University of Glasgow | Experimental Particle Physics

Measurement of the top-quark mass with the ATLAS detector using $t\bar{t}$ events with a boosted top quark

Elliot Watton

School of Physics and Astronomy
College of Science and Engineering
University of Glasgow

July 21, 2025

A dissertation submitted to the University of Glasgow
for the degree of Doctor of Philosophy

Abstract

As the heaviest fundamental particle observed to date, the top quark plays a crucial role in particle physics. Precise determination of its mass is essential for testing the internal consistency of the Standard Model and probing potential new physics. This thesis presents a measurement of the top-quark mass using the full Run 2 dataset, corresponding to an integrated luminosity of 140 fb^{-1}, collected during proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV by the ATLAS experiment at the Large Hadron Collider between 2015 and 2018. The analysis focuses on top-antitop-quark pair decays that produce a highly energetic top quark, reconstructed using a single large-radius jet. The average invariant mass of this jet is used to extract the top-quark mass using a profile likelihood fit, incorporating two additional observables to constrain and reduce systematic uncertainties. This approach yields the most precise top-quark mass measurement by ATLAS in a single channel to date: 172.95 ± 0.53 GeV. In addition, this thesis presents the development and integration of the Fast Track Finder algorithm into the ATLAS trigger system. This represents the first successful implementation of a CPU-based charged-particle track reconstruction chain within the ATLAS trigger that is intended for use during the operation of the High-Luminosity Large Hadron Collider, which is scheduled to commence in 2030.

Declaration

I declare that the research presented in this thesis is the result of my own work within the Particle Physics Experiment Group in the School of Physics and Astronomy at the University of Glasgow, as well as the Particle Physics Department at the Rutherford Appleton Laboratory. The research contained within this thesis has not been submitted for another degree or qualification at the University of Glasgow or any other institution.

My research was performed as part of the ATLAS Collaboration at the Large Hadron Collider (LHC), meaning that I have benefited from the expertise, tools, and hard work of thousands of individuals over many years. During my research I have contributed to the collaboration by undertaking ATLAS control room shifts, as well as data quality expert shifts. Additionally, I have been a leading developer for a charged particle track reconstruction pipeline using CPU hardware, and have taken a year-long position as the statistics liaison for the Top working group.

This thesis is split into three parts. Part I (Chapters 1-5) covers an introduction to the work presented in this thesis, as well as discussing the theoretical background, the ATLAS detector, and the relevant experimental techniques where I did not contribute any work.

Part II (Chapters 6-7) covers the details of the Phase-II upgrade planned for the ATLAS detector ahead of the High-Luminosity LHC (HL-LHC) era, as discussed in Chapter 6. Chapter 7 then describes my contributions to the development of a CPU-based charged particle tracking pipeline intended for use by ATLAS following this upgrade. The work presented in Chapter 7 is my own, except for the following:

- The implementation of the Fast Track Finder algorithm before Run 4 (Section 7.2.1).

- Original design and development of the track seeding based on a graph neural network (Section 7.2.2).

Part III (Chapters 8-11) presents the measurement of the top-quark mass using events with a high transverse momentum top quark, for which I was the lead analyst in a team of five people. This work is published by ATLAS in Ref. [1]. Part III also includes the conclusion to the thesis. The work presented in Part III is my own and places where other analysis team members made significant contributions are:

- The method for estimating the contribution of multijet events to the signal background (Section 8.1).
- Initial development of the profile likelihood fit (Section 9.3).
- Testing whether statistical correlations between observables used in the profile likelihood fit were significant (Section 9.5.1).
- Test of fit linearity and non-closure (Section 9.5.2).
- Expected precision of the top-quark mass result after the operation of the HL-LHC (Section 10.4). This extrapolation is published by ATLAS in Ref. [2].

Elliot Watton

Acknowledgements

The past few years as a PhD student have been some of the most rewarding and memorable of my life. This journey would not have been possible without the support, guidance, and friendship of many people. While I cannot possibly name everyone who has played a role, I hope those not mentioned here know how much I appreciate them.

First and foremost, I would like to express my deepest gratitude to my supervisors, Mark Owen and Stewart Martin-Haugh. Being a joint Glasgow-RAL student has provided many advantages, the most significant of which has been the invaluable support and mentorship of two exceptional supervisors. Their experience, knowledge, insight, and willingness to engage with my many questions have played an important role in shaping me into a more well-rounded physicist.

I am also sincerely grateful to John Baines for his unwavering support and expertise in the tracking software for the ATLAS Phase-II upgrade. His guidance, encouragement, and deep knowledge of the trigger system were instrumental to my work at RAL, and I truly appreciate the time and effort he invested in helping me navigate this project.

I would also like to express my gratitude to Federica Fabbri for her allowing me to pester her time and again about top-quark modelling and the intricacies of the GlasgowAna software. Alongside this, I have benefitted from her guidance and encouragement throughout my PhD and as I embark on the next stage of my academic journey.

My thanks also go to Jonathan Jamieson, whose experience working on a closely related analysis made him an invaluable source of knowledge. Whether over a coffee in the Kelvin Building or a beer on a beach in sunny Spain, his advice and support were always appreciated.

Living in Glasgow, Oxford, and Geneva throughout my PhD has allowed me to form an incredible network of friends, both within and beyond the world of physics. While I cannot list you all by name, I hope you know how much your friendship means to me. From floating down the Rhône to sipping wine at Caves Ouvertes, from conference-adjacent holidays to countless pub trips, you have made this experience truly unforgettable.

Finally, I want to express my deepest appreciation to my family for their love and support. Being farther from home than ever before has been challenging, but their encouragement has been a constant source of strength.

Contents

List of acronyms

BSM Beyond the Standard Model. 2, 8–10, 31, 70, 72

CPV Charge Parity Violation. 8, 9, 15, 41

CR Colour Reconnection. 41, 42, 109–111, 119, 142, 155, 169

CSC Cathode Strip Chambers. 28, 30, 47, 51

CTP Central Trigger Processor. 78, 80

ECal Electromagnetic Calorimeter. 26, 27

EF Event Filter. 73, 78, 81, 83–85, 92, 98–100, 178

EM Electromagnetic. 5–7, 20, 26, 28, 48, 60, 79, 113

EW Electroweak. 5, 6, 9–11, 37, 72, 73, 111, 177

FCal Forward Calorimeter. 28

FELIX Front-End Link Exchange. 80

FEX Feature Extractors. 79

FPGA Field Programmable Gate Arrays. 78, 79, 81, 99

FSR Final-State Radiation. 53, 110, 119, 120, 145, 148, 149, 156

FTF Fast Track Finder. 84–88, 91–101, 178

GNN Graph Neural Network. 84, 88, 92, 96, 98, 100, 178

GSC Global Sequential Calibration. 60, 61, 64

HCal Hadronic Calorimeter. 27

HL-LHC High-Luminosity Large Hadron Collider. 3, 72, 73, 76–78, 81, 82, 84–86, 88, 99, 105, 151, 174, 176–178

HLT High-Level Trigger. 31, 32, 73, 81, 83, 86, 178

IBL Insertable B-Layer. 23, 76

ID Inner Detector. 20, 22, 23, 25, 26, 43, 44, 47–52, 54, 70, 72–74, 77, 78, 86, 108, 114, 178

IP Interaction Point. 17, 20, 22, 26

IR Infrared. 37, 39

IRC Infrared and Collinear. 57

IS Inner System. 75

ISR Initial-State Radiation. 110, 119, 145

ITk Inner Tracker. 73, 74, 76–78, 81–86, 88, 90, 92, 93, 97–100, 178

JER Jet Energy Resolution. 62, 63, 67, 117, 118, 137

JES Jet Energy Scale. 58, 62–64, 66, 117, 118, 122, 123, 134, 135, 143, 145, 148, 156, 169, 175

JVT Jet Vertex Tagger. 57, 118

L0 Level-0. 78, 80, 85, 99

L1 Level-1. 31, 32, 78, 86, 87

LEP Large Electron-Positron collider. 17, 65

LHC Large Hadron Collider. 2, 3, 9, 10, 15–20, 23, 30, 31, 33, 34, 37, 38, 42, 53, 70, 72, 151, 173, 176, 177

LO Leading Order. 10, 37, 103, 105, 111, 113

MC Monte Carlo. 3, 13, 34, 38–43, 48, 58, 62, 65–67, 86, 88, 92, 93, 98, 103, 104, 107–109, 112, 113, 118, 121, 123, 125, 136, 137, 142, 153, 178

MDT Monitored Drift Tubes. 28, 30, 46, 51

ME Matrix Element. 36, 38–40, 109–111, 113, 118

MJB Multijet Balance. 61, 63

MPF Missing-E_{T} Projection Fraction. 61, 63

MPI Multiple Partonic Interactions. 41, 42, 110, 111

MS Muon Spectrometer. 20, 22, 28, 31, 43, 44, 46–48, 51–53, 80, 85

MUCTPI Muon-to-CTP-Interface. 78, 80

NLO Next-To-Leading Order. 11, 13, 37–39, 111–113, 118, 125

NNLL Next-To-Next-To-Leading Logarithm. 37, 109, 110, 112

NNLO Next-To-Next-To-Leading Order. 37, 109–113, 120, 125, 161

NP Nuisance Parameter. 104, 108, 127–129, 131–137, 142, 143, 145, 148, 149, 151, 152, 155–157, 159, 161, 162, 169, 177

OB Outer Barrel. 75

OE Outer End-Cap. 75

PDF Parton Distribution Function. 36, 38, 108–113, 120, 121

PFlow Particle Flow. 44, 53, 54, 56, 57, 64, 69, 85, 117

POI Parameter of Interest. 127–129, 132

QCD Quantum Chromodynamics. 6, 10, 11, 36, 37, 40–42, 53, 106, 109, 111–113, 119, 123

QFT Quantum Field Theory. 5

RC Reclustered. 69, 104, 114, 117, 118, 177

RoI Regions of Interest. 31, 32, 79, 80, 85–87, 92

RPC Resistive Plate Chambers. 28, 30, 31, 46, 80

SCT Semiconductor Tracker. 22, 24, 25, 44, 45, 54, 86

SM Standard Model. 2, 3, 5–9, 11, 17, 28, 31, 34, 37, 53, 62, 67, 70, 72, 104, 127, 132, 141, 142, 177, 178

TDAQ Trigger and Data Acquisition. 31, 33, 73, 78, 79, 81–83, 99, 178

TGC Thin Gap Chambers. 28, 30, 31, 46, 80

TRT Transition Radiation Tracker. 22, 25, 28, 86, 87

UE Underlying Event. 41, 57, 61, 109, 110, 119

UV Ultraviolet. 37

PART I

The top quark and measurements with the ATLAS detector

CHAPTER 1

Introduction

The Standard Model (SM) of elementary particle physics stands as the most successful theory we have for describing the fundamental building blocks of the Universe. To test this framework, scientists have constructed the largest and one of the most complex machines ever built: the Large Hadron Collider (LHC) at CERN. This extraordinary instrument smashes particles together at nearly the speed of light, reaching unprecedented energies. Along the LHC ring are four major experiments, one of which is ATLAS. The ATLAS detector functions much like an incredibly sophisticated and expensive camera, capturing snapshots of these high-energy particle collisions. By analysing these interactions, physicists can probe the most fundamental forces of nature, rigorously test our theories, and search for physics Beyond the Standard Model (BSM).

Particle physics has fascinated me for many years, a passion that first began when I visited CERN in 2017. At the time, I was an 18 year old high school student preparing to study physics at university, entirely new to the world of particle physics. My understanding barely extended beyond knowing that quarks existed and that they were somehow important. I vividly remember standing outside the ATLAS control room, peering through the glass at the busy physicists inside. It looked exciting, fast paced, focused, and full of purpose. I thought to myself that one day I wanted to be on the other side of that glass, right in the middle of the action. Eight years later, I have not only stepped inside, I have taken shifts working in the ATLAS control room. There is something deeply rewarding about contributing to the smooth operation of the experiment and knowing that the data we help collect will lead to significant results and potentially new discoveries. Now, I am about to begin a two year position as

part of the ATLAS central trigger group at CERN, and I could not be more excited to continue contributing to the field of particle physics.

This thesis presents my contributions to the ATLAS collaboration over the past three and a half years, including a precision measurement of the heaviest fundamental particle in the SM: the top quark. Due to its large mass, the top quark plays a crucial role in the SM, closely linked to the Higgs mechanism, and a precise measurement of its mass provides an essential consistency test of the SM. The result presented here is of particular significance, as it represents the most precise measurement of the top-quark mass performed by the ATLAS collaboration to date in a single decay channel, and is published in Ref. [1].

The ATLAS detector is scheduled to undergo a major upgrade in preparation for the high-luminosity upgrade of the LHC, known as the High-Luminosity Large Hadron Collider (HL-LHC), which is expected to begin operations in 2030. In response to the increasingly challenging particle tracking environment associated with higher luminosity, the ATLAS detector will be equipped with a new inner tracker, which will replace the existing inner detector. The system responsible for determining which collision events are retained for analysis will also be upgraded. This thesis presents my contribution to this effort, focusing on the development and integration of a charged particle track reconstruction algorithm designed to operate within this upgraded system with the new detector geometry.

Given the dual focus of this thesis and the need to establish the theoretical and experimental context of the SM and the ATLAS detector at the LHC, the work is divided into three main parts outlined below.

Part I lays the theoretical and experimental foundations for conducting a particle physics measurement with the ATLAS detector. Chapter 2 introduces the SM and the top-quark's place within it. Chapter 3 provides an overview of the CERN accelerator complex, the LHC, and the ATLAS detector. Chapter 4 describes the use of Monte Carlo (MC) methods for simulating particle collisions. Finally, Chapter 5 explains how the detector's recorded signals are reconstructed into calibrated physics objects such as leptons and hadronic jets.

Part II focuses on the Phase-II upgrade of the ATLAS detector, designed to ensure ATLAS can fully exploit the physics potential of the HL-LHC. Chapter 6 discusses some specifics of this upgrade, such as a new all-silicon inner tracker detector, while

Chapter 7 details my contributions to the integration and development of a charged particle track reconstruction algorithm within the ATLAS software framework, Athena.

Part III presents the high-precision measurement of the top-quark mass. Chapter 8 outlines the analysis strategy, Chapter 9 describes the profile likelihood fitting technique used to determine the top-quark mass, and Chapter 10 presents the final results. The thesis concludes with a summary in Chapter 11.

CHAPTER 2

Theoretical Framework

To understand the significance of measuring the top-quark mass, it is essential to first discuss the underlying theory. This chapter introduces the SM, highlights its limitations, and explores how the top quark can be used to test its internal consistency. The production of top quarks at hadron colliders and their decay modes will then be examined. Finally, the chapter discusses the issue associated with relating the top-quark mass measured in experiments to the parameters used in Quantum Field Theory (QFT).

2.1 The Standard Model

The SM is currently the most comprehensive theory describing the matter and anti-matter in the Universe, as well as the fundamental forces acting between them. It is a QFT that unifies special relativity and quantum mechanics, where matter and anti-matter particles are described by fields. The SM consists of twelve spin-$\frac{1}{2}$ particles, known as *fermions*, each interacting through at least one of the three fundamental forces in the framework: the Electromagnetic (EM) interaction, the weak nuclear force, and the strong nuclear force. At sufficiently high energies, the EM and weak nuclear forces combine into the unified Electroweak (EW) interaction. Fermions interact via the exchange of integer-spin particles, called *bosons*, that mediate these forces. Gravity, however, is not part of the SM and is instead described by general relativity. It remains unknown whether a viable quantum theory of gravity exists that could be incorporated into the SM. Nevertheless, in high-energy particle physics, gravity is typically negligible due to its relative weakness compared to the other forces.

The SM is governed by its Lagrangian, which remains invariant under multiple local gauge transformations. These symmetries necessitate the inclusion of additional gauge fields, each corresponding to a fundamental force within the SM. The symmetry structure of the SM is expressed as $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_I \times \mathrm{U}(1)_Y$, where each group is associated with a conserved quantity. The $\mathrm{SU}(3)_C$ symmetry corresponds to the strong force, described by Quantum Chromodynamics (QCD), with the conserved quantity being the colour charge, denoted by C. The strength of this interaction is governed by the strong coupling constant, α_s, which decreases with increasing energy. The strong force is mediated by eight massless spin-1 bosons known as *gluons*. The $\mathrm{SU}(2)_I \times \mathrm{U}(1)_Y$ symmetry corresponds to the unified EW interaction, with weak isospin, I, and hypercharge, Y, as conserved quantities. The EW interaction is mediated by four spin-1 bosons: the massive $W^{\pm}$ and Z gauge bosons, which mediate the weak force, and the massless photon, γ, which mediates the EM interaction. The charged $W^{\pm}$ bosons are associated with the $\mathrm{SU}(2)_I$ term, while the neutral Z boson and photon are associated to the transformations involving the $\mathrm{SU}(2)_I \times \mathrm{U}(1)_Y$ symmetry. The scalar Higgs boson emerges from the process of *spontaneous symmetry breaking* [3, 4], which ensures that the SM Lagrangian remains gauge-invariant. The Higgs boson itself is an excitation of the scalar Higgs field, which is responsible for the masses of the SM fermions via *Yukawa couplings* [5]. For a fermion with mass m_f, it couples to the Higgs field with the Yukawa coupling, y_f, given by:

$$y_f = \sqrt{2}\frac{m_f}{v}, \tag{2.1}$$

where v is the *vacuum expectation value* of the Higgs field, and has a value of 246 GeV. Therefore, the strength of each coupling is proportional to the mass of the fermion, meaning that more massive fermions interact more strongly with the Higgs boson. However, the masses of fermions are free parameters within the SM; they cannot be derived theoretically and must instead be determined experimentally. Interactions with the Higgs field generates the masses of the W and Z bosons.

