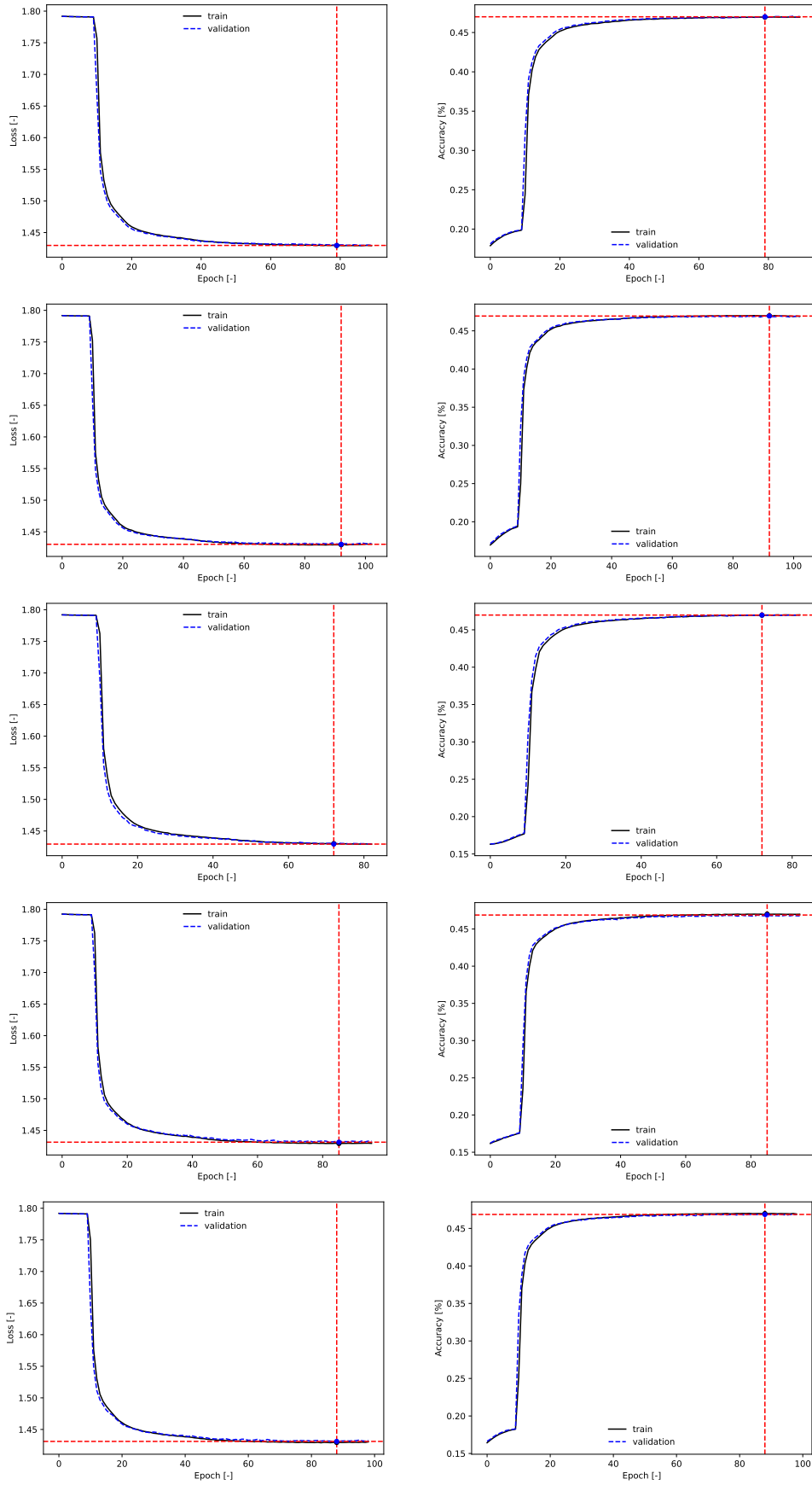


Figure A.2: Dilepton Classification Network

Appendix A. Additional MVA Material

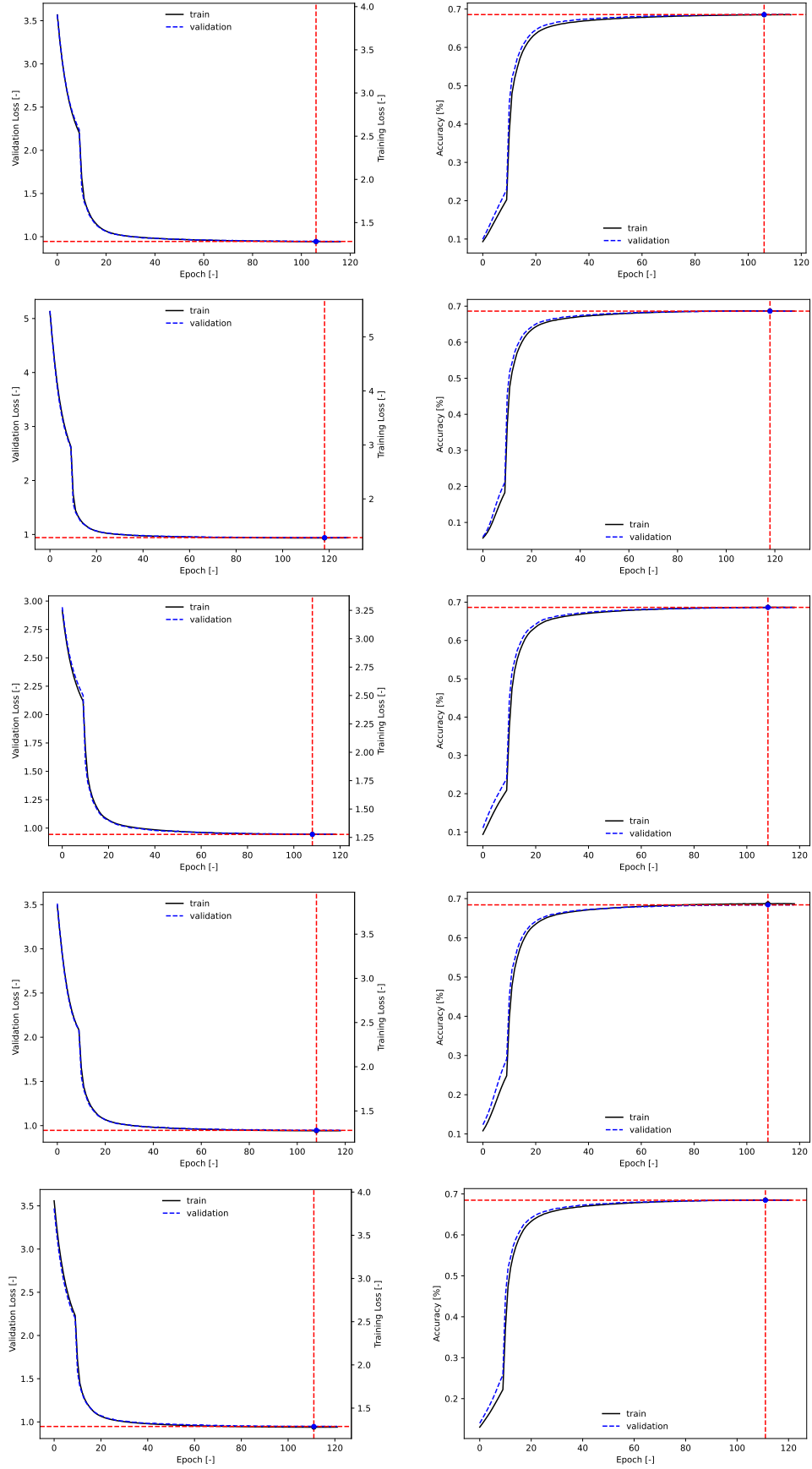


Figure A.3: Single Lepton Reconstruction Network

A.1. Network Training

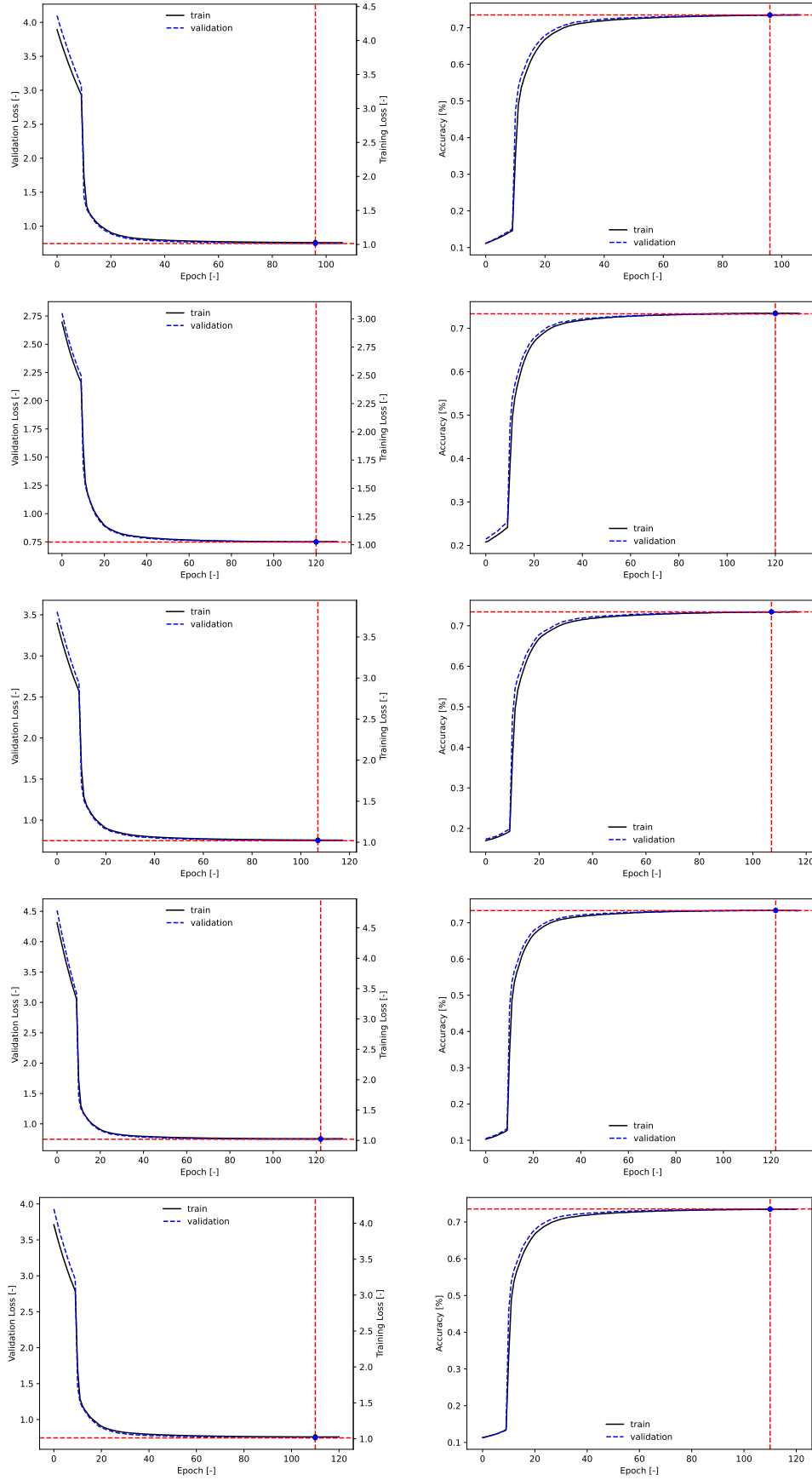


Figure A.4: Dilepton Reconstruction Network

A.1.2. Mc Statistics

The samples used for training the Transformer models are listed in Tables A.1 and A.2 with the corresponding MC statistics.

Table A.1: List of MC samples used during the classification Transformer training with associated combined MC statistics of the training samples. Several alternative samples are included in the training in addition to the nominal POWHEGBOX + PYTHIA8 samples to enhance training statistics.

Process	Combined MC Statistics		Samples
	Single Lepton	Dilepton	
$t\bar{t}H$	4 438 478	2 511 844	POWHEGBOX + PYTHIA8 POWHEGBOX + HERWIG7
$t\bar{t} + \geq 2b$	2 761 118	996 653	POWHEGBOX + PYTHIA8
$t\bar{t} + 1b$	2 390 870	1 332 277	POWHEGBOX + HERWIG7
$t\bar{t} + 1B$	837 097	429 334	POWHEGBOX + PYTHIA8, p_T^{hard} variation
$t\bar{t} + \geq 1c$	696 998	908 461	
$t\bar{t} + \text{light}$	1 431 191	1 121 241	