The twelve fundamental fermions in the SM are divided into two categories: *quarks* and *leptons*. The six quarks are organised into three generations of increasing mass, with each generation containing an up-type and a down-type quark. The up-type quarks are named the *up*, *charm*, and *top* quarks, each carrying a fractional electric charge of $+\frac{2}{3}e$. The corresponding down-type quarks are the *down*, *strange*, and *bottom* quarks, each with a fractional electric charge of $-\frac{1}{3}e$. All quarks carry both electric and colour charge, enabling them to participate in the EW and strong interactions,

respectively. Similarly, the six leptons are grouped into three generations of increasing mass, with each generation comprising a pair: one electrically charged lepton and one neutral lepton. The charged leptons are known as the *electron*, *muon*, and *tau*, while the corresponding neutral leptons are called *neutrinos*, with one neutrino associated with each charged lepton. Neutrinos interact exclusively via the weak force, whereas the electrically charged leptons interact through both the weak and EM forces. Unlike quarks, leptons do not carry a colour charge and therefore cannot interact via the strong force. Every fermion and boson in the SM has a corresponding antiparticle, characterised by the same mass, but opposite charge and quantum numbers. The gluon, photon, Z boson, and the Higgs boson are their own antiparticles, as they are invariant under charge conjugation, while the W^+ and W^- bosons are each other's antiparticles. Although neutrinos are both electrically and colour neutral, they are not their own antiparticles due to their distinguishable chirality.[1] Figure 2.1 provides a summary of the fundamental particles of the SM, including their masses, charges, and spin values, as reported by the Particle Data Group [6].

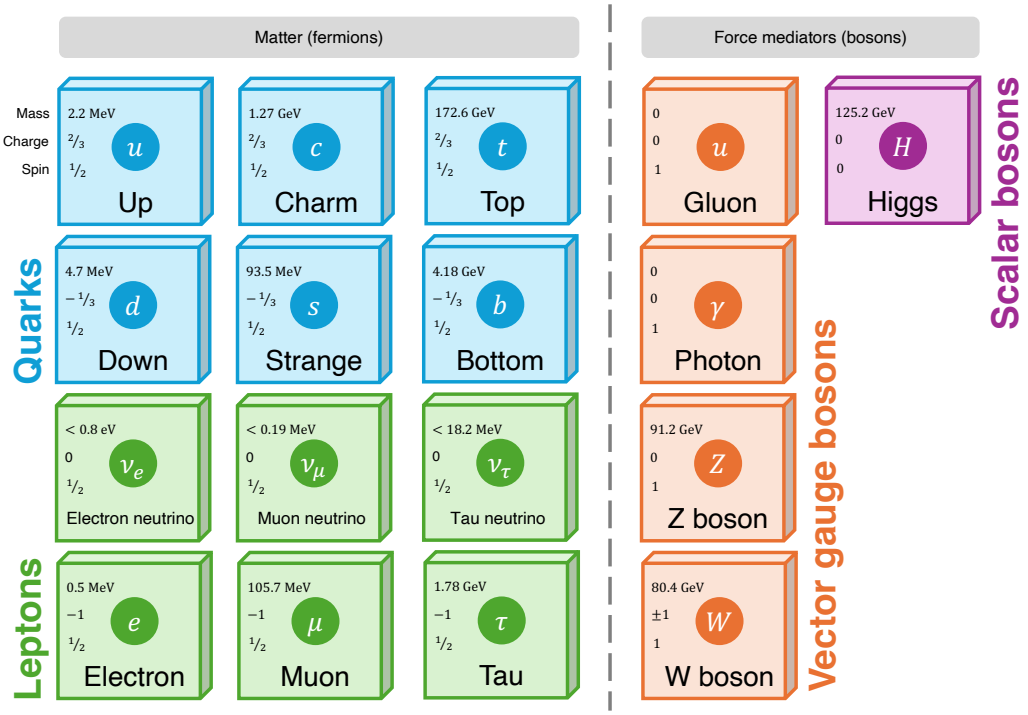


Figure 2.1: Summary of the fundamental particles described by the SM of particle physics. Each particle has a corresponding mass, charge and spin value taken from the Particle Data Group [6].

[1] If neutrinos are discovered to possess a Majorana nature, it would imply that they are their own antiparticles [5].

Colour-charged quarks interact via the strong force through the exchange of gluons. Uniquely, gluons themselves carry colour charge, enabling them to interact with each other. This property gives rise to the phenomenon of *colour confinement*, which prevents quarks from existing in isolation. Instead, quarks combine with other quarks to form colour-neutral states known as *hadrons* through a process called *hadronisation*. Hadrons are categorised into two types: *baryons*, consisting of three quarks, and *mesons*, composed of a quark-antiquark pair. These particles form in such a way that their electric charge is always zero or an integer value.

Although the SM is currently the most successful theory for describing nature, it is not without its challenges. Some of the key issues include:

- The SM fails to account for the observed asymmetry between matter and anti-matter in the Universe. While this asymmetry could arise from the phenomenon of Charge Parity Violation (CPV) (the violation of charge conjugation and parity symmetries), the extent of CPV predicted by the SM is insufficient to explain the imbalance observed in nature.
- There is no coherent way of incorporating gravity into the SM.
- The SM lacks a clear explanation for the apparent fine-tuning required to cancel the divergent quadratic contributions to the Higgs-boson mass, resulting in a mass significantly lower than expected. This issue is referred to as the *hierarchy problem*.
- The SM only describes the nature of around 5% of the Universe. The remaining 95% is composed of dark matter and dark energy, neither of which can be described using the SM.
- The SM predicts that neutrinos are massless. However, the neutrino oscillation has been observed experimentally [7–9], which proves that neutrinos must have a non-zero mass. These masses could be explained by the concept of neutrinos possessing right-handed chirality, however, there is no evidence for their existence in nature.

The limitations of the SM have driven extensive research into alternative models and searching for BSM physics. One prominent idea is *supersymmetry*, which predicts a symmetry between bosons and fermions resulting in a partner for each SM particle [10]. Another interesting area of research is *effective field theory*, which treats the SM as a

low-energy approximation of a more fundamental theory operating at higher energy scales that remain inaccessible to current experiments [11].

2.2 The top quark

The existence of the top quark was first predicted in 1973 by Makoto Kobayashi and Toshihide Maskawa, who proposed that a third generation of fermions was necessary to explain realistic models of CPV [12]. Within this third generation, the tau lepton was the first particle to be discovered, in 1975 [13]. Shortly thereafter, in 1977, the Y resonance was observed at Fermilab and interpreted as a bound state of third-generation bottom quarks ($b\bar{b}$) [14]. This discovery implied the existence of a third-generation partner to the bottom quark - the top quark [15]. The top quark itself was finally discovered in 1995 by the D0 and CDF collaborations at the Tevatron using collisions between protons and anti-protons at $\sqrt{s} = 1.8$ TeV [16, 17].

The 22-year delay between the prediction of the top quark and its discovery can be attributed to its large mass, requiring colliders with sufficiently high centre-of-mass energies, such as the Tevatron. According to the Particle Data Group [6], the top-quark mass, m_t, is measured as (172.57 ± 0.29) GeV, based on combined results from the Tevatron and LHC Run 1 measurements. The top quark's large mass gives it a very short lifetime of $\mathcal{O}(10^{-25}\text{s})$, which is shorter than the typical timescale for hadronisation, $\mathcal{O}(10^{-24}\text{s})$. Consequently, the top quark does not form bound states and instead exhibits the properties of a bare quark. As the heaviest fermion in the SM, the top quark is predicted to have the strongest coupling to the Higgs boson, with a Yukawa coupling close to one. This strong interaction with the Higgs field makes the top quark play a key role in searches for BSM physics as well as for testing the consistency of the SM. Specifically, precise measurements of m_t provide critical input for global fits of EW parameters [18, 19]. One example is the consistency test of the SM for the simultaneous indirect determination of m_t and the W-boson mass, M_W [20]. Figure 2.2 presents the results of confidence levels profiled from scans of M_W versus m_t, where the global fit either includes or excludes direct Higgs-boson mass (M_H) measurements. Both M_W and M_H are sensitive to the value of m_t, as the top quark contributes through quantum loop corrections. Figure 2.2 demonstrates that the direct measurements of m_t and M_W are consistent with the indirect measurements obtained from the global fit, highlighting the consistency of the SM.

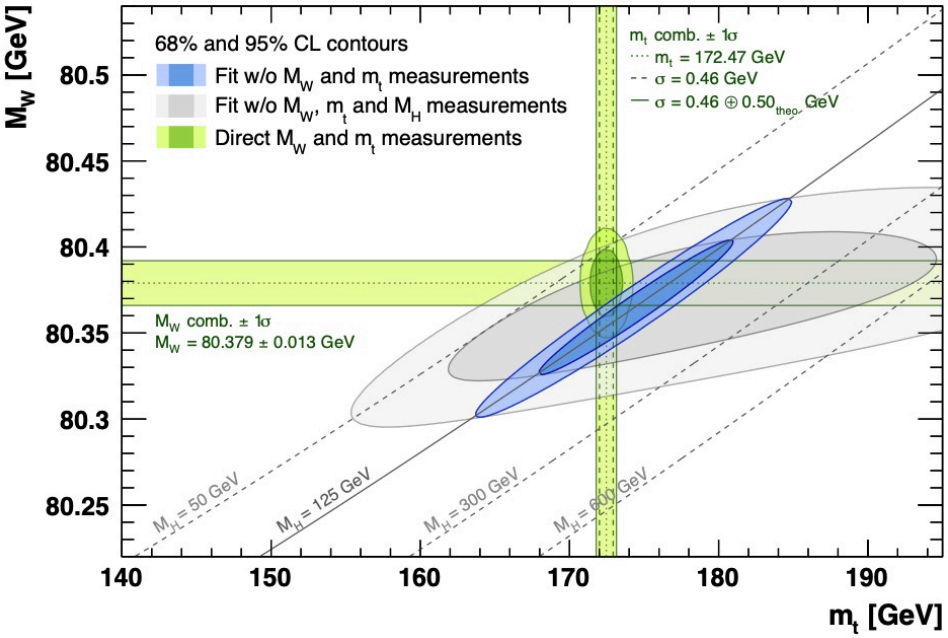


Figure 2.2: Contours at 68% and 95% confidence level obtained from scans of M_W versus m_t for the fit excluding (grey) and including (blue) the direct measurement of M_H. This is compared to the results of direct measurements (green vertical and horizontal bands, and two-dimensional ellipses). This is taken from Ref. [20].

Considering the dependence of M_H on m_t, the large Yukawa coupling associated with m_t ensures that the top quark provides the dominant contribution to the correction of M_H. Specifically, this quantum loop correction to the Higgs-boson mass, ΔM_H, is sensitive to the energy-scale cut-off used in the theoretical calculation, Λ, such that $\Delta M_H^2 \propto \Lambda^2$. As a result, the correction becomes divergent at higher energy scales, necessitating a precise fine-tuning in order to yield the observed Higgs-boson mass at the EW scale. This is the hierarchy problem mentioned in Section 2.1, which suggests the possible need for BSM physics. One proposed solution is supersymmetry, which predicts the existence of a supersymmetric partner for the top quark [10]. The loop correction from this partner would cancel the top-quark contribution, thereby eliminating the requirement for fine tuning.

Single top quarks can be produced via the EW interaction. However, at hadron colliders such as the LHC, top quarks are primarily produced in pairs of top and anti-top quarks ($t\bar{t}$) through the strong force. Figure 2.3 illustrates two Leading Order (LO) Feynman diagrams for $t\bar{t}$ production in QCD: quark-antiquark annihilation and gluon-gluon fusion. At the LHC, gluon-gluon fusion is the dominant production mechanism, contributing approximately 90% of the top quarks produced [21].

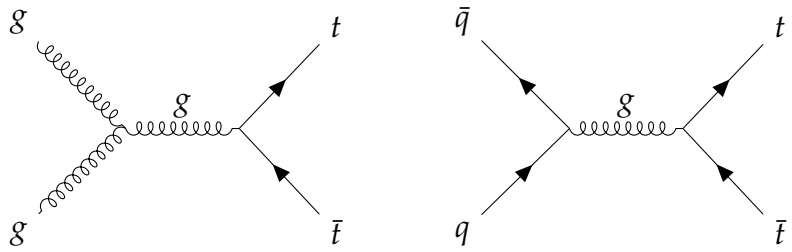


Figure 2.3: Example Feynman diagrams for the production of $t\bar{t}$ pairs at LO in QCD: gluon-gluon fusion (left) and quark-antiquark annihilation (right).

Due to its large mass, the top quark is the only quark that can decay via the flavour-changing weak interaction into an on-shell W boson and a down-type quark. These decays are governed by elements of the CKM matrix [12, 22], with the $|V_{tb}|$ element being significantly larger than $|V_{ts}|$ and $|V_{tu}|$. The top quark can also decay into an up-type quark and a neutral gauge boson or Higgs boson. However, these decays are governed by flavour-changing neutral currents, which are forbidden at tree level in the SM. These processes, which involve W-boson loops, are further suppressed by the GIM mechanism [23] and the scales of the relevant CKM matrix elements. As a result, top-quark decays are predominantly governed by the $t \to Wb$ flavour-changing charged current. Top quarks can be classified based on how the W-boson decays. If the W-boson decays into a charged lepton and a neutrino ($\ell\nu$), the top quark is referred to as a *leptonically decaying top quark*. Alternatively, if the W-boson decays into a quark and an anti-quark pair ($q\bar{q}'$), the top quark is classified as a *hadronically decaying top quark*. Consequently, $t\bar{t}$ events can be categorised into three types based on their decay products: *all-hadronic*, where both top quarks decay hadronically; *semi-leptonic* (lepton + jets), where only one top quark decays hadronically; and *dilepton*, where both top quarks decay leptonically. The various decay modes and their branching ratios, as obtained from the Particle Data Group [6], are summarised in Figure 2.4.

2.3 The top-quark mass interpretation problem

The short lifetime of the top quark ensures it decays before it can hadronise. However, because the top quark carries a strong colour charge, it cannot exist as a physically observable free particle. To define its mass consistently in a theoretical framework, one must account for QCD and EW quantum fluctuations. Unfortunately, divergences arise when including Next-To-Leading Order (NLO) terms. These divergences must be

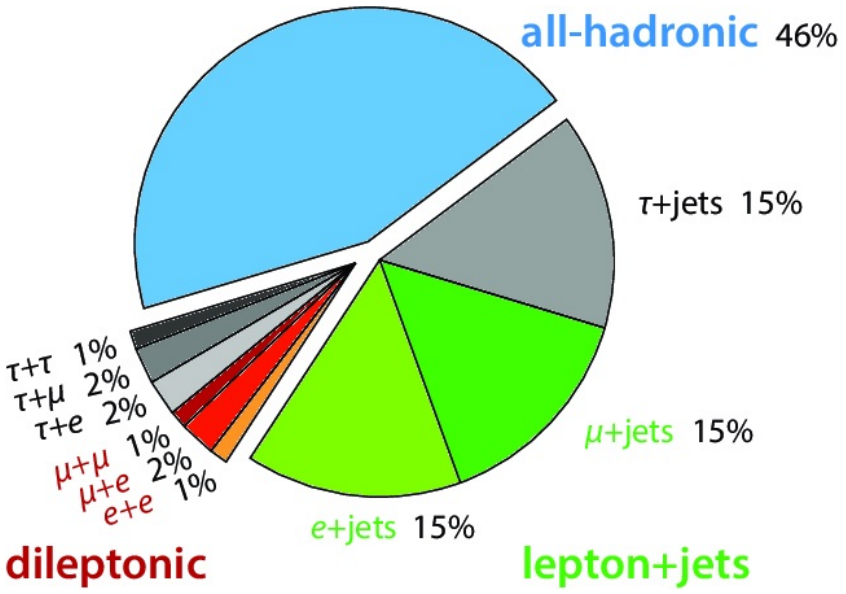


Figure 2.4: Summary of the various $t\bar{t}$ decay modes included with branching ratios taken from Particle Data Group [6]. Diagram taken from [24].

addressed systematically through the process of *renormalisation*. The various methods for renormalising the mass are known as *renormalisation schemes*, which result in different definitions of the m_t. Examples include the $\overline{\text{MS}}$ mass, the MSR mass, and the pole mass [25, 26]. The pole mass, m_t^{pole}, is commonly used for calculations in perturbation theory, corresponding to the rest mass of the top quark as a real on-shell particle. The pole mass is defined by subtracting the divergent higher-order terms, leaving the pole in the top-quark propagator fixed within the perturbative framework [6]. It is worth noting that the relation between any two m_t renormalisation schemes can be described by a perturbative series such that one can measure the mass in one scheme and translate it to another scheme.

The value of m_t can be measured using both *direct* and *indirect* techniques. Indirect methods involve measuring observables that strongly depend on m_t, and then extracting the m_t by comparing these measurements with fixed-order calculations that incorporate theoretical predictions for m_t in a specified renormalisation scheme. As such, indirect measurements provide a clear definition of the m_t parameter being determined. For example, ATLAS has performed an indirect measurement by interpreting differential cross-section data for $t\bar{t}$ + 1 jet events at $\sqrt{s} = 8$ TeV in terms of m_t^{pole} and the $\overline{\text{MS}}$ mass [27]. The advantage of indirect measurements is the clear understanding of the specific m_t being measured. However, these methods typically have uncertainties of at least 1 GeV, making them less precise than direct measurements.

Direct techniques involve the kinematic reconstruction of the top-quark decay products to build observables sensitive to m_t. These methods rely heavily on the accuracy of MC simulations, using templates derived from simulated events with varying values of the generator m_t parameter, m_t^{MC}, alongside a well-calibrated detector simulation. The m_t obtained from these measurements corresponds to the generator parameter m_t^{MC}. ATLAS has used these techniques to measure m_t with the full Run 2 dataset at $\sqrt{s}$ = 13 TeV [28]. The key advantage of direct measurements is their potential for higher precision compared to indirect methods. However, a the main drawback is the challenge in identifying which m_t parameter is being determined. For example, the MC event generator used to simulate the signal $t\bar{t}$ events used in the measurement described in Part III uses the m_t^{pole} definition at NLO. The issue is that when using a parton shower to simulate the many emissions coming from the resulting partons from the $t\bar{t}$ event, there is an energy cut-off at around 1 GeV. However, the definition of m_t^{pole} absorbs all higher-order contributions down to zero momentum. This discrepancy could, in principle, disrupt the definition of m_t obtained in experiments, and it remains an open question whether the low-energy effects from the parton shower and hadronisation (see Chapter 4) cause a shift in m_t^{MC} relative to m_t^{pole}. The challenge of reconciling the top-quark mass parameters defined within specific renormalisation schemes with m_t^{MC} is known as the *top-quark mass interpretation problem*, and it is an ongoing challenge in the field [25, 26, 29, 30]. In practice, m_t^{MC} is often associated with m_t^{pole} to a precision of 0.5 GeV.

Figure 2.5 presents a selection of previous direct m_t measurements, including the combined results from the Tevatron [31] and from Run 1 of the LHC [32], the first ATLAS Run 2 result [28], the most precise single measurement in a single decay channel, performed by CMS [33], and the latest CMS result based on top quarks with high transverse momentum (boosted) [34]. These measurements exhibit a degree of scatter, underscoring the need for additional high-precision m_t measurements to clarify the current picture. In response to this, Part III presents a recent measurement of m_t, which is the most precise determination performed by the ATLAS collaboration in a single decay channel to date [1].

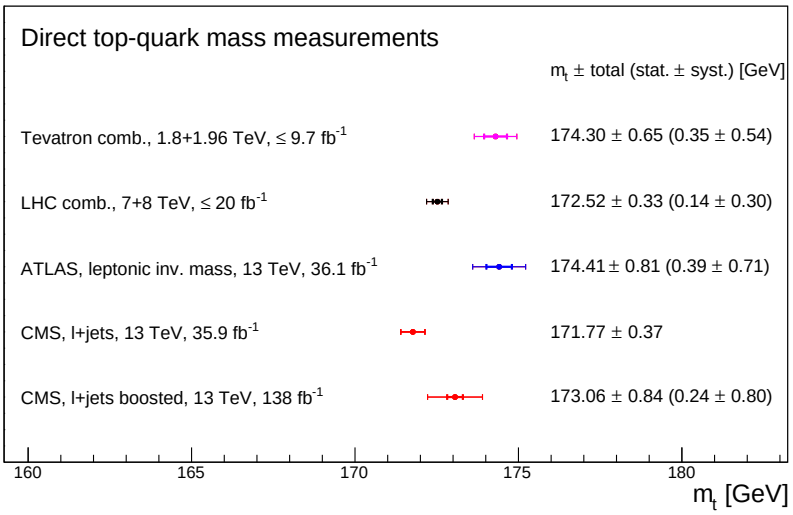


Figure 2.5: A selection of precise direct m_t results from different experimental collaborations.

CHAPTER 3

The ATLAS experiment at the LHC

The experimental apparatus in particle physics has undergone a remarkable transformation, evolving from early tabletop cloud chambers that were used to discover the positron in 1932 [35], the muon in 1937 [36], and the kaon in 1947 [37], to the LHC [38], the largest machine ever built.

This chapter provides an introduction to the particle accelerator complex of CERN, the world's leading research facility in particle physics, which hosts the LHC and has been the source of numerous groundbreaking discoveries. These include the observation of neutral currents in 1974 [39], the discovery of the W and Z bosons in 1983 [40–43], direct CPV in 1999 [44], and the Higgs boson in 2012 [45, 46].

This chapter will then introduce the ATLAS detector, outlining the design and function of each subsystem. As the measurement of the m_t presented in this thesis is based on data collected during Run 2 of the LHC, the discussion will cover the CERN accelerator complex, the LHC, and the ATLAS detector as they operated during this period.

3.1 The CERN accelerator complex

The high-energy particle beams used in the LHC do not originate within the ring itself. Instead, a series of accelerators progressively increase the beam's energy before injecting it into the LHC. These accelerators are part of the CERN accelerator complex, as shown in Figure 3.1.

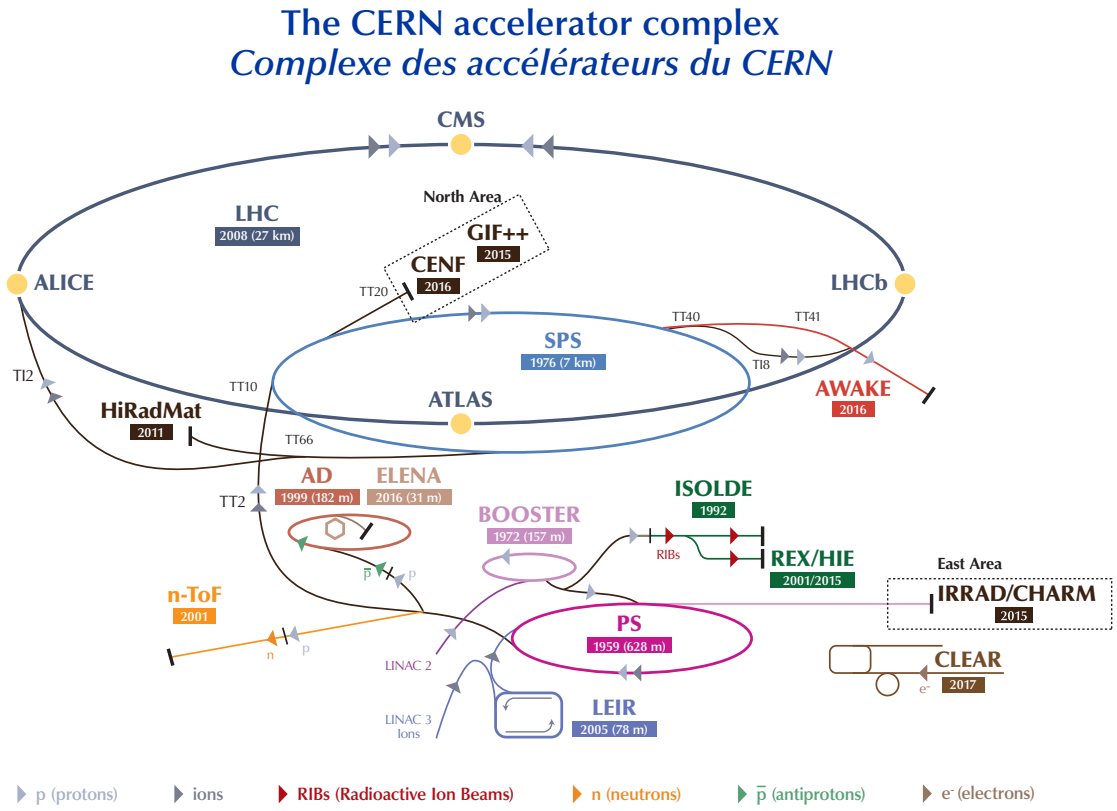


Figure 3.1: A diagram of the CERN accelerator complex during Run 2 taken from Ref. [47].

Most collisions in the LHC result from high-energy proton beams colliding almost head-on. While heavy ions such as lead, gold, and xenon are also used, this thesis will focus exclusively on proton-proton (pp) collisions.

The protons originate from a simple hydrogen gas source, which is ionized to produce H^- ions. These ions are first accelerated to 50 MeV using the LINAC2 [48, 49] linear accelerator[1]. An electric field then strips the two electrons from each ion, leaving a single proton. The protons are subsequently injected into the Proton Synchrotron Booster [51, 52], where they are accelerated to 1.4 GeV, followed by the Proton Synchrotron [53], which boosts their energy to 25 GeV. They then enter the Super Proton

[1] LINAC2 was replaced with LINAC4 in 2020, increasing the energy of the accelerated ions to 160 MeV [50].

Synchrotron [54], where their energy increases further to 450 GeV. Finally, the protons are injected into the LHC, split into two beams, and directed into opposite circulatory paths at two separate injection points. During Run 2, each beam reached a final energy of 6.5 TeV. Throughout the acceleration chain, both electromagnets and radio frequency cavities are used to steer, focus, and accelerate the protons.

3.2 The Large Hadron Collider

3.2.1 Overview

The LHC [38] is a synchrotron housed in a 27 km-circumference tunnel, originally built for the Large Electron-Positron collider (LEP) [55], approximately 100 m underground. Proton beams are injected into the LHC in a structured *bunch* formation, where each bunch contains approximately 1.2×10^{11} protons, with a spacing of 25 ns between them [56]. This structured arrangement is maintained using 392 quadrupole magnets. The two counter-rotating beams are kept on a circular trajectory[2] by 1,232 superconducting dipole magnets, which are cooled using liquid helium. These magnets generate strong magnetic fields exceeding 8 T, perpendicular to the direction of the beams, ensuring that the Lorentz force directs them toward the centre of the LHC ring. Additionally, numerous quadrupole and higher-order magnets correct for beam spreading over time, maintaining beam stability and focus.

Collisions at the LHC occur at four designated locations along the ring, each known as an Interaction Point (IP). Four major experiments are positioned at these locations: two general-purpose detectors: ATLAS [57] and CMS [58] performing a broad range of SM measurements and searches, and two specialised detectors: ALICE [59], dedicated to heavy-ion physics, and LHCb [60], focused on heavy-flavour physics.

3.2.2 Luminosity and centre-of-mass energy

One of the most important parameters of any particle collider is the *instantaneous luminosity*, $\mathcal{L}_{\text{inst}}$, which quantifies the rate of collisions. It is defined such that the

[2] Although the LHC is often described as a circular ring, it is actually an octagonal structure consisting of eight straight sections connected by arcs.

interaction rate over time is given by

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \sigma \mathcal{L}_{\text{inst}}, \tag{3.1}$$

where σ is the *cross-section* for the interaction, representing the probability that the interaction will occur as a result of the collision. In terms of collisions between two opposing bunched beams of particles with Gaussian profiles, $\mathcal{L}_{\text{inst}}$ depends only on the properties of the colliding bunches and is given by

$$\mathcal{L}_{\text{inst}} = \frac{N_1 N_2 N_b f}{4\pi\sigma_x\sigma_y}, \tag{3.2}$$

where N_1 and N_2 are the number of protons per bunch in each beam, N_b is the number of bunches in each beam, f is the frequency of collisions determined by the revolution frequency of the LHC, and σ_x and σ_y are the horizontal and vertical bunch dimensions [38]. The LHC was originally designed to achieve an instantaneous luminosity of 10^{34} cm^{-2}s^{-1}, however, this has been surpassed during Run 2, and has achieved a peak of 2.3×10^{34} cm^{-2}s^{-1} during Run 3 data-taking [61].

A related metric is the *integrated luminosity*, $\mathcal{L}$, used as a measure of the amount of recorded data over a given period. It is obtained by integrating the instantaneous luminosity over time and is typically expressed in inverse femtobarns (fb^{-1}) where one barn is equal to 10^{-24} cm^2. Given the integrated luminosity and cross-section for a specific interaction, the expected number of observed events is

$$N = \sigma\mathcal{L}. \tag{3.3}$$

It is important to note that the integrated luminosity delivered by a collider is not necessarily the same as the luminosity recorded by an experimental detector due to detector inefficiencies.

Another important parameter of a particle collider is the *centre-of-mass energy*, $\sqrt{s}$, which is a measure of the energy of the colliding beams and therefore the energy available for the creation of new particles. A collider with a higher $\sqrt{s}$ is able to produce more massive or energetic particles. For the case of a collider with two opposing beams of equal energy, $\sqrt{s}$ is given by the total energy of the two beams.

3.2.3 Operation

Collisions at the LHC are conducted in distinct periods known as *Runs*, each spanning multiple years. So far, two major Runs have been completed, with a third currently ongoing. These Runs are separated by periods of *Long Shutdowns*, during which the LHC is powered down for maintenance and upgrades. LHC operations began in 2009 with Run 1, during which $\sqrt{s}$, was increased from 900 GeV to 7 TeV and later 8 TeV, before concluding in 2012. This was followed by Long Shutdown 1, during which upgrades to the magnet structure allowed Run 2 to operate at 13 TeV from 2015 to 2018. Long Shutdown 2 brought further maintenance and enhancements, increasing $\sqrt{s}$ to 13.6 TeV for Run 3, which began in 2022 and is scheduled to continue until 2026. The cumulative daily luminosity delivered to ATLAS between 2011 and 2024 is shown in Figure 3.2.

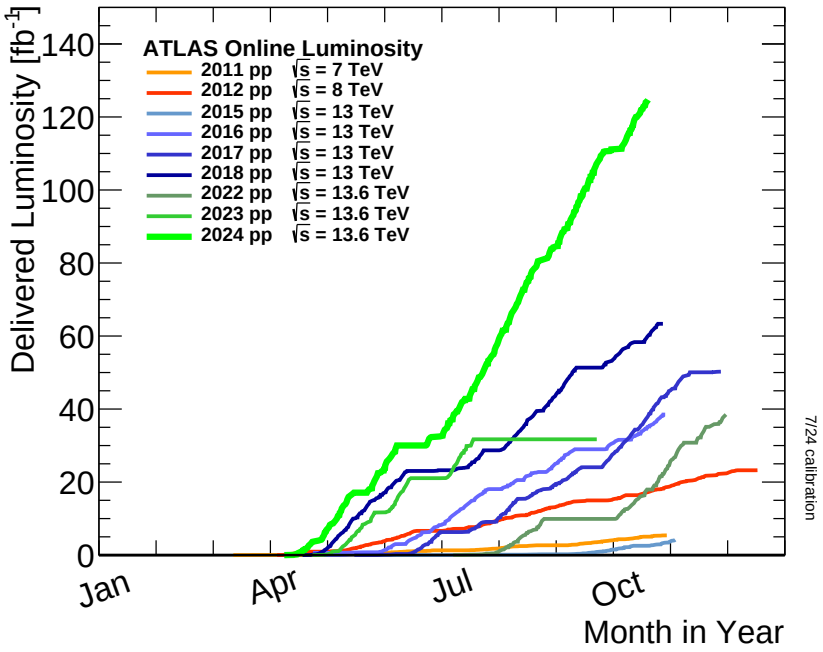


Figure 3.2: The daily cumulative luminosity delivered to ATLAS during stable beams and for high energy proton-proton collisions in the 2011-2024 period. This is taken from Ref. [61].

3.3 The ATLAS detector

The ATLAS detector [57] is a general-purpose particle detector with forward-backward cylindrical geometry aligned along the beamline, providing nearly full 4π solid-angle coverage around the IP. It is the largest particle detector ever built, measuring 44 m in length and 25 m in height, with a total weight of approximately 7,000 tonnes, comparable to the weight of the Eiffel Tower. This immense size is necessary to accommodate multiple sub-detector layers, each designed to identify particles and measure their properties. The Inner Detector (ID), situated closest to the IP, is designed to be radiation-hard and highly granular, allowing precise tracking of electrically charged particles. Surrounding the ID are the EM and hadronic calorimeters, which measure the energy of both charged and neutral particles, with the exception of muons and neutrinos. Since neutrinos pass through all detector layers without interacting, they can only be inferred by calculating the missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$. The outermost layer of the detector is the Muon Spectrometer (MS), which records the tracks of muons. Muons pass through all layers of the detector while depositing only a portion of their energy. To enable accurate particle tracking and momentum measurements, ATLAS also incorporates superconducting magnet systems, which provide the magnetic fields in the ID and MS sub-detector systems. A schematic diagram of the ATLAS detector, highlighting its main sub-detector systems, is shown in Figure 3.3.

3.3.1 Coordinate system

The ATLAS detector employs a right-handed Cartesian coordinate system (x, y, z) centred at the IP. In this system, the z-axis aligns with the beam axis, the x-axis is directed to the centre of the LHC ring, and the y-axis extends vertically upwards. The x-y plane is defined as the plane *transverse* to the direction of the beam, and is used to define quantities such as a particle's transverse momentum, p_{T}.