Table A.2: MC statistics for the $t\bar{t}H$ training samples of the reco Transformers. Stats are listed both split by truth STXS bin and as a combined total. The same samples as listed in Table A.1 for the classification models are used with reduced MC statistics due to the requirement of two distinct jets truth-matched to the Higgs boson decay products.

Truth p_T^H [GeV]			Combined MC Statistics	
			Single Lepton	Dilepton
0	–	60	659 432	340 025
60	–	120	1 031 029	518 148
120	–	200	919 209	441 637
200	–	300	484 514	225 734
300	–	450	206 597	94 472
450	–	∞	51 627	23 394
Total			3 352 408	1 643 410

A.1.3. Training Hyperparameters

The optimal training hyperparameters for the Mk3 retraining campaign are listed in Table A.3. The parameters were determined in a series of grid search optimisations informed by the converged validation losses and accuracies. Care was taken that the associated training curves, similar to those displayed in Section A.1.1, did not exhibit problematic behaviour such as overtraining or mode collapse.

Table A.3: List of optimal training hyperparameters used in the Mk3 networks.

Parameter	Classification	Reconstruction
Max. number of jet objects	11 (single lepton) 10 (dilepton)	
Dropout Rate		0.04
Pooling query regulariser	L1, 1×10^{-6}	n/a
Pairing layer regulariser	n/a	L1, 1×10^{-6}
Training batch size		2048
Validation batch size		2048
AdamW weight decay	5×10^{-6}	1×10^{-6}
AdamW first moment decay rate		0.9
AdamW second moment decay rate		0.999
Label smoothing		0.1
Max. number of training epochs		300
Early stopping min. loss decrease		0
Early stopping patience		10
Initial learning rate		1×10^{-8}
Max. learning rate		3.5×10^{-5}
Final learning rate		1×10^{-7}
Learning rate burn-in duration		10
Learning rate ramp-up duration		10
Learning rate plateau duration		15
Learning rate ramp-down duration		100

A.2. Transformer Performance

A.2.1. Mk3 Fraction Performance

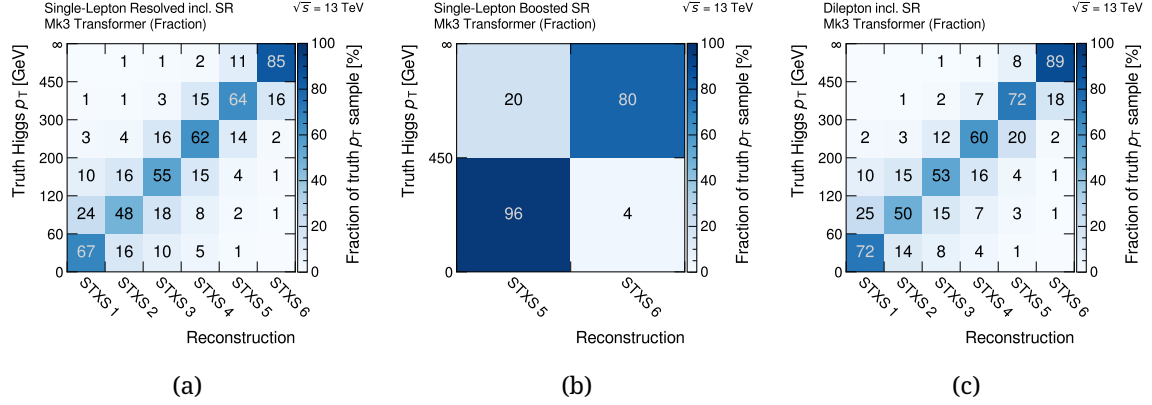


Figure A.5: STXS confusion matrices for the Mk3 pairing layer p_T^H reconstruction in the ttH SRS defined via the fraction classifiers of Equation (7.1). Overall, no significant difference to the STXS confusion matrices in Figure 7.6 resulting from SRS defined via the discriminant classifiers of Equation (7.2) is visible.