In addition to the Cartesian system, ATLAS also uses a cylindrical polar coordinate system (r, ϕ, θ), with the radial distance r and the azimuthal angle ϕ (measured around the z-axis) used in the transverse plane. Instead of using the polar angle θ (measured

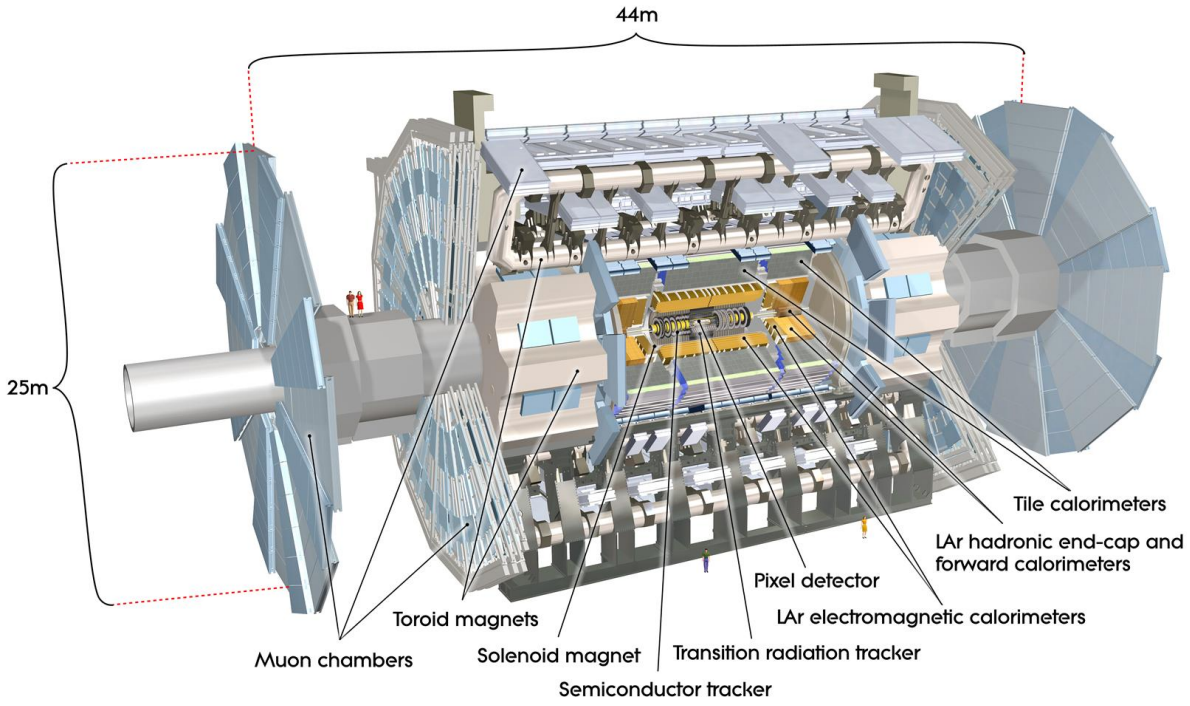


Figure 3.3: Schematic diagram of the ATLAS detector during Run 2, taken from Ref. [57].

from the z-axis), the *pseudorapidity*, η is used, defined as

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

In the relativistic limit, this is equivalent to the *rapidity*, y, given by

$$y = \frac{1}{2}\ln\left(\frac{E + p_z c}{E - p_z c}\right), \tag{3.5}$$

where E is the particle energy and p_z is its momentum along the z-direction. Pseudorapidity is defined such that a value of zero refers to a particle travelling entirely within the transverse plane, and $|\eta| \to \infty$ for a particle moving in a direction parallel to the beamline. Differences in rapidity (or pseudorapidity for the relativistic limit) are invariant under Lorentz boosts in the z-direction. Using the difference in azimuthal angle, $\Delta\phi$, and pseudorapidity, $\Delta\eta$, between two particles, their *angular separation* can be expressed as

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}. \tag{3.6}$$

3.3.2 Magnet system

The ATLAS detector utilises two superconducting magnet systems [57] to generate the magnetic fields necessary for particle charge identification and precise momentum measurements within the ID and MS. A schematic representation of these magnet systems is shown in Figure 3.4. The ID magnetic field is produced by a solenoid magnet, generating a uniform field strength of 2 T. In contrast, the MS relies on a system of three sets of eight toroidal magnets: one set for the barrel region and one for each end-cap region. The field strength varies depending on the region, averaging 0.5 T in the barrel and 1 T in the end-caps.

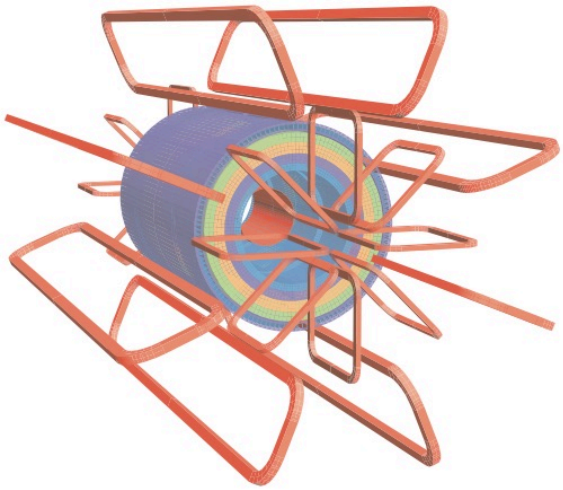


Figure 3.4: Schematic diagram of the superconducting magnet system used for the ATLAS detector, taken from Ref. [57]. The solenoid lies in the centre shown by the coloured cylinders, and the toroidal system is shown in red.

3.3.3 Inner detector

The ATLAS ID [62,63], illustrated in Figure 3.5, is the closest sub-detector system to the IP. It is specifically designed to deliver exceptional tracking and momentum resolution, as well as vertex reconstruction for charged particles above a certain p_{T} threshold. The detector operates within a pseudorapidity range of $|\eta| < 2.5$. The ID is composed of three distinct technologies, each arranged in layers immersed in a magnetic field. The innermost layer is the Pixel detector, followed by the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT), which forms the outermost layer.

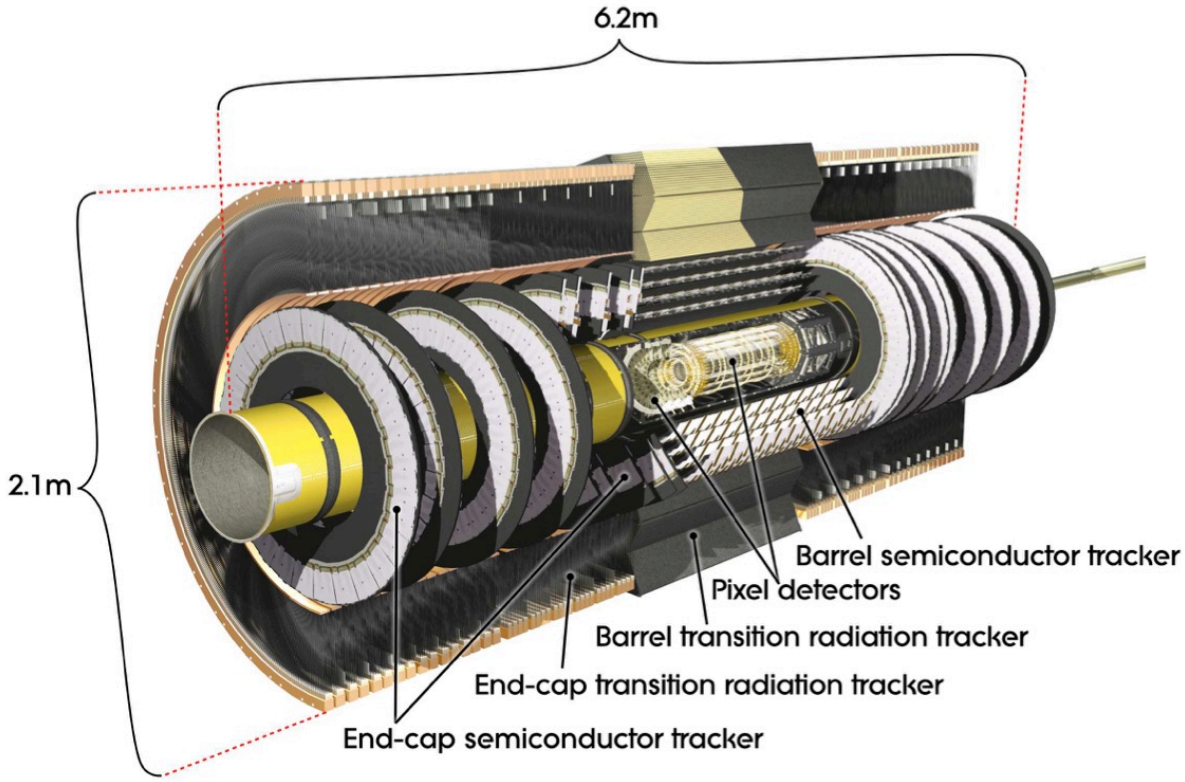


Figure 3.5: Schematic diagram of the ID system used for the ATLAS detector, taken from Ref. [57]. The different layers of technologies are labelled, but the surrounding solenoid is not shown.

Pixel detector

The Pixel detector [64] is a high-granularity system composed of four layers of silicon sensors in the barrel and three disks in the end-caps. Following Long Shutdown 1 of the LHC, the Insertable B-Layer (IBL) [65, 66] was added to the detector. This is shown schematically in Figure 3.6, along with the different layers of the ID.

Situated closest to the beamline, the IBL is designed to improve the identification of secondary vertices associated with B-hadron decays, whilst prolonging the detector's lifespan in regions expected to endure high radiation damage. The IBL uses pixels with size 50 μm $\times$ 250 μm, while the other layers use pixels with size 50 μm $\times$ 400 μm.

The Pixel detector consists of approximately 92 million pixels that cover an area of around 1.9 m^2, encompassing the entire azimuthal angle range and a pseudorapidity range of $|\eta| < 3.0$. The high-granularity of the Pixel detector is reflected in its spatial resolution: 115 μm along the beamline and 10 μm perpendicular to the beamline.

Charged particles are detected via the creation of electron-hole pairs in the doped silicon substrate of the pixel sensors. These pairs generate an electrical signal when the electrons reach an anode, and this signal is recorded as a *hit* in the detector.

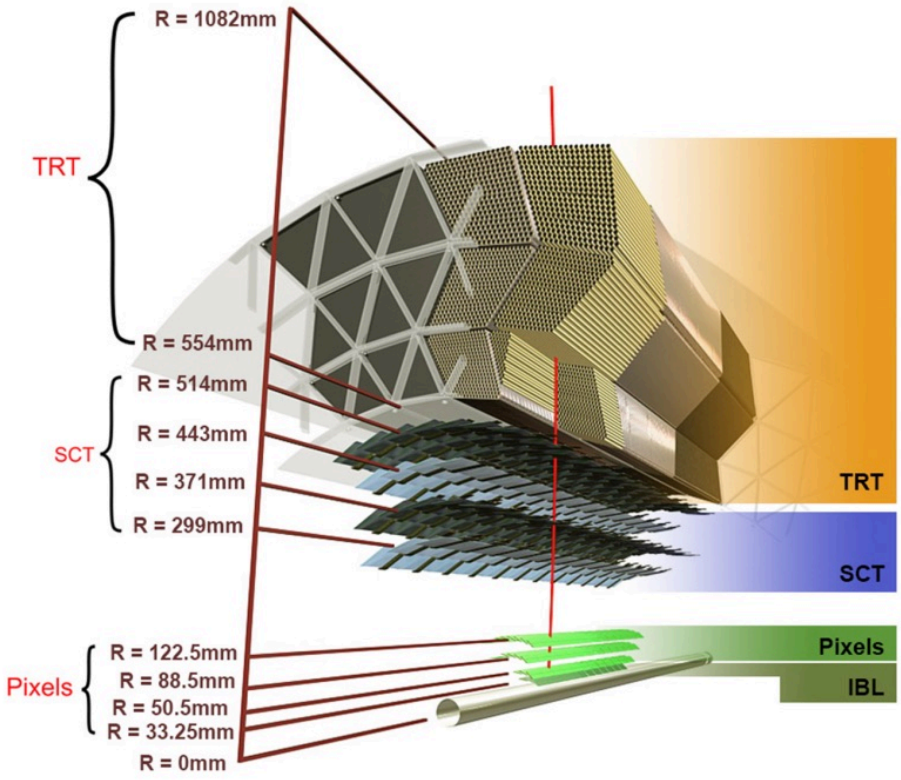


Figure 3.6: Schematic diagram of layers of the ID system used for the ATLAS detector, taken from Ref. [67].

Semiconductor tracker

The ATLAS SCT [68] consists of silicon sensors similar to those in the Pixel detector but uses long strip sensors instead of pixels. It surrounds the Pixel detector, as illustrated in Figure 3.6, and covers a pseudorapidity range of $|\eta| < 2.5$. The SCT is made up of 4088 strip modules arranged across four concentric cylindrical layers in the barrel region and nine discs in the end-caps, covering a total area of approximately 60 m^2.

Each SCT module contains two strip sensors placed at slight angles to improve spatial resolution. The choice to use long silicon strips rather than pixels helped keep manufacturing costs low, though it comes at the price of reduced granularity. The detector resolution along the beamline is 580 μm, while the resolution perpendicular to

the beamline is 17 μm. Like the Pixel detector, charged particles are detected through the creation of electron-hole pairs in the silicon substrate, which generate an electrical signal when the electrons reach an anode.

Transition radiation tracker

The ATLAS TRT [69] surrounds the Pixel and SCT sub-detector systems, as shown in Figure 3.6, and is the largest section of the ID. It consists of gas-filled straws with a diameter of 4 mm, each containing a single positively charged gold-plated tungsten wire of 30 μm diameter at its centre, running along the length of the straw. The straws were originally filled with a mixture of xenon, carbon dioxide, and oxygen gases; however, due to significant gas leaks, the expensive xenon was replaced with argon in parts of the TRT.

When a charged particle passes through the straws, it ionises the gas, causing the resulting electrons to travel towards the anode and generate an electrical signal. The straws are arranged in 73 layers in the barrel and 122 layers in the end-caps, allowing for approximately 30-40 particle hits per track. With an acceptance range of $|\eta| < 2.0$, the TRT provides additional spatial hit information, complementing the already precise track measurements from the Pixel and SCT sub-detectors, thereby enhancing the accuracy of momentum measurements.

The TRT spatial resolution in the direction perpendicular to the beamline is 130 μm, which is notably larger than the resolutions achieved by the Pixel and SCT sub-detectors. However, the gases used in the TRT absorb transition radiation from charged particles, with the effect being most pronounced for electrons, due to their low mass [70]. This capability allows the TRT to aid in the identification of electrons, helping to distinguish them from pion tracks.

3.3.4 Calorimeters

Surrounding the ATLAS ID is a system of calorimeters [71], designed to measure the energy of electrons, photons, and hadrons. The calorimeters are constructed with alternating layers of *absorber* and *active* materials. When a particle reaches an absorber layer, it interacts with the material and loses energy, producing cascades of

secondary lower-energy particles, known as *showers*. The energy of these showers is then measured by the active layers.

Particles interact with the absorber material, losing energy in different ways. Electrons and photons primarily lose energy through EM interactions to produce EM showers. Electrons with energies above 10 MeV predominantly lose energy via Bremsstrahlung[3], while photons lose energy mainly through the production of electron-positron pairs. On the other hand, hadrons lose energy through strong interactions, producing hadronic showers.

The characteristics of these hadronic showers depend on the material's *nuclear interaction length* (λ_I), which is typically much larger than the *radiation length* (X_0) that describes the frequency of EM interactions. This difference leads to the expectation that photons and electrons will interact with detector material before hadrons.

The ATLAS calorimeter system consists of two types of calorimeters: an EM calorimeter to capture and measure the energy of EM showers, and a hadronic calorimeter to measure the energy of hadronic showers. Additionally, a forward calorimeter is included to enhance the precision in the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$, covering the very high $|\eta|$ region.

A schematic diagram of the ATLAS calorimeter system is shown in Figure 3.7.

Electromagnetic calorimeter

The Electromagnetic Calorimeter (ECal) [72] is the closest calorimeter system to the IP. It is designed using lead as the absorber material and liquid argon as the active material. Electrons and photons that produce showers in the lead plates ionise the liquid argon, generating an electrical signal. The ECal has an acceptance of $|\eta| < 3.2$ over the full range of ϕ. It is composed of three fine-granularity layers in the barrel region ($|\eta| < 1.475$) and two coarser-granularity layers in the end-caps ($1.375 < |\eta| < 3.2$). In terms of thickness, the ECal is $22X_0$ thick in the barrel and over $24X_0$ thick in the end-caps, ensuring that it captures as much of the EM showers as possible.

Additionally, a *presampler* layer was included in the barrel region, situated between the ID and the ECal, to correct for energy losses of electrons and photons in the ID.

[3]The process by which a charged particle emits a photon when it is decelerated or deflected by an electric field.

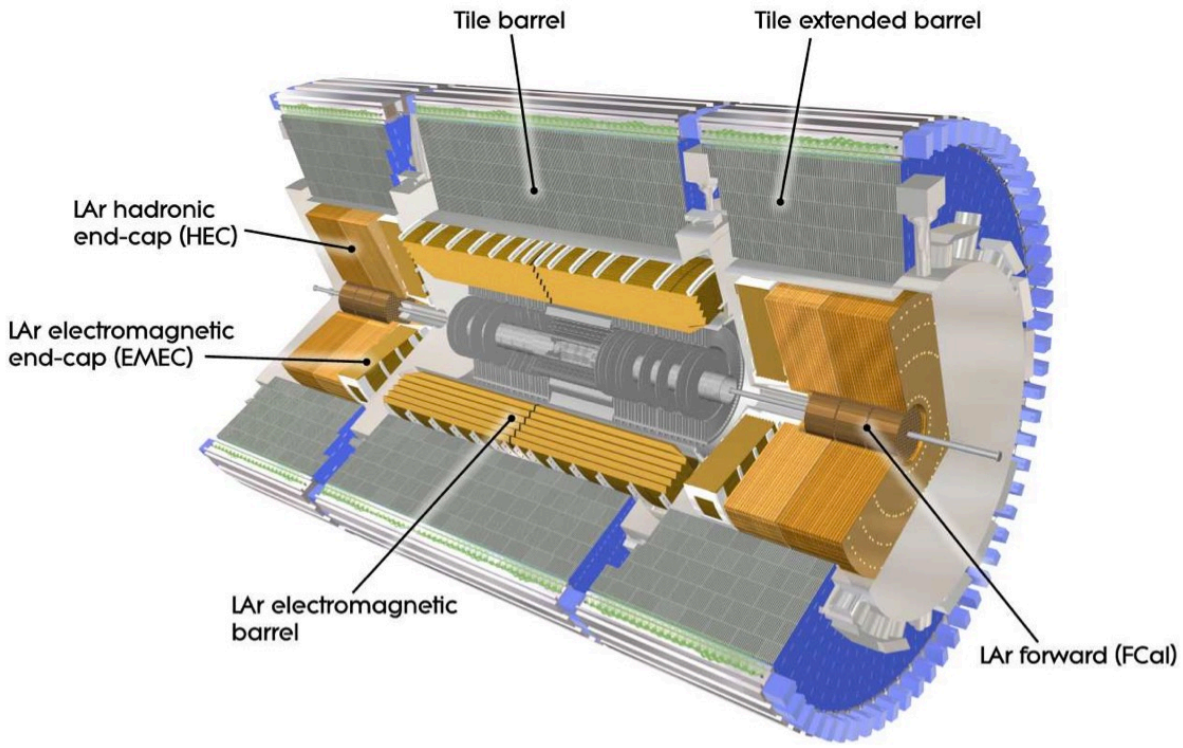


Figure 3.7: Schematic diagram of the ATLAS calorimeter system, taken from Ref. [57].

Hadronic calorimeter

The ATLAS Hadronic Calorimeter (HCal) system [73] surrounds the ECal and is designed to capture the remaining energy of hadrons that pass through the entire ECal. The HCal utilises both liquid argon and plastic scintillating tile technologies for its active layers.

In the three barrel layers ($|\eta| < 1.7$), the HCal uses steel as the absorber material and scintillating tiles as the active material. When hadrons interact with the steel, they create showers, and the particles in these showers interact with the scintillator material to produce photons, which are then converted into an electrical signal. For the four end-cap layers ($1.7 < |\eta| < 3.2$), copper is used as the absorber material, and liquid argon serves as the active material. Although the HCal has lower granularity compared to the ECal, it is larger in size to ensure that high-energy particles are fully absorbed and to reduce the chance of them passing through the calorimeter system undetected.

Forward calorimeter

The final component of the ATLAS calorimeter system is the dedicated Forward Calorimeter (FCal) [72], which extends the calorimeter coverage to the $3.2 < |\eta| < 4.9$ region. This extension is specifically designed to enhance the precision in calculating the $E_{\mathrm{T}}^{\mathrm{miss}}$.

The FCal consists of three layers, all using liquid argon as the active material, but with different absorber layers. The first absorber layer is made of copper, chosen for its sensitivity to EM showers. The second and third absorber layers are made of tungsten, selected for its sensitivity to hadronic showers.

3.3.5 Muon spectrometer

The ATLAS MS [74] is the outermost part of the detector, located around the calorimeters. It is designed to measure the momentum of charged particles that pass through the calorimeters without depositing all of their energy. Since neutrinos escape the detector without interacting, the only SM particles to behave in this way are muons. However, sufficiently energetic jets can punch through the calorimeters and reach the MS, leading to fake muon signatures.

The MS is composed of four sub-detector systems: Monitored Drift Tubes (MDT) [75] and Cathode Strip Chambers (CSC) [76] for precision tracking of muons, and Resistive Plate Chambers (RPC) [77] and Thin Gap Chambers (TGC) [78] for fast identification of muon candidates used in the ATLAS hardware trigger (see Section 3.3.6). The MS is structured in three layers: the barrel, consisting of concentric cylindrical layers, and the transition region and end-caps, both formed of wheels that surround the beamline. Each of the sub-detectors is immersed in a magnetic field produced by the superconducting toroidal magnet system. A schematic of the ATLAS MS during Run 2 is shown in Figure 3.8. Figure 3.9 shows the layout of the four sub-detector technologies for one quarter of the MS during Run 2, relative to the beamline.

Precision muon tracking detectors

Precision muon tracking is achieved through the use of the MDT and CSC. The MDT are employed in both the barrel and end-caps, using a similar technology to the TRT.

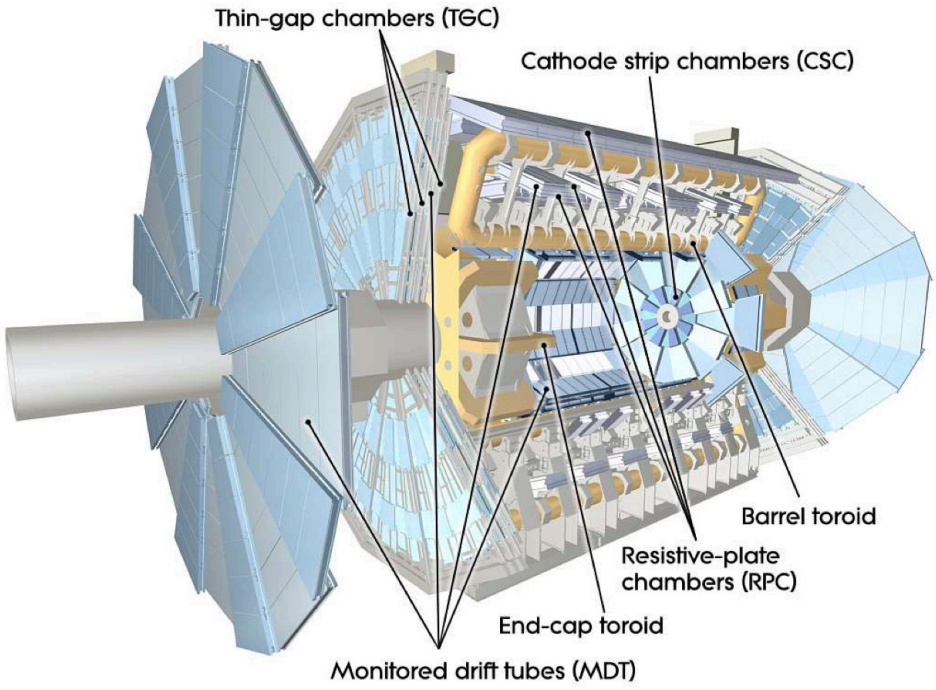


Figure 3.8: Schematic diagram of the ATLAS MS system during Run 2, taken from Ref. [57].

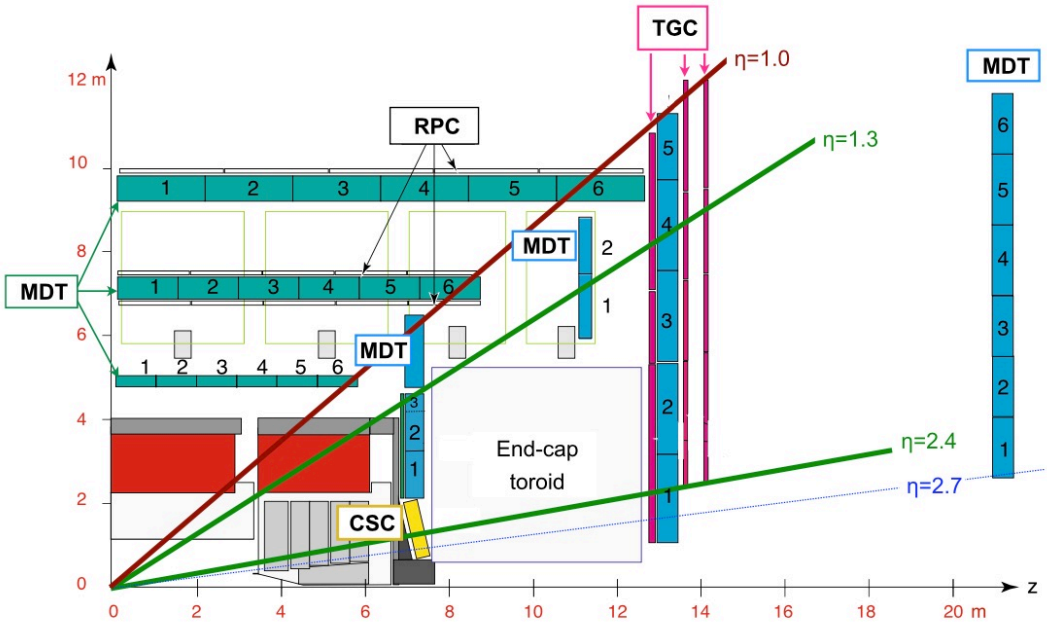


Figure 3.9: Schematic diagram of one quarter of the ATLAS MS system during Run 2, taken from Ref. [79].

These tubes are made of aluminium with a diameter of 3 cm and are filled with a mixture of argon and carbon dioxide gases. When a muon passes through a drift tube,

it ionises the gas, releasing electrons that then travel toward the positive tungsten anode wire at the centre of the tube, generating an electrical signal. The MDT are arranged into chambers, where each chamber consists of two groups of either three or four layers of tubes, covering a pseudorapidity range of $|\eta| < 2.7$, except for the innermost MDT layer in the end-caps, which covers $|\eta| < 2.0$. The precision tracking capability of the MDT results in a spatial resolution of approximately 35 μm.

In addition to the MDT, the CSC provides precision muon tracking in the $2.0 < |\eta| < 2.7$ region. The CSC are multi-wire proportional chambers that utilise cathode strips surrounding anode wires. The gap between the strips is filled with a mixture of argon and carbon dioxide gases. The strips are arranged into layers oriented either perpendicular or parallel to the anode wires, allowing for position determination in both the ϕ and η directions. This configuration improves spatial resolution, achieving a precision of approximately 40 μm. As with the MDT, when a muon passes through a CSC, the gas is ionised, causing the ions and electrons to travel toward the cathode and anode, respectively.

Fast muon tracking detectors

Despite the impressive spatial resolution of the MDT and CSC, they are still limited by their response time. The fast identification of muons is ensured by the use of the RPC and TGC. The gaseous RPC detectors are arranged into three layers within the $|\eta| < 1.05$ region, providing crucial information for the hardware trigger (see Section 3.3.6). Each RPC consists of two parallel electrodes separated by an electric field, covering a gap of only a few millimetres, which is filled with a mixture of $C_2H_2F_4$, Iso-C_4H_1C, and SF_6 gases. When a muon traverses this detector, the gas is ionised, creating avalanches that generate a large and almost instantaneous electrical signal. Although the spatial resolution of these RPC is relatively coarse at 10 mm, much worse than the precision detectors, they are capable of achieving an impressive time resolution as low as 1.5 ns.

These RPC are supplemented by the TGC, which are multi-wire proportional chambers similar to the CSC. However, unlike the CSC, the distance between the anode wires and cathodes in the TGC is shorter than the distance between the wires themselves. This design enables the TGC to achieve greater than 99% efficiency when providing muon tracking information at the 25 ns interval, matching the LHC bunch spacing.

3.3.6 Trigger and data acquisition system

The ATLAS Trigger and Data Acquisition (TDAQ) system [80] is a critical component that enables the detector to record interesting[4] physics collisions for offline analysis. Since proton bunches at the LHC collide up to every 25 ns, recording every event would produce approximately 60 TB of data per second. This is computationally infeasible, as it is neither possible to read out data at such a rate nor to permanently store this volume of data. Additionally, a large fraction of these events would not be suitable for further analysis, as they wouldn't contain interesting physics. To address this challenge, ATLAS employs a real-time system to analyse collisions and decide which events should be stored for offline analysis.

The TDAQ system includes a Level-1 (L1) hardware trigger, which then feeds into a High-Level Trigger (HLT). The HLT performs more complex algorithms to reconstruct full events and make the final decision about whether to retain the event for permanent storage. The L1 trigger's primary task is to reduce the event rate from the initial maximum input of 40 MHz down to less than 100 kHz. The HLT further reduces this rate to approximately 1 kHz, a manageable event rate that is feasible to process and store. The ATLAS TDAQ system, as used during Run 2, is shown in Figure 3.10, and the details of the L1 trigger and HLT are discussed in the following sections.

Level-1 trigger

The ATLAS L1 trigger is a hardware-based system that processes events at a rate of 40 MHz, with a latency of 2.5 μs per event. It uses reduced-granularity data from the calorimeters, along with information from the RPC and TGC in the MS. The system is designed to identify collisions containing interesting physics objects, such as high-p_{T} leptons, jets, or large $E_{\mathrm{T}}^{\mathrm{miss}}$, by scanning the full $(\eta - \phi)$ space and selecting candidates within smaller Regions of Interest (RoI), which are then passed to the HLT for further processing.

[4] The definition of "interesting" physics is inherently subjective, but it typically refers to processes involving rare or exotic signatures, interactions with the Higgs boson or top quark, or phenomena that enable precision tests of the SM or searches for BSM physics.

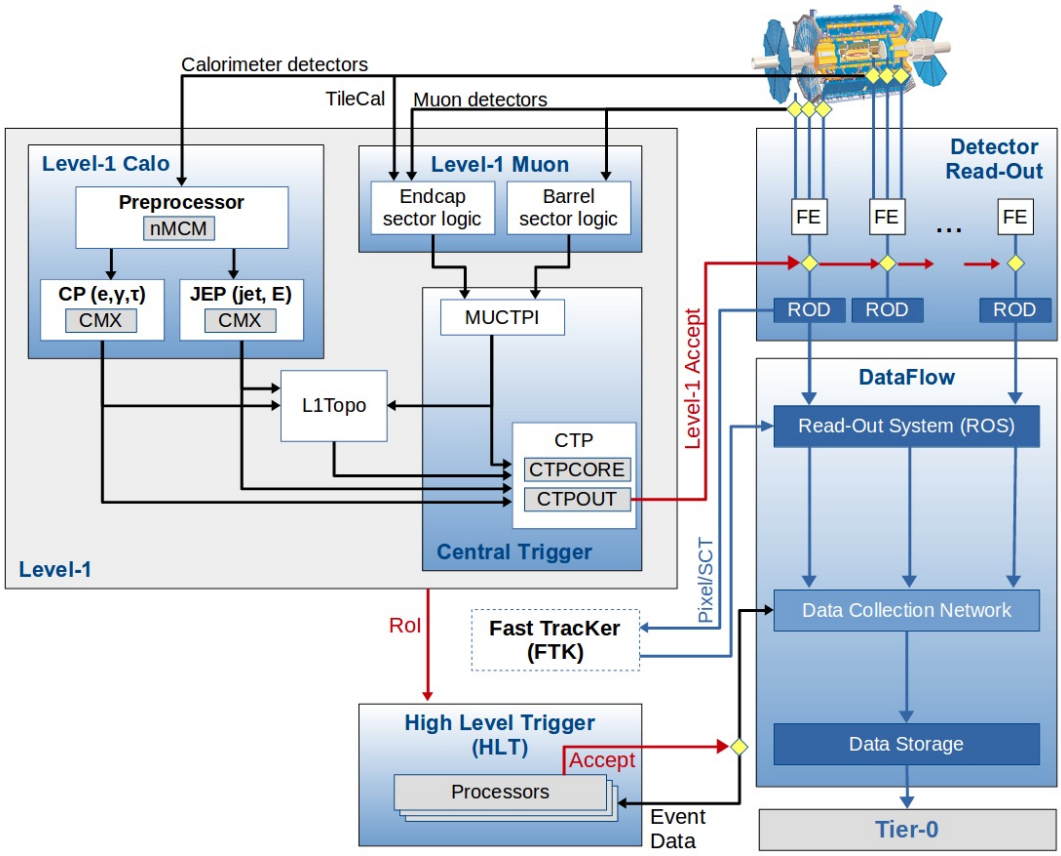


Figure 3.10: Flowchart of the ATLAS TDAQ system during Run 2, taken from Ref. [81].

High level trigger

The ATLAS HLT is a software-based system that processes data from the L1 trigger. With access to the full event information, it applies increasingly sophisticated algorithms to reconstruct events within the RoIs identified by the L1 trigger and makes the final decision on whether to retain an event for offline analysis. The HLT is a farm of CPU cores, designed to process each event and make a trigger decision within an average time of 200 μs. This is achieved using *trigger chains*, which are sequences of progressively more computationally intensive algorithms for object and event reconstruction, where events may be rejected at each stage to conserve CPU resources. The chains are defined to target specific physics objects, such as electrons, muons, or hadronic jets (see Section 5.5). Together, these chains form a *trigger menu*, ensuring that the full ATLAS physics programme is covered. The menu is modular and optimised to maximise the use of the 1 kHz output event rate. To control the event rate, *prescales* are applied to reduce the frequency of triggers that are expected to fire more

often, reserving CPU resources for rarer events. Events that pass the reconstruction requirements of at least one chain are kept for permanent storage, while events failing all reconstruction steps in every chain are discarded and lost forever. Thus, the ATLAS TDAQ system must be effective, fast, and robust to fully exploit the physics potential of the LHC.

CHAPTER 4

Monte Carlo event simulation

A critical component of any physics analysis is the comparison between experimental results and theoretical predictions. This requires a method for generating the expected detector output of a pp collision event based on theoretical frameworks such as the SM. Event generators facilitate this process by incorporating MC techniques, which use random sampling to simulate collision events from the initial interaction to the formation of stable particles and their interaction with detector material.

This chapter introduces the key aspects of a typical pp collision event at the LHC and discusses how each can be modelled using MC event simulation. These methods are then employed to simulate the various types of events used in Parts II and III of this thesis. A schematic representation illustrating the complexity of a pp collision event is provided in Figure 4.1. The following sections describe the simulation of each aspect of a pp collision including the hard-scatter calculation (Section 4.1), parton showering (Section 4.2), hadronisation (Section 4.3), the underlying event (Section 4.4), pileup (Section 4.5), and simulation of the ATLAS detector (Section 4.6).

4.1 Hard scatter

At the LHC, individual quarks are not accelerated and collided in isolation. Instead, protons, which are composite particles made up of *partons*, serve as the colliding entities. A proton in the SM consists of three valence quarks (uud) along with a surrounding *sea* of additional quarks and gluons. Each parton carries a fraction of the proton's total energy and can participate in collisions.