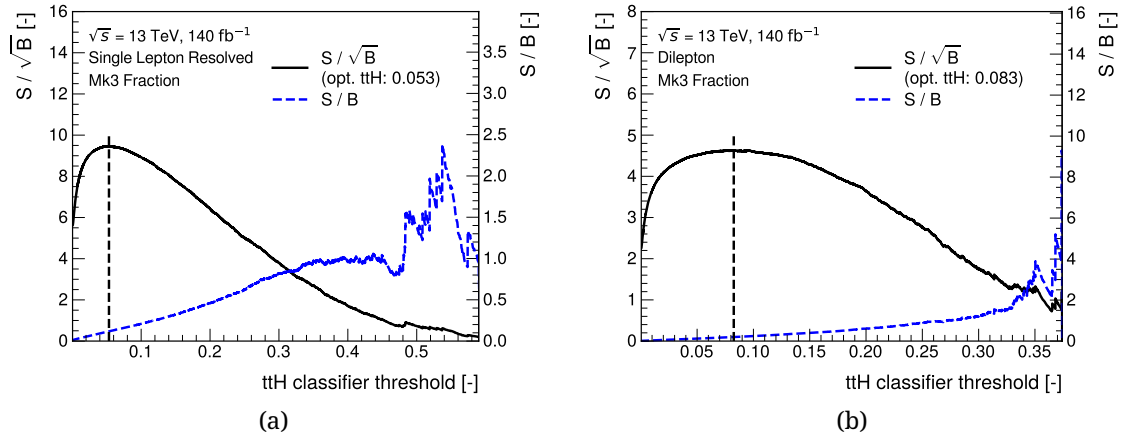


Figure A.6: Dependence of the signal-over-background noise ratio $S/\sqrt{B}$ and of the signal-over-background ratio S/B on the Mk3 ttH fraction classifier threshold in the (a) single-lepton resolved and (b) dilepton channel. Optimal thresholds were determined by finding the maxima of $S/\sqrt{B}$ in each case, while the distribution of S/B does not help to define usable thresholds. The optimal thresholds are plotted as dashed vertical lines in each plot.

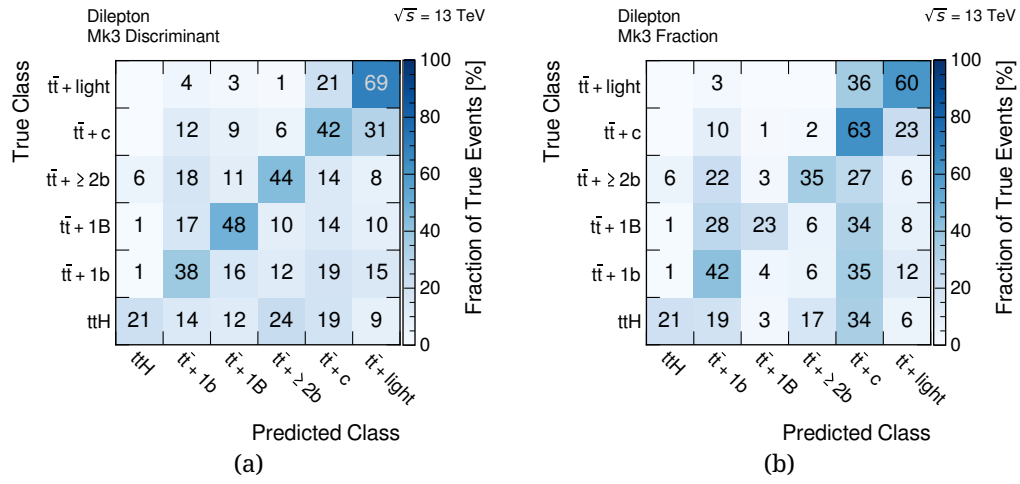


Figure A.7: Mk3 classification confusion matrices at the optimised $t\bar{t}H$ (a) discriminant and (b) fraction classifier thresholds in the dilepton channel. The matrices are normalised per row.