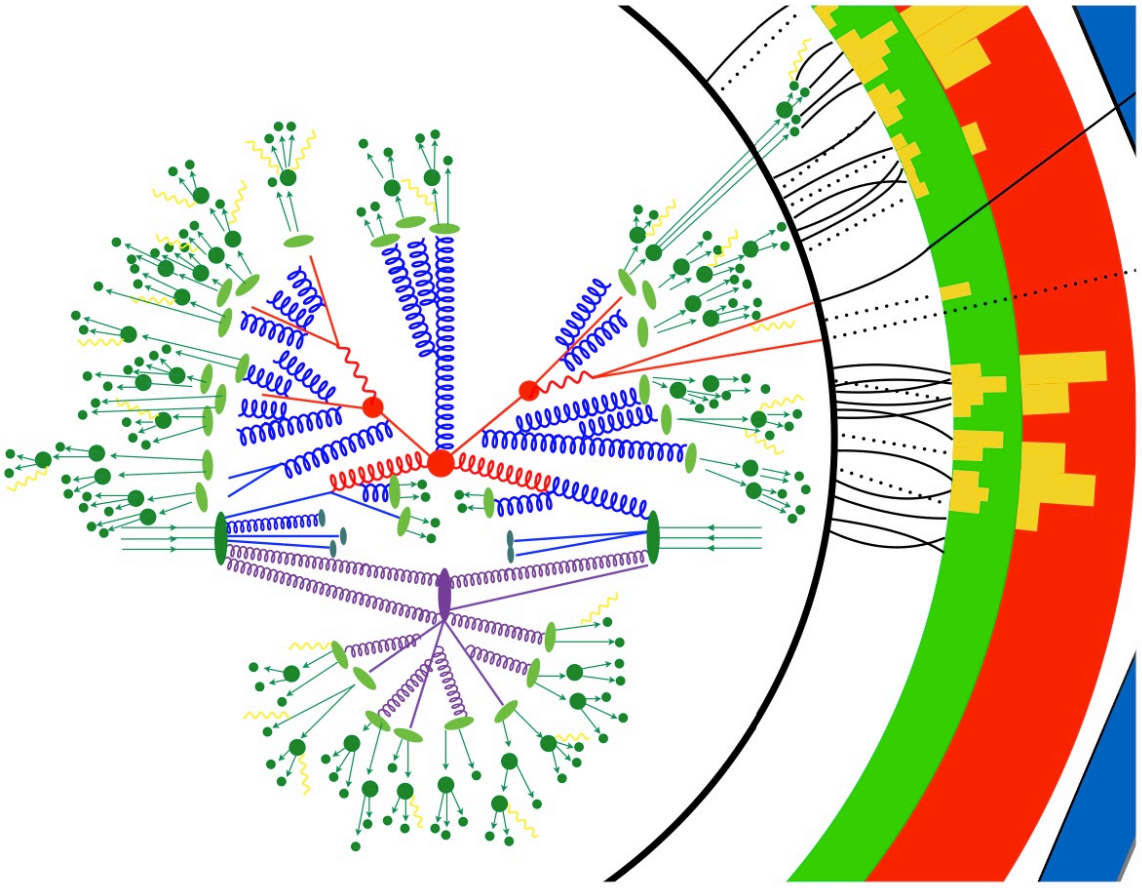


Figure 4.1: A schematic diagram illustrates the event generation of the $gg \to t\bar{t} \to q\bar{q}b\mu\nu_{\mu}\bar{b}$ process using MC methods. The dark green ovals represent the incoming protons, while the $t\bar{t}$ pair is produced in the hard-scatter process, shown in red. The parton shower, depicted in blue, occurs before the hadronisation process, which is indicated by the light green ovals. The final-state particles are represented by green circles, with photons shown in yellow. The lower section of the schematic highlights the underlying event, which arises from multiple parton interactions shown in purple, as well as the presence of proton remnants. On the right-hand side, the interactions of the final-state particles with the various layers of the ATLAS detector are depicted. This schematic has taken from Ref. [82].

An important observable to calculate for a given interaction is the cross-section, σ, which represents the probability of the interaction occurring at a given energy (see Section 3.2.2). In the context of pp collisions, calculating cross-sections requires careful consideration of the wide range of energy scales involved. The primary collision between two partons from opposing protons happens at high energy, while additional interactions between partons within each proton occur at lower energy. This presents a challenge because the strong interaction is governed by α_s, which decreases as the

energy increases. For the high-energy parton collision, α_s is sufficiently small to allow for perturbative calculations. However, for the lower-energy interactions within the proton, α_s becomes large, rendering perturbation theory inapplicable. To address this, the factorisation theorem [83] is employed, which states that the cross-section for a process can be factorised into separate high-energy and low-energy components. These components are defined using a factorisation energy scale, μ_f, which separates the perturbative and non-perturbative regimes of the process. Applying this theorem to the example of $t\bar{t}$ production ($pp \to t\bar{t}$), the cross-section is expressed as:

$$\sigma_{pp\to t\bar{t}} = \sum_{i,j\in\{g,q,\bar{q}\}} \int_0^1 dx_1 dx_2 f_i\left(x_1, \mu_\mathrm{f}^2\right) f_j\left(x_2, \mu_\mathrm{f}^2\right) \hat{\sigma}_{ij\to t\bar{t}}\left(x_1, x_2, \alpha_s, \frac{Q^2}{\mu_\mathrm{f}^2}\frac{Q^2}{\mu_\mathrm{r}^2}\right), \quad (4.1)$$

where i and j denote the species of the incoming partons, which collide with fractions x_1 and x_2 of the total momenta of the first and second protons, respectively. The different components of Equation 4.1 are now discussed.

The function $f_i\left(x_\mathrm{k}, \mu_\mathrm{f}^2\right)$ is referred to as a Parton Distribution Function (PDF), which describes the probability of finding a parton of type i carrying a momentum fraction x_k within the proton. The PDFs encapsulate the non-perturbative aspects of the interaction and depend on the choice of the factorisation scale, which is typically set according to the hard-scatter process. For $t\bar{t}$ production, the factorisation scale is commonly chosen to be proportional to the square of the top-quark mass, $\mu_\mathrm{f} \propto m_t^2$. The PDFs are not static; they evolve with energy, a dependence governed by the DGLAP equations [84–86]. However, the PDFs themselves cannot be computed analytically and must instead be extracted from global fits to experimental data [87].

The perturbative aspects of the cross-section are encapsulated in the *partonic cross section*, $\hat{\sigma}$, corresponding to the interaction between the two partons from incoming protons. The partonic cross-section for a given process is proportional to the square of the Matrix Element (ME) for that process, which is determined by calculating the *scattering amplitude* by applying Feynman rules and summing over all Feynman diagrams that contribute to the process (i.e. all diagrams with the same initial- and final-state particles). For any given process, the number of possible Feynman diagrams is infinite as it is always possible to add an arbitrary number of additional real emissions and virtual gluon self-interactions (also known as *loops*). Each additional vertex in QCD brings an extra factor of $\sqrt{\alpha_s}$ to the ME calculation. Given this, the partonic

cross-section for the $ij \to t\bar{t}$ interaction, can be written as a power series in α_s:

$$\hat{\sigma}_{ij\to t\bar{t}} = H^{(0)} + \alpha_s H^{(1)} + \alpha_s^2 H^{(2)} + \ldots, \tag{4.2}$$

where $H^{(k)}$ represents terms of increasing order in α_s. The different terms are labelled according to the order of α_s associated with their contribution. The first term is the LO contribution, the second term represents the NLO contribution, the third term denotes the Next-To-Next-To-Leading Order (NNLO) contribution, and higher-order terms follow accordingly. Figure 4.2 shows example Feynman diagrams for $t\bar{t}$ production, including higher-order contributions with the presence of an additional real emission or loop. At the high energies of the LHC, $\alpha_s < 1$ meaning that the higher-order terms associated with the diagrams exhibiting more real emissions and loops are suppressed relative to those at lower energies, resulting in a reduced impact on the partonic cross-section. For a calculation at a specific order in α_s, the power series is truncated after the term that contributes at that order. For example, the series is truncated after the third term for a cross-section calculated to NNLO in QCD. The current best predictions for the $t\bar{t}$ cross-section have NNLO accuracy in QCD [88–91], incorporating NLO EW corrections [92] and Next-To-Next-To-Leading Logarithm (NNLL) corrections in re-summed soft gluon terms [91, 93, 94].

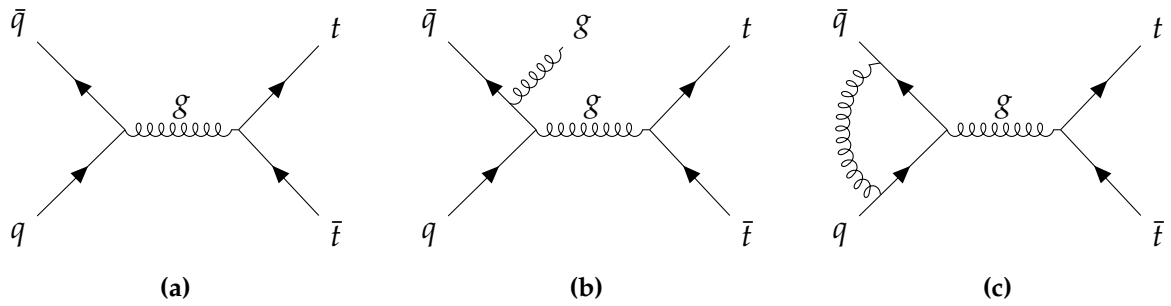


Figure 4.2: Example Feynman diagrams for $t\bar{t}$ production at (a) LO, and NLO in QCD including (b) a real emission and (c) a virtual gluon loop.

When calculating cross-sections at NLO and above, real soft (energy approaching zero) and collinear (emitted in the same direction as emitting particle) radiation contributions introduce Infrared (IR) divergences. These are handled via different methods including phase-space slicing [95, 96] or subtraction methods [97–100]. Similarly, Ultraviolet (UV) divergences arise from terms involving integrals of virtual loops with large momenta. These UV divergences are handled through the renormalisation [101] of the SM Lagrangian, which involves introducing a renormalisation scale parameter,

μ_r. The value of μ_r is typically chosen based on the energy scale of the hard-scatter process. For example, in the case of $t\bar{t}$ production, the renormalisation scale is often set proportional to the square of the top-quark mass, $\mu_r \propto m_t^2$. As shown in Equation 4.1, the partonic cross-section depends on the choices of both μ_r and μ_f, as well as the square of the transferred momentum between the two partons, denoted by Q^2.

Given the knowledge of the PDFs, the cross-section for an interaction can be predicted by a MC event generator to a fixed order in α_s. However, in practice, fixed-order calculations become increasingly complicated when accounting for additional real emissions and virtual loops. To address the computational challenges of incorporating these processes within fixed-order calculations, MC event generators use parton shower algorithms (see Section 4.2). Typically, measurements employ MC event generators that compute MEs at NLO precision, which are subsequently combined with a parton shower.

4.2 Parton shower

To compare theoretical predictions with experimental pp collision data at the LHC, all aspects of the collision must be accounted for, not just the primary hard-scatter process. Firstly, the incoming and outgoing high-energy partons can radiate many soft and collinear emissions, leading to the formation of particle showers. These subsequent emissions must be accounted for to improve the accuracy of theoretical predictions in comparison with observed data. Rather than expending significant computational resources to calculate the MEs of increasingly complex Feynman diagrams, parton shower algorithms are employed to simulate these emissions using MC methods.

Consider the splitting of a parton i into two partons j and k with momentum fractions z and $(1-z)$ with angle θ between them. The parton shower simulates how partons split in this manner using an iterative process that relies on the fact that, in both the soft and collinear regimes, the calculation of an additional emission is independent of the hard-scatter process [101]. This independence allows the parton shower to be computed separately from the hard-scatter interaction.

Emissions are ordered either by decreasing angle from the parent parton or by decreasing *virtuality*, t, which is a Lorentz invariant quantity that is closely related to the squared four-momentum. At each step, MC sampling is used to generate the type of emission and momentum fractions, using predefined splitting functions $P_{ij}(z)$. These

probability density functions encode the probabilities of a parton i splitting to into a parton j with momentum fraction z. The evolution of these splittings is governed by the *Sudakov factor*, which defines the probability that a parton evolves from an initial virtuality t_0 (or angle) to a different virtuality t (or angle) without emitting any resolvable partons with momentum fractions in the range $z_{\text{min}} \leq z \leq z_{\text{max}}$. This factor is given by:

$$\Delta_i\left(t_0, t\right) = \exp\left[-\sum_j \int_{t_0}^{t} \frac{\mathrm{d}t'}{t} \int_{z_{\text{min}}}^{z_{\text{max}}} \mathrm{d}z P_{ij}(z)\right], \tag{4.3}$$

where the summation is over the possible resulting partons j. Thus, given the starting point of the hard-scatter process, the subsequent emissions can be simulated, progressively reducing the energy scale of the system. A cut-off is imposed to terminate the shower process to prevent IR divergences, making it an essential feature of any parton shower algorithm. The cut-off is typically chosen to be around 1 GeV.

Both angle-ordered and virtuality-ordered parton shower algorithms are employed in MC event generators. Another approach is the momentum-ordered *dipole shower*, in which emissions recoil against a pair of partons [102]. Dipole showers offer an accurate treatment of soft, wide-angle radiation and correctly capture the colour flow, just like angle-ordered parton showers. However, they are known to handle partonic recoil less accurately. Conversely, angle-ordered parton showers provide a more accurate description of partonic recoil, but may misrepresent certain azimuthal correlations [103].

Parton shower algorithms are implemented in different MC event generators, with variations in shower ordering and splitting functions depending on the generator. The HERWIG [104] generator supports both angle-ordered parton showers and momentum-ordered dipole showers, while PYTHIA [105] and SHERPA [106] employ momentum-ordered dipole showers.

To properly integrate the parton shower with the hard-scatter ME calculation, it is essential to avoid double counting of emissions. This is achieved through dedicated algorithms in a procedure known as *matching*. One widely used technique at NLO accuracy is POWHEG [107], which restricts the parton shower to produce emissions that are softer than those generated by the ME. An alternative method, MC@NLO [108, 109], avoids double counting by subtracting the parton shower calculated to fixed order from the NLO ME, and then applying the shower to the final result. It is

also possible for the hard-scatter process to include a limited number of additional parton emissions incorporated directly within the ME. This enables the generator to simulate a small number of additional partons with higher-order precision. However, since these additional emissions are also described by the parton shower, any overlap between the two methods must be removed. This technique is referred to as *multijet-merging* [110, 111].

4.3 Hadronisation

After the parton shower algorithms generate additional emissions beyond the hard-scatter process, the resulting products must undergo hadronisation, forming colour-singlet states. This process occurs at energy scales comparable to the QCD confinement scale, where α_s becomes too large for perturbative methods to remain valid. Consequently, non-perturbative models must be employed, with parameters tuned to experimental data. While different hadronisation models exist, they share certain fundamental features. Firstly, the hadronisation process should not significantly alter the directions of highly energetic partons emerging from the hard-scatter interaction. Secondly, each model must determine how to arrange the colour of the final-state partons to form colour-singlet hadrons. Lastly, hadronisation models incorporate parameters that require tuning to experimental data, allowing them to make reliable predictions for future experiments.

Two primary hadronisation models are implemented in MC event generators: the *Lund string model* [112], used in PYTHIA, and the *cluster model* [113], employed by HERWIG and SHERPA.

The Lund string model represents quark-antiquark pairs as being connected by a *colour string*, whose potential energy increases as the quarks move apart. At sufficiently large distances, the energy stored in the string reaches a point where it becomes energetically favourable for the string to break, forming a new quark-antiquark pair. Due to the asymptotic freedom of QCD, the energy of the produced pairs decreases, which in turn reduces the total energy of the system. This process continues iteratively until further splitting is no longer kinematically possible. In its basic form, this mechanism only produces mesons. The production of baryons is achieved by allowing quarks to form colour strings with *diquarks* states. In this model, gluons manifest as kinks in the colour string.

While the Lund string model is grounded in the concept of asymptotic freedom in QCD, the cluster hadronisation model is based on the principle of *preconfinement* [101]. This principle states that colour-singlet clusters of radiated partons naturally emerge following the parton shower. Since parton showers evolve according to the flow of colour charge, partons with the same colour charge tend to remain close in phase space. These clusters subsequently fragment into smaller clusters, with their masses adjusted to correspond to those of known hadrons.

The decays of hadrons containing *b*- or *c*-quarks (referred to as *B*- and *D-hadrons*, respectively) are simulated by constructing each branch of the decay chain using decay amplitudes, while incorporating global angular and time-dependent correlations to enable the simulation of processes that exhibit CPV. For the measurement presented in Part III, these decays are simulated using the EvtGen program [114].

4.4 Underlying event and colour reconnection

Beyond the primary interaction, additional effects must be considered. Equation 4.1 describes the case where a single parton from one proton interacts with a single parton from another proton. However, when protons collide, Multiple Partonic Interactions (MPI) can occur and must be taken into account. In addition to this, the remnants of the protons may also interact. These contributions, along with all other particles that do not originate from the hard-scattering process simulated according to Sections 4.1–4.3, are collectively referred to as the Underlying Event (UE). The UE is simulated in MC event generators using models that are tuned to experimental data.

One crucial feature of the UE is Colour Reconnection (CR). In MC event generators, each quark is colour-connected to one other parton via a colour string, while each gluon carries both a colour and anti-colour charge and can connect to two. The presence of MPIs introduces many additional coloured partons beyond those originating from the hard-scatter, leading to a dense network of colour strings. As a result, colour strings from the hard-scatter final states may overlap in phase space with those from separate MPIs. In simple hadronisation and MPI models, colour exchange is not permitted between distinct hadron clusters or separate scatterings. A CR model is therefore introduced to resolve these ambiguities by reassigning the colour connections among partons in a physically motivated way. The default CR model in Pythia, known as the *MPI-based model* (CR0), adds the partons from MPIs of lower-p_{T} to the colour dipole

defined by the MPI systems of higher-p_{T}, in a way that minimises the total length of the colour string for each MPI system that undergoes a reconnection. An alternative model, known as the a *QCD-based model* (CR1), provides a more complete treatment of the QCD multiplet structure and enhances the production of baryons. A third model, known as the *gluon-move model* (CR2), only allows gluons to be moved to reduce the total colour string length.

Accurately modelling these complex colour reconnections is vital to avoid limiting the precision of physics analyses. For instance, previous measurements of m_t have shown that uncertainties related to CR modelling can contribute up to 0.4 GeV [115]. The role of the top quark in CR remains ambiguous. In PYTHIA, CR occurs by default before the top quark decays, meaning that its decay products do not take part in the process. As a result, the b-quarks are always colour-connected to the same partons as the parent top quark, while the W-boson decay products remain unaffected by CR. Alternatively, a configurable setting in PYTHIA allows resonance decays to occur before CR, enabling the top-quark decay products to participate in the process.