A.3. Additional Layer Schematics

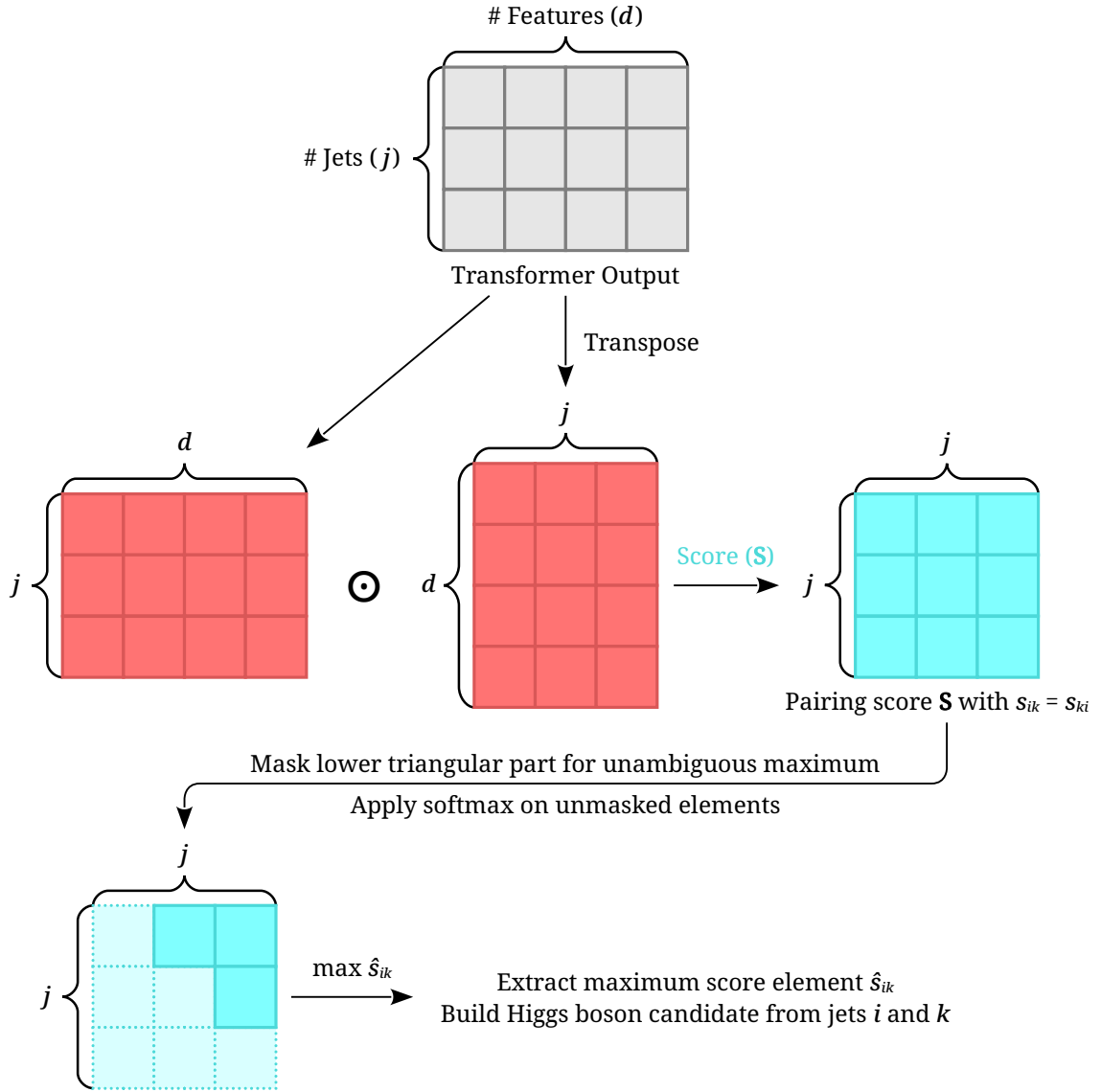


Figure A.8: Schematic of the information flow in a direct pairing layer. The computation follows Equations (5.12) and (5.13).