4.5 Pileup

At the LHC, protons are accelerated and collided in bunches, resulting in multiple pp collisions per bunch crossing, a phenomenon known as *pileup*. Pileup is modelled by overlaying additional hard-scatter events onto the primary interaction, incorporating all previously discussed steps. The pileup profile, defined by the distribution of the average number of collisions per bunch crossing, $\langle\mu\rangle$, is re-weighted separately for the different Run 2 data-taking periods (2015-2016, 2017, and 2018) to ensure consistency with the observed data.

4.6 ATLAS detector simulation

After the previously discussed steps, the interactions of the final-state stable particles with the detector must be simulated to allow for comparison with experimental data. This is accomplished using the GEANT4 software package [116], which models the response of the different layers of the ATLAS detector to simulated particles. MC methods are employed to track each particle's trajectory and energy deposition until the particle either loses all its energy or exits the detector. The output of this simulation

consists of recorded hits in the ID, calorimeters, and MS. These hits are then digitised into voltage and current signals, such that the same object reconstruction techniques can be applied to both simulated and real data (see Chapter 5).

Due to the large number of particles per event, this detector simulation step is highly CPU-intensive. To mitigate this computational cost, the dedicated Atlfast-II (AFII) software package [117] was developed, utilising a parametrised description of the calorimeter system.

Most of the MC samples used in this thesis to obtain the results presented in both Part II and Part III were generated using the GEANT4 detector simulation. Only a subset of samples, used specifically to evaluate certain systematic uncertainties in the m_t measurement, rely on the AFII simulation.

CHAPTER 5

Object definition and reconstruction

To conduct physics analyses using data from the ATLAS detector, the raw data must be processed to reconstruct physics objects such as leptons, hadronic jets, and the $E_{\mathrm{T}}^{\mathrm{miss}}$. This data comprises particle hit and energy deposition information, which are used to reconstruct particle trajectories and topological clusters. A Particle Flow (PFlow) algorithm then combines track and cluster information to construct objects that serve as input to jet clustering algorithms. This chapter details the reconstruction of particle trajectories within the ATLAS detector, along with the reconstruction and calibration procedures for physics objects used in analysis.

5.1 Tracks

As charged particles traverse the various layers of the ATLAS detector, they interact with the detector material, producing spatial hits and energy deposition data. The strong magnetic fields induce helical trajectories, enabling the determination of particle charge and momentum. These trajectories are reconstructed from spatial hit information to form *tracks* in the ID and MS.

5.1.1 Inner detector tracks

The first step in constructing a track object in the ID is the generation of *space-points* [118]. These three-dimensional measurements are formed by clustering hits in the Pixel and SCT sub-detectors. *Clusters* are created by grouping adjacent pixels or strips in a sensor that share a common edge or corner, provided the deposited energy

exceeds a predefined threshold. In the Pixel detector, each cluster corresponds to a single space-point, whereas in the SCT, clusters from both sides of a strip layer are required to define a space-point.

Before assembling a track candidate, triplets of space-points within the same sub-detector are grouped to form *track seeds*. This approach provides an initial estimate of the momentum and impact parameters while maximising the number of possible track combinations. The impact parameters consist of the *transverse impact parameter* (d_0), defined as the closest approach of the track to the beamline in the transverse plane, and the *longitudinal impact parameter* (z_0), which measures the displacement along the z-axis from the primary interaction vertex (see Section 5.2). The *purity* of these seeds, defined as the fraction that correspond to a true particle trajectory, is improved by requiring an additional space-point to be compatible with the track seed's trajectory.

Track candidates are then constructed through an iterative fitting process, starting from the set of track seeds. A combinatorial Kalman filter [119] extrapolates each seed's trajectory through the remaining Pixel and SCT sub-detector layers, incorporating additional space-points that align with the predicted path. This filter propagates the initial track parameters outward, estimating the location of the next space-point with an uncertainty that accounts for multiple scattering and energy loss due to interactions with the detector material. The trajectory is then updated using a weighted mean of the predicted and measured positions of the next space-point, factoring in the detector resolution. This process continues through successive layers, refining the track parameters at each step. If multiple space-points in a given layer are compatible with the track seed, multiple track candidates are generated, leading to the idea that this is a "combinatorial" filter exploring different possible trajectories. This approach is crucial for the ATLAS detector, where additional interactions from pileup events introduce extraneous space-points. Given the large number of potential combinations, the resulting set of track candidates includes some with overlapping or incorrectly assigned space-points. These ambiguities are addressed in a subsequent *ambiguity-solving* step.

In the ambiguity-solving stage, a score is assigned to each track candidate, where a higher score indicates a higher probability of representing a true charged particle trajectory. The score is influenced by the number of assigned space-points, the number of layers in the Pixel and SCT sub-detector layers missing compatible space-points (known as *holes*), and the χ^2 value of the fitted track. Track momentum is also considered to suppress reconstructed tracks with incorrectly assigned clusters, typically of

lower p_{T}. Track candidates are then processed in descending order of score, deciding which tracks to retain, favouring those with higher scores. A neural network [120] assists in distinguishing cases where shared space-points correspond to merged clusters from multiple particles (in which case both tracks retain the hits) versus shared clusters that result from incorrect assignments (where the track with the higher score is favoured).

During a precision tracking stage [121], track candidates are extended into the TRT by matching them with compatible hits to improve momentum resolution.

Once these steps are completed, the remaining track candidates must satisfy the criteria outlined in Table 5.1. Tracks meeting these requirements undergo a high-precision fit [121], further refining the final track parameters.

Table 5.1: The requirements that ID tracks must satisfy after the ambiguity-solving step. Here, θ denotes the track's polar angle.

Quantity	Requirement
p_{T}	$p_T > 400$ MeV
η	$\lvert\eta\rvert < 2.5$
Minimum number of clusters	7 between pixel and SCT layers
Number of holes	≤ 2 overall with ≤ 1 in the pixel layers
d_0	$\lvert d_0\rvert < 2$ mm
z_0	$\lvert z_0 \sin(\theta)\rvert < 3$ mm

5.1.2 Muon spectrometer tracks

Tracks in the MS are constructed [122] using the same fundamental principle of connecting space-points to form track candidates. However, due to the more complex toroidal magnetic field, different fitting methods are employed. In each layer of the MDT chambers, a Hough transform [123] is used to identify short, straight-line local track segments from the recorded hits. This process entails mapping hit coordinates in the detector space into a parameter space, where collinear points belonging to a track form a peak. These measurements in the bending plane of the detector are complemented by hits in the orthogonal (r, ϕ) plane, provided by the RPC or TGC sub-

detectors [122]. In the CSC, track segments are reconstructed using a combinatorial search in the η and ϕ detector planes, with a loose requirement on track compatibility with the luminous region.

Next, a segment-seeded combinatorial search algorithm is used to build MS tracks from the identified track segments in different layers. This process begins with segments in the middle layers, where more trigger hits are available, and extends outward to the inner and outer layers. Additional segments are selected based on their relative positions and angles, as well as their hit multiplicity and fit quality. A track candidate requires at least two matching track segments, except in the transition region between the barrel and end-caps ($1.0 < |\eta| < 1.4$), where a single high-quality segment is sufficient.

As in ID track reconstruction, a single seed can generate multiple track candidates. To address this, an overlap removal algorithm is applied to optimally assign segments to individual tracks while allowing limited segment sharing between two tracks when necessary. In cases where two muons are close in the detector and share identical track segments, both tracks are retained if they include segments from at least three MS layers, provided that they share segments in the first two layers but not in the outermost layer.

5.2 Vertices

In the detector, the intersection of multiple tracks is referred to as a *vertex*, representing either the interaction of multiple particles or the decay of a single particle. Vertices are reconstructed by extrapolating ID tracks backward to determine their intersection points. The *primary vertex* is assumed to correspond to the hard-scatter process of the event and is identified as the vertex with the highest sum of the squared transverse momenta of all associated tracks, $\sum {p_{\mathrm{T}}}^2$. Other reconstructed vertices, known as *secondary vertices*, arise from either the decay of long-lived particles, such as B-hadrons, or pileup interactions.

5.3 Topological clusters

In addition to producing hits in the ID and MS for track reconstruction, particles also deposit energy in the calorimeter cells. While track objects in the ID and MS are built from three-dimensional space-points, topological clusters (known as *topo-clusters*) in the calorimeters are constructed by grouping energy deposits in individual cells into three-dimensional clusters [124]. This process enables the measurement of both the energy and direction of incident particles. However, topo-clusters may only account for a fraction of the energy deposited or conversely include contributions from multiple particles.

Topo-clusters are constructed using an algorithm that connects neighbouring calorimeter cells based on their *signal significance*. A cell's signal significance, $\zeta_{\text{cell}}^{\text{EM}}$, is defined as the ratio of its energy deposit, $E_{\text{cell}}^{\text{EM}}$, to the expected average noise, $\sigma_{\text{noise,cell}}^{\text{EM}}$, which is estimated separately for each data-taking year using MC simulations of events triggered using a zero-bias trigger, as discussed in [124]:

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}. \tag{5.1}$$

Both $E_{\text{cell}}^{\text{EM}}$ and $\sigma_{\text{noise,cell}}^{\text{EM}}$ are measured on the EM scale, ensuring accurate reconstruction of energy depositions from electrons and photons.

The clustering algorithm first selects all cells with $\zeta_{\text{cell}}^{\text{EM}} > 4$. For each selected seed cell, neighbouring cells with $\zeta_{\text{cell}}^{\text{EM}} > 2$ are grouped to form a *proto-cluster*. Neighbouring cells are defined as those adjacent within the same layer or overlapping in the (η,ϕ) plane across adjacent layers. If a neighbouring cell also has $\zeta_{\text{cell}}^{\text{EM}} > 4$, making it a seed for another proto-cluster, the two proto-clusters are merged. If a neighbouring cell belongs to another proto-cluster and has $\zeta_{\text{cell}}^{\text{EM}} > 2$, the proto-clusters are also merged. This process iterates until no remaining neighbours satisfy $\zeta_{\text{cell}}^{\text{EM}} > 2$. Finally, additional neighbouring cells with $\zeta_{\text{cell}}^{\text{EM}} > 0$ are included, ensuring that the proto-cluster consists of a core of high-significance cells surrounded by lower-significance cells. This clustering configuration maintains a dynamic noise suppression by allowing signals close to the noise threshold to be included while ensuring a higher noise threshold, thereby improving calorimeter performance.

During topo-cluster formation, spatial signal structures are not explicitly considered. As a result, local maxima within a cluster may indicate multiple particles

depositing energy in close proximity, leading to topo-clusters that are too large. To mitigate this, proto-clusters are split based on local maxima, defined by $E_{\text{cell}}^{\text{EM}} > 500$ MeV. A local maximum must have at least four neighbouring cells, each with lower energy. The clustering algorithm is then rerun, beginning with pairs of maxima, and the energy of shared cells is distributed based on the relative energy of each maximum and the distance to the centre of mass of each resulting cluster. The final set of topo-clusters are then used for the reconstruction of other physics objects, including leptons (see Section 5.4) and jets (see Section 5.5).

5.4 Leptons

Leptons are reconstructed using either tracks or a combination of tracks and topo-clusters. This discussion focuses on the reconstruction and calibration of electrons and muons, as tau leptons are not considered in the work presented in Part II or the analysis in Part III. However, the results may include events where electrons and muons originate from tau lepton decays.

5.4.1 Electrons

Electron reconstruction

Electron candidates are reconstructed using *superclusters*, which incorporate energy from low-energy Bremsstrahlung photons [125]. Superclusters are formed from the highest-p_{T} topo-clusters with $p_{\text{T}} > 1$ GeV that are matched to a track in the ID containing at least four space-points. An electron candidate is defined as a supercluster with an associated track, with its energy determined from the cluster's energy.

Electron candidates must have a transverse energy of $E_T > 27$ GeV and be reconstructed within the active calorimeter regions ($|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$). To suppress pileup effects, the ratio of the transverse impact parameter to its uncertainty must satisfy $|d_0|/\sigma_{d_0} < 5$, and the longitudinal impact parameter relative to the primary vertex must satisfy $|\Delta z_0 \sin(\theta)| < 5$ mm.

To ensure electrons are distinguishable from other physics objects, such as low-energy jets (see Section 5.5), candidates must satisfy identification criteria based on

a multivariate likelihood method [125, 126]. This method incorporates track and calorimeter information to define different selection working points. The analysis in Part III employs the *Tight* working point, which achieves an average selection efficiency of 80% with a background rejection factor of 3.5 [125].[1] Additionally, in Section 8.1, a sample with looser electron identification requirements is used to estimate the multijet background contribution for the m_t measurement discussed in Part III. This sample applies the *Medium* working point, corresponding to an 88% selection efficiency and a background rejection factor of 2 [125], and does not impose isolation criteria, which are discussed next.

Electron candidates must also be sufficiently isolated from other physics objects to improve the accuracy of their reconstruction. This is achieved by evaluating the total transverse momentum of surrounding tracks within a variable cone of maximum size $\Delta R = 0.2$ in the ID, ${p_{\mathrm{T}}}^{\mathrm{cone20}}$, and the total surrounding transverse energy within a cone of $\Delta R = 0.2$ in the calorimeters, E_T^{cone20}. The analysis in Part III applies the *Tight* isolation working point, requiring ${p_{\mathrm{T}}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.06$ and $E_T^{\mathrm{cone20}}/E_T < 0.06$.

Electron calibration

Reconstructed electron candidates must be calibrated to correct the electron energy scale in data, and the electron energy resolution in simulation. This is performed by comparing $Z \to e^+e^-$ events in simulation and data [127]. Additionally, it is important to account for efficiency differences between data and simulation. This calibration is applied through scale factors that correct the simulation to better match data. These scale factors are derived by comparing reconstruction, identification, and isolation efficiencies in $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ events between data and simulation [125].

Electron uncertainties

Since the reconstruction, identification, and isolation efficiencies directly affect electron candidate reconstruction, they also influence physics analyses that use these objects. To account for these effects in a physics analysis, systematic uncertainties from the efficiency measurements are propagated to the analysis [125, 128]. Systematic uncer-

[1] A background rejection factor is given by the inverse of the background efficiency. Therefore, a background rejection factor of 3.5 means that approximately 28.6% of background events pass the criteria.

tainties associated with the electron energy scale and resolution that arise from pileup modelling, calorimeter layer inter-calibration, and discrepancies between data and simulation, are propagated to the energy scale and resolution corrections [127].

5.4.2 Muons

Muon reconstruction

Muon candidates are reconstructed by matching tracks in the ID with those in the MS [129]. For the measurement discussed in Part III, two types of muon candidates are identified: *combined* and *inside-out*. Combined muons are constructed by performing a track fit that incorporates hits from both the ID and MS while accounting for energy loss in the calorimeters. Inside-out muons, on the other hand, are reconstructed by extrapolating ID tracks into the MS and searching for at least three loosely aligned hits. These hits, together with the ID track and calorimeter energy loss, are then used in a combined track fit to form the final inside-out muon candidate.

Muon candidates must have $p_{\mathrm{T}} > 27$ GeV and $|\eta| < 2.5$. Just as for electrons, muon candidates are required to satisfy criteria ensuring association to the primary vertex, reducing the impact of pileup. These are $|d_0|/\sigma_{d_0} < 3$ and $|\Delta z_0 \sin(\theta)| < 0.5$ mm.

Similar to electron candidates, muon candidates must satisfy identification criteria [129] to ensure they are sufficiently distinguishable from other physics objects, such as pions and kaons. Different selection working points are defined based on different requirements related to the number of hits in various ID and MS subdetectors, track fit properties, and the compatibility of ID and MS measurements. The measurement discussed in Part III uses the *Medium* working point for $t\bar{t}$ signal events, requiring at least three hits in at least two layers of the MDT or the CSC. This requirement is relaxed in the region where $|\eta| < 0.1$, where muons only need at least three hits in one MDT layer, provided they do not have more than one layer with fewer than three hits where more hits would be expected given the muon's trajectory. A requirement is also placed on the agreement between ID and MS measurements, defined using the charge-to-momentum compatibility. This is determined by the absolute difference between the measured charge-to-momentum ratio in the ID and MS, divided by the sum in quadrature of the associated uncertainties in both measurements. For the *Medium* working point, this compatibility must be less than seven.

Muon candidates must also be sufficiently isolated from other physics objects. Muon isolation is determined using both track-based and calorimeter-based criteria, similar to the isolation methods applied to electrons [129]. Several working points are defined for muon isolation, with the *Tight* working point used for the signal $t\bar{t}$ events in the measurement discussed in Part III. This criterion requires muon candidates to have tracks with $p_{\mathrm{T}} > 1$ GeV, ${p_{\mathrm{T}}}^{\mathrm{varcone30}}/p_{\mathrm{T}} < 0.04$, and $E_T^{\mathrm{topocone20}}/p_{\mathrm{T}} < 0.15$. Here, ${p_{\mathrm{T}}}^{\mathrm{varcone30}}$ represents the scalar sum of the transverse momenta of the ID tracks associated with the primary vertex within a cone of radius $\Delta R = \min(10\,\mathrm{GeV}/p_{\mathrm{T}}, 0.3)$ around the muon. Similarly, $E_T^{\mathrm{topocone20}}$ denotes the sum of the transverse energy in topo-clusters within a cone of radius $\Delta R = 0.2$ around the muon, with contributions from the muon itself and pileup effects subtracted.

As mentioned in Section 5.4.1, a sample with looser identification requirements is used to estimate the multijet background contribution for the m_t measurement discussed in Part III. This sample applies the same *Medium* working point for muon identification as the $t\bar{t}$ sample. Like the electron isolation requirements in this sample, no muon isolation requirements are applied.

Muon calibration

As with electron candidates, muon candidates require calibration to correct the simulation for discrepancies in the muon momentum scale and resolution when compared to data. This calibration is achieved using resonance decays from $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ events [130]. In addition, procedures are implemented to correct the data for charge-dependent biases in the momentum measurement that arise from residual uncertainties in the MS chamber alignment, which persist despite dedicated alignment efforts, as described in Ref. [130]. To address efficiency differences between data and simulation, calibration scale factors are applied. These scale factors are derived from observed differences in reconstruction, identification, isolation, and vertex association efficiencies between data and simulation, using $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ events [130].