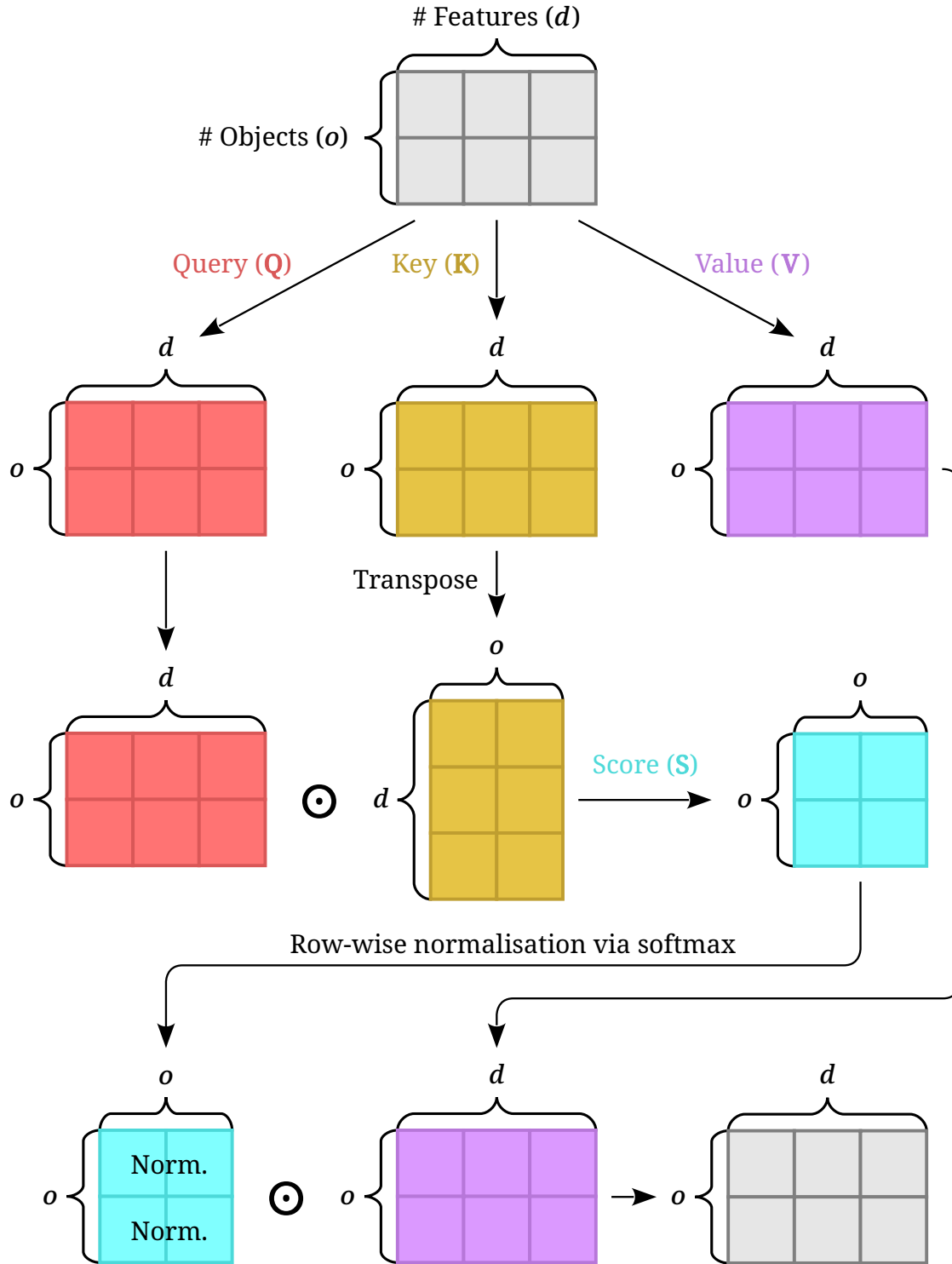


Figure A.9: Schematic of the information flow in a self-attention layer. The query, key, and value tensors are obtained from the $o \times d$ dimensional input feature tensor via three feed-forward layers. The computation involving the three tensors follows Equation (5.11).

Terms & Acronyms

2HDM

Two Higgs Doublet Model. [14](#), [15](#)

4FS

Four-Flavour Scheme. [23](#), [24](#), [26](#), [27](#), [78](#), [82–86](#)

5FS

Five-Flavour Scheme. [23](#), [24](#), [26](#), [78](#), [82–84](#), [86](#)

AFII

Fast simulation of the ATLAS detector response. [80](#), [84](#), [85](#), [90](#)

ALICE

A Large Ion Collider Experiment, one of the experiments at the LHC. [29](#)

AOD

Analysis Object Data. [81](#)

ASIC

Application-Specific Integrated Circuit. [42](#), [47](#), [48](#), [53](#)

ATLAS

One of the experiments at the LHC. [13](#), [17–19](#), [22](#), [24–27](#), [29](#), [32–37](#), [39–43](#), [47](#), [49](#), [50](#), [52–54](#), [58](#), [61](#), [77](#), [80–82](#), [84](#), [85](#), [87](#), [89](#), [94](#), [99](#), [117](#), [135–138](#), [179](#), [181](#), [185](#), [186](#), [188](#), [189](#)

AUC

Area Under the ROC Curve. [110](#), [114](#)

BCF

Boc Control FPGA. [49](#)

BCM

Beam Conditions Monitor. [43](#)

BDT

Boosted Decision Tree. [65](#), [77](#), [114](#), [115](#), [185](#)

BMF

Boc Main FPGA. [49](#)

Boc

Back-of-Crate Card. [49](#), [50](#), [53](#), [57–61](#), [180](#)

BR

Branching Ratio. [17](#), [20](#), [21](#), [25](#), [98](#), [119](#), [127](#)

BSM

Physics Beyond the Standard Model. [14](#), [23](#), [24](#), [40](#)

B-Layer

Innermost layer of the original PIXEL detector. [37](#), [53](#), [56](#)

CDF

Collider Detector at Fermilab, one of the experiments at the TEVATRON. [15](#)

CERN

Conseil Européen pour la Recherche Nucléaire
(European Organisation for Nuclear Research). [13](#), [29–31](#), [34](#), [49](#), [50](#), [58](#), [181](#), [188](#)

CI

Continuous Integration. [58](#), [60–62](#)

CKM

Cabibbo-Kobayashi-Maskawa (Matrix). [10](#), [12](#), [14](#)

CMB

Cosmic Microwave Background. [14](#)

CMS

Compact Muon Solenoid, one of the experiments at the LHC. [13](#), [18](#), [19](#), [22](#), [24](#), [25](#), [27](#), [29](#), [32](#), [135](#), [137](#)

CR

Control Region. [78](#), [91](#), [93](#), [94](#), [99](#), [100](#), [114](#), [120](#), [121](#), [123](#), [124](#), [128](#), [131](#), [132](#), [134](#)

CSC

Cathode Strip Chamber. [41](#)

CTP

Central Trigger Processor. [42](#), [53](#)

CVMFS

CERNVM File System. [58](#), [61](#)

CoM

Centre-of-Mass (Energy). [7](#), [15](#), [16](#), [30](#), [34](#), [35](#)

DAC

Digital-to-Analogue Converter. [52](#)

DAOD

Derived Analysis Object Data. [81](#), [188](#)

DAQ

Data Acquisition. [47](#), [58–62](#)

DGLAP

Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (Evolution Equations). [8](#)

DIS

Deep Inelastic Scattering. [6](#), [8](#)

DL1r

RNN-based ATLAS Run 2 *b*-tagging algorithm. [77](#), [88](#), [90](#), [91](#), [94](#), [96](#), [98](#), [99](#), [102](#), [112](#), [115](#), [118](#)

DNN

Deep Neural Network. [88](#), [90](#)

DR

Diagram Removal (Scheme). [17](#), [23](#), [83–85](#)

DS

Diagram Subtraction (Scheme). [17](#), [84](#)

DØ

One of the experiments at the TEVATRON. [15](#)

ECAL

Electromagnetic Calorimeter. [39](#), [40](#), [42](#), [88–90](#)

ECR

Event Counter Reset. [53](#), [55](#)

EFT

Effective Field Theory. [9](#)

EWSB

Electroweak Symmetry Breaking. [11](#), [12](#), [23](#)

EW

Electroweak. [11–13](#), [18](#), [23](#), [85](#)

FCNC

Flavour-Changing Neutral Current. [13](#), [15](#)

FDAC

Pixel-level register for fine control of the ToT response. [52](#), [60](#)

FE-I4B

FE read-out chip used for IBL. [47](#), [54](#), [59](#)

FE-I3

FE read-out chip used for the PIXEL detector. [47](#), [54](#), [59](#)

FE

Front-End. [47–52](#), [54](#), [55](#), [60](#), [61](#), [182–184](#)

FPGA

Field-Programmable Gate Array. [42](#), [43](#), [49](#), [50](#), [57](#), [58](#), [180](#)

FSR

Final State Radiation. [82](#), [83](#), [119](#), [127](#)

FTK

Fast Tracker. [43](#), [49](#)

FW

Firmware. [49](#), [50](#), [53](#), [54](#), [57](#), [58](#), [60–62](#)

GDAC

FE-level register for control of the discriminator threshold. [52](#), [60](#)

ggF

Gluon–Gluon Fusion. [15](#), [19](#), [20](#), [22](#), [23](#), [26](#), [41](#)

GIM

Glashow-Iliopoulos-Maiani (Mechanism). [13](#), [15](#)

GRL

Good Runs List. [82](#)

GUI

Graphical User Interface. [58](#), [59](#), [184](#)

GUT

Grand Unified Theory. [14](#)

HCAL

Hadronic Calorimeter. [40](#), [42](#)

HERA

Hadron-Elektron-Ringanlage
(Hadron-Electron Ring Accelerator). [8](#)

HLT

High-Level Trigger. [42](#), [43](#), [49](#), [87](#)

HW

Hardware. [49](#), [50](#), [57](#)

IBL

Insertable *B*-Layer. [36–38](#), [47–49](#), [54–56](#), [182](#), [186](#)

ID

Inner Detector. [34–39](#), [43](#), [88–90](#)

IFDAC

FE-level register for control of the ToT response. [52](#)

IGUI

Integrated GUI. [58](#), [61](#)

ISR

Initial State Radiation. [82](#), [83](#), [119](#), [127](#)

LB

Luminosity Block. [43](#)

LEIR

Low Energy Ion Ring. [31](#)

LEP

Large Electron-Positron Collider. [30](#)

LHCb

Large Hadron Collider beauty, one of the experiments at the LHC. [29](#)

LHC

Large Hadron Collider. [8](#), [15–17](#), [19](#), [21](#), [22](#), [26](#), [29–35](#), [37](#), [41](#), [43](#), [49](#), [53](#), [55](#), [56](#), [81](#), [82](#), [117](#), [138](#), [179](#), [181](#), [184](#), [189](#)

Lo

Leading Order. [17](#), [19](#), [80–82](#), [84](#), [85](#)

LUCID

Luminosity Cherenkov Integrating Detector. [43](#)

LINAC

Linear Accelerator. [31](#)

L1

Level-1 (Trigger). [41](#), [42](#), [48](#), [49](#), [52–54](#), [56](#), [61](#), [87](#)

Mcc

Module Controller Chip. [48](#), [53](#), [55](#)

Mc

Monte-Carlo. [8](#), [15](#), [17](#), [23](#), [26](#), [27](#), [67](#), [77](#), [79–82](#), [84–87](#), [92](#), [93](#), [96](#), [97](#), [106–109](#), [117–119](#), [122](#), [126](#), [128](#), [130–134](#), [172](#)

MDT

Monitored Drift Tube. [41](#)

ME

Matrix Element. [79](#), [82](#), [84](#), [85](#)

MLP

Multilayer Perceptron. [65–67](#), [71](#), [73](#), [105](#)

MSE

Mean-Squared Error. [67](#), [70](#), [106](#)

MSSM

Minimally Supersymmetric Standard Model. [15](#)

MS

Muon Spectrometer. [34](#), [40](#), [42](#), [88](#)

MVA

Multivariate Analysis (Technique). [26](#), [77](#), [78](#), [101](#), [138](#)

Mv2c10

BDT-based ATLAS *b*-tagging algorithm. [77](#), [115](#)

Micromega

Micro Mesh Gaseous Structure, a type of wireless gaseous drift chamber able to cope with environments of high particle fluxes. [41](#)

NLO

Next-to-Leading Order. [16–18](#), [23](#), [78](#), [82–85](#), [119](#)

NNLL

Next-to-Next-to-Leading-Log Order. [16](#), [85](#)

NNLO

Next-to-Next-to-Leading Order. [16](#), [18](#), [83–85](#)

N³LO

Next-to-Next-to-Next-to-Leading Order. [19](#)

NP

Nuisance Parameter. [92](#), [93](#), [118](#), [119](#), [125](#), [126](#)

NSW

New Small Wheel. [35](#), [41](#)

ONNX

Open Neural Network Exchange. [107](#)

PCA

Principal Component Analysis. [118](#)

PCBT

Pseudo-Continuous *b*-Tagging. [88](#), [102](#), [108](#)

PDF

Parton Distribution Function. [8](#), [16](#), [17](#), [23](#), [80–85](#), [119](#)

PIXEL

Original three-layered ATLAS pixel detector, complemented by IBL in Long Shut-down 1. [37](#), [38](#), [47–49](#), [53](#), [54](#), [180](#), [182](#)

PMNS

Pontecorvo-Maki-Nakagawa-Sakata (Matrix). [14](#)

PMT

Photomultiplier Tube. [40](#), [43](#)

POI

Parameter of Interest. [92](#), [93](#), [124](#), [126](#), [128](#), [129](#)

PPC

PowerPC. [50](#), [60](#)

PRM

Program Reset Manager. [50](#)

PSB

Proton Synchrotron Booster. [31](#)

PS

Proton Synchrotron. [31](#)

Ps

Parton Shower. [78–80](#), [82–85](#), [94](#), [119](#)

QCD

Quantum Chromodynamics. [4–9](#), [16–20](#), [23](#), [25](#), [26](#), [79](#), [82–85](#), [90](#), [119](#), [127](#)

QED

Quantum Electrodynamics. [3](#), [9](#)

QFT

Quantum Field Theory. [5](#), [6](#)

RNN

Recurrent Neural Network. [65](#), [77](#), [88](#), [89](#), [181](#)

ROB

Read-Out Buffer. [49](#)

ROC

Receiver-Operating Characteristic. [110](#), [111](#), [114](#), [115](#), [180](#)

ROD

Read-Out Driver Card. [49–51](#), [53](#), [54](#), [57](#), [58](#), [60–62](#)

ROS

Read-Out System. [42](#), [47](#), [49](#)

RPC

Resistive Plate Chamber. [41](#)

SCT

Semiconductor Tracker. [36](#), [38](#), [49](#)

SET

Single Event Transient. [54](#)

SEU

Single Event Upset. [54](#)

SGD

Stochastic Gradient Descent. [68](#)

SM

Standard Model (of Particle Physics). [4–6](#), [10](#), [13–15](#), [17](#), [19–25](#), [79](#), [90](#), [99](#), [119](#), [125](#), [128](#), [135](#)

SPS

Super Proton Synchrotron. [31](#)

SR1

Facility at CERN housing the ATLAS Pixel detector test system. [50](#), [58](#)

SR

Signal Region. [78](#), [91](#), [94](#), [99](#), [100](#), [106](#), [111](#), [112](#), [114–116](#), [120–124](#), [128](#), [130](#), [132](#), [133](#), [174](#)

STXS

Simplified Template Cross-Section. [26](#), [73](#), [78](#), [99](#), [100](#), [106–108](#), [110–112](#), [115](#), [117](#), [119](#), [124](#), [127–129](#), [135](#), [137](#), [172](#), [174](#)

SUSY

Supersymmetry. [14](#)

SW

Software. [49](#), [50](#), [53](#), [57–62](#)

SppS

Super Proton–Anti-proton Synchrotron. [13](#)

TDAC

Pixel-level register for fine control of the discriminator threshold. [52](#), [60](#)

TDAQ

Trigger and Data Acquisition. [58](#), [61](#)

TEVATRON

Circular proton–anti-proton accelerator situated at Fermilab. [8](#), [15](#), [16](#), [180](#), [182](#)

TGC

Thin Gap Chamber. [41](#)

TOPQ1

DAOD format used by the ATLAS $t\bar{t}(H \rightarrow b\bar{b})$ Run 2 legacy analysis. [81](#), [86](#)

TRT

Transition Radiation Tracker. [36](#), [38](#)

ToT

Time-over-Threshold. [48](#), [51–53](#), [56](#), [59](#), [60](#), [182](#), [184](#)

USA15

Service cavern of the ATLAS detector containing the counting room. [49](#)

VBF

Vector Boson Fusion. [19](#), [20](#), [26](#)

WIMP

Weakly Interacting Massive Particle. [14](#)

WLCG

Worldwide LHC Computing Grid. [81](#)

WP

Working Point. [88–91](#), [94](#), [96](#), [99](#)

YAML

YAML Ain't Markup Language. [58](#), [189](#)

YETS

Year-End Technical Stop. [30](#), [33](#), [49](#)

Declaration on the Use of ChatGPT and Comparable Tools in the Context of Examinations

In this thesis, I have used ChatGPT or comparable AI tools as follows:

- ☐ not at all
- ☐ during brainstorming
- ☐ when creating the outline
- ☐ to write individual passages
- ☐ for the development of software source texts
- ☒ for optimising or restructuring software source texts
- ☒ for proofreading or optimising
- ☐ further, namely: ...

I hereby declare that I have stated all uses completely. Missing or incorrect information will be considered as an attempt to cheat.