Muon uncertainties

Variations in the scale factors associated with identification, isolation, trigger, and vertex association corrections are applied as systematic uncertainties in the reconstruction

of muon candidates [129, 131]. The muon momentum scale and resolution corrections are sensitive to various systematic uncertainties [130], including resonance decay modelling, Final-State Radiation (FSR) modelling, the range of p_{T} used to construct the Z-boson and J/ψ templates, the Z-boson and J/ψ mass windows, template fit stability, and MS alignment. Additionally, systematic uncertainties are included to account for the accuracy of the sagitta measurement in the MS.

5.5 Jets

Unlike leptons, quarks and gluons cannot be reconstructed as single objects. Due to colour confinement in QCD, only colour-neutral states can exist. Consequently, quarks and gluons produce sprays of colour-singlet states through the process of hadronisation, which are captured using objects known as *jets*. These jets are defined by a jet clustering algorithm and are constructed from PFlow objects, which combine track information with topo-clusters [132].

To ensure consistency between data and simulation, jets are calibrated to account for the *jet response* of the calorimeters, defined as the ratio of the jet p_{T} after detector simulation to that before detector simulation, which is less than one. This calibration corrects for detector effects and contributions from pileup. A subsequent calibration step is then applied to bring the data and simulation into agreement to preserve an accurate energy scale and resolution.

The high-energy pp collisions at the LHC result in final states that are typically dominated by jets, which play a crucial role in defining and reconstructing the signal process, describing background processes, and accounting for additional pileup interactions. As a result, precision measurements of the SM involving jets are sensitive to their reconstruction, the calibration of their energy and resolution, and their associated systematic uncertainties. In particular, the m_t measurement discussed in Part III relies heavily on jets, with uncertainties in the calibration of their energy scale being a factor that limits the measurement's precision.

5.5.1 Jet reconstruction

Particle flow objects

A PFlow algorithm [132] is used to define objects to be used for jet reconstruction, which incorporates both topo-cluster and ID track measurements. These *PFlow objects* consist of energy-subtracted topo-clusters with matched tracks representing charged particles, while topo-clusters without track associations represent neutral particles. By combining tracking information with topo-clusters, jet reconstruction benefits from the better momentum resolution of the ID for low-energy particles compared to the energy resolution of the calorimeters, the improved angular resolution of the ID, and the suppression of pileup effects through vertex association. Since the calorimeters provide significantly better resolution than the ID for high-energy particles, a combination of topo-clusters and track information is used to construct PFlow objects, rather than relying solely on one or the other.

Before creating PFlow objects, an event-by-event origin correction is applied to each topo-cluster to account for the position of the primary vertex [133]. This correction improves the η resolution of the reconstructed jets by more accurately associating their origin with the primary vertex.

When combining topo-cluster and tracking information to construct PFlow objects, it is essential to avoid double-counting energy. To address this, the energy contributions of topo-clusters corresponding to matched tracks must be removed from the calorimeter cell-by-cell. The PFlow algorithm [132] is summarised in Figure 5.1 and follows these steps:

1. Select well-measured tracks that have at least nine hits across the Pixel and SCT sub-detectors, with no holes in the layers of the Pixel sub-detector.

2. Match each track with momentum p^{trk} to a single topo-cluster with energy E^{clus} such that $E^{\text{clus}}/p^{\text{trk}} < 0.1$, and minimise the angular distance $\Delta R'$, defined as:

$$\Delta R' = \sqrt{\left(\frac{\Delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\Delta\eta}{\sigma_\eta}\right)^2}, \tag{5.2}$$

where $\Delta\phi$ and $\Delta\eta$ are the distances between the centre-of-mass of the topo-cluster and the track extrapolated to the calorimeter, and σ_ϕ and σ_η represent the widths of the topo-cluster in ϕ and η.

3. For a given track, use the topo-cluster position and track momentum to calculate the expected energy deposit in the calorimeter, given by:

$$\langle E_{\text{dep}} \rangle = p^{\text{trk}} \langle E_{\text{ref}}^{\text{clus}} / p_{\text{ref}}^{\text{trk}} \rangle, \tag{5.3}$$

for a track with momentum p^{trk} where $\langle E_{\text{ref}}^{\text{clus}} / p_{\text{ref}}^{\text{trk}} \rangle$ represents the expectation determined from simulation.

4. Determine whether additional topo-clusters are needed to recover the total expected energy using a discriminant that distinguishes between single and multiple topo-cluster scenarios, defined as:

$$S(E^{\text{clus}}) = \frac{E^{\text{clus}} - \langle E_{\text{dep}} \rangle}{\sigma(E_{\text{dep}})}, \tag{5.4}$$

where $\sigma(E_{\text{dep}})$ is the width of the expected energy deposition, corresponding to the standard deviation of the $E_{\text{ref}}^{\text{clus}} / p_{\text{ref}}^{\text{trk}}$ distribution in simulation. Additional topo-clusters are required if $S(E^{\text{clus}}) < -1$.

5. The expected energy deposited in the calorimeter is subtracted cell-by-cell from the set of matched topo-clusters. This is achieved by forming rings in η and ϕ around the extrapolated track in each calorimeter layer of the topo-clusters. These rings are evenly spaced in radius and are designed to always contain at least one calorimeter cell. The rings are ranked in descending order in energy density. Starting with the ring of highest energy density, rings are iteratively removed if their total energy is less than the amount needed to reach the required $\langle E_{\text{dep}} \rangle$, removing energy from the system at each iteration. This process continues until the current ring contains more energy than what remains to be subtracted, at which point the energy of each cell in that ring is scaled down to exactly match $\langle E_{\text{dep}} \rangle$.

6. Remove any remaining topo-cluster energy if it is consistent with the expected shower fluctuations of a single particle's signal. This condition is met if the total energy of the remaining cells after subtraction is less than $1.5\sigma(E_{\text{dep}})$, which assumes the system was produced by a single particle and that the residual

energy arises from shower fluctuations. If this criterion is not satisfied, the remnant topo-cluster energy is retained, as it is likely to originate from multiple particle deposits.

After the final step, any topo-clusters that are not matched to a track and were not modified during the subtraction process are identified as energy deposits from neutral particles and are retained. Any topo-clusters matched to tracks that are inconsistent with the primary vertex are removed, along with the associated tracks. The remaining tracks and topo-clusters, now free from double-counted energy, are used for jet reconstruction via a jet clustering algorithm.

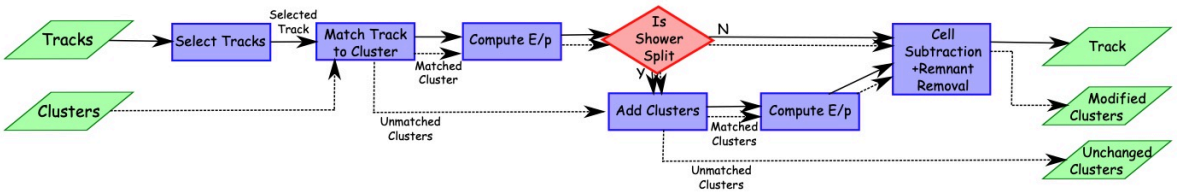


Figure 5.1: A flow chart for the PFlow algorithm, taken from Ref. [132]. The algorithm begins with a selection of tracks and topo-clusters and concludes with a set of tracks associated with modified clusters representing charged particle information, along with a set of unchanged clusters that do not have associated tracks representing the information of neutral particles.

The anti-k_t jet clustering algorithm

Jet clustering algorithms determine how particles are combined to form jets. Various algorithms exist [134], but the measurement in Part III employs the anti-k_t algorithm [135]. This algorithm utilises sequential recombination to merge nearby PFlow objects, i and j, based on the distance parameter d_{ij}, defined as:

$$d_{ij} = \min\left(k_{t,i}^{-2}, k_{t,j}^{-2}\right) \frac{\Delta R_{ij}^2}{R^2}, \tag{5.5}$$

where $\Delta R_{ij} = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, and R is a variable radius parameter. Here, $k_{t,i}$, y_i, and ϕ_i represent the transverse momentum, rapidity, and azimuthal angle of object i, respectively. Additionally, the algorithm defines the distance parameter d_{iB}, which represents the distance between object i and the beamline:

$$d_{iB} = k_{t,i}^{-2}. \tag{5.6}$$

For each pair of input objects i and j, both d_{ij} and d_{iB} are calculated. If $d_{ij} < d_{iB}$, objects i and j are clustered together to form a *pseudo-jet*. This process repeats iteratively until $d_{ij} > d_{iB}$, at which point there are no remaining unclustered objects within a distance R from object i. At this stage, the pesudo-jet is identified as a jet candidate and removed from the remaining set of objects. The procedure continues with the remaining objects until all PFlow objects have been processed, resulting in a final set of jet candidates.

The anti-k_t algorithm is designed to be Infrared and Collinear (IRC) safe, meaning that its reconstruction remains invariant under the introduction of additional soft radiation or collinear radiation emitted along the same axis as the jet [135]. Although other IRC-safe algorithms exist [134], the reconstruction dynamics of the anti-k_t algorithm ensure that it produces jets with a regular circular shape. This characteristic simplifies energy calibration and the consideration of systematic uncertainties. Additionally, the anti-k_t algorithm ensures that the hardest particles (those with the largest p_{T}) act as seeds for the reconstruction, around which softer particles cluster. As a result, the produced jets are less sensitive to pileup and UE effects [135].

The measurement in Part III utilises jets formed using the anti-k_t algorithm with $R = 0.4$ referred to as *small-radius jets*, or *small-R jets*. These jets are calibrated using the procedures outlined in Section 5.5.2 and are required to be within the region $|\eta| < 2.5$ of the detector and have $p_{\mathrm{T}} > 26$ GeV. Jets originating from pileup are removed by applying a requirement on the output of the multivariate Jet Vertex Tagger (JVT) for jets with $p_{\mathrm{T}} < 60$ GeV and $|\eta| < 2.4$ [136]. To avoid double-counting energy contributions from leptons, overlap criteria are employed. Jets within $\Delta R < 0.2$ of an electron are removed, as are jets within $\Delta R < 0.2$ of a muon if they have fewer than three tracks or if any track is part of the muon candidate. After these requirements, jets and leptons are ensured to be well separated by removing any lepton within $\Delta R < 0.4$ of a jet.

5.5.2 Jet calibration

Just as lepton candidates must be calibrated before being used in physics analyses, jets also require calibration. This process involves a series of simulation-based corrections to remove pileup and detector effects, followed by an *in-situ* calibration to account for differences between data and simulation [133].

Jet energy scale

The energy scale of the small-R jets must be calibrated to ensure that the reconstructed jet energy matches the expected energy of the jet before simulating the detector response, where this jet is known as the *truth jet*. The Jet Energy Scale (JES) calibration [133] consists of a sequence of simulation-based corrections applied to both data and simulation followed by an in situ correction applied exclusively to data. These corrections align the reconstructed jet four-momenta with jets simulated using the PYTHIA8 generator [105]. However, the jet response varies among different MC generators used for systematic uncertainty evaluation where comparisons between alternative generators will lead to the double-counting of detector response effects. To address this, a final MC-to-MC scale factor calibration is applied to alternative generators to match the jet response in PYTHIA8. The complete JES calibration chain is illustrated in Figure 5.2, where each step corrects the four-momentum of the reconstructed jets, and is now discussed in turn.

The first step in the calibration chain involves a pileup correction. Since pileup contributes additional energy throughout the calorimeter, its impact can be estimated on an event-by-event basis by examining the energy of all reconstructed jets. The correction is implemented by subtracting the median jet p_{T} density, ρ, calculated for each event [133]. A jet's susceptibility to pileup is related to its area, A, with larger jets being more affected. The jet area is determined by assessing the relative number of *ghost* particles associated with the jet after clustering. These ghost particles are artificial soft particles that are uniformly distributed in (η, ϕ) space and are added to the event before jet reconstruction. The median p_{T} density in the y-ϕ plane is then computed as $\rho = \langle p_{\mathrm{T}}/A \rangle$ using jets reconstructed via the k_t jet clustering algorithm [137] within the central region ($|\eta| < 2$). The k_t jet clustering algorithm is used as it naturally reconstructs jets with a uniform soft background [138]. Although the correction is derived using central jets, it is applied to all jets by using the ratio of the ρ subtracted jet p_{T} to the original jet p_{T}, as discussed in Ref. [133]. This modifies the jet's four-momentum while preserving its direction.

Following this correction, some residual dependence on pileup remains, particularly in the forward region ($|\eta| > 2$) where the lower detector occupancy compared to the central region causes ρ to drop to zero. This is addressed in the second stage of the calibration chain, which corrects for the effects of additional pp interactions occurring within the same bunch crossing or in nearby bunch crossings, referred to as *in-time*

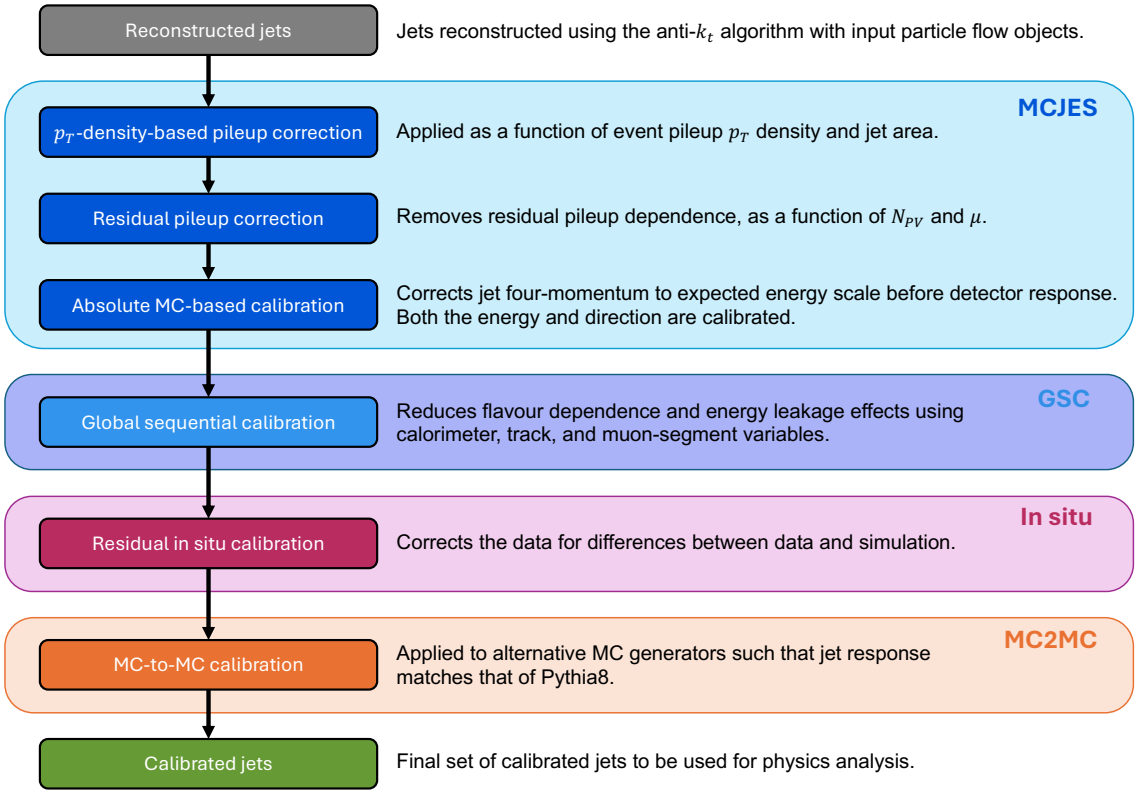


Figure 5.2: The full chain of the JES calibration, where each stage corrects the four-momentum of the reconstructed jet. No MC-to-MC calibration is applied to jets reconstructed using the PYTHIA8 generator.

and *out-of-time* pileup, respectively. The number of primary vertices in the event, N_{PV}, is sensitive to in-time pileup, whilst the number of interactions per bunch crossing, μ, is sensitive to out-of-time pileup. Therefore, a residual correction is applied based on the jet response, parametrised in terms of N_{PV} and μ. The pile-up corrected jet p_{T}, after both the ρ-subtraction based and residual corrections, is given by:

$$p_{\mathrm{T}}^{\mathrm{corr}} = p_{\mathrm{T}}^{\mathrm{reco}} - \rho \times A - \alpha \times (N_{\mathrm{PV}} - 1) - \beta \times \mu, \tag{5.7}$$

where $p_{\mathrm{T}}^{\mathrm{reco}}$ is the jet transverse momentum before any correction is applied, and α and β are coefficients determined from independent linear fits in bins of $p_{\mathrm{T}}^{\mathrm{true}}$ and $|\eta_{\mathrm{det}}|$. Here, $p_{\mathrm{T}}^{\mathrm{true}}$ denotes the transverse momentum of the geometrically matched truth jet within $\Delta R = 0.3$, and η_{det} refers to the jet pseudorapidity measured relative to the geometric centre of the detector.

The transition between different calorimeter technologies and granularities is a primary source of bias in the reconstruction of jet η. Additionally, jet energy and η are affected by variations in the reconstructed energy scale for jets that deposit energy in both the EM and hadronic calorimeters, energy losses in passive material, and the effect of energy not being fully reconstructed within the calorimeter system, known as *leakage*. The next step in the calibration chain corrects for these effects by calibrating both the jet energy and direction. The jet energy is corrected by applying a scale factor equal to the inverse of the average jet energy response, $1/\mathcal{R}$. The average response, $\mathcal{R}$, is defined as the mean of a Gaussian fit to the central region of the $E^{\text{reco}}/E^{\text{true}}$ distribution, measured in bins of E^{true} and η_{det}, where E^{reco} is the reconstructed jet energy and E^{true} is the energy of the geometrically matched truth jet within $\Delta R = 0.3$. The response is calculated as a function of E^{true} rather than E^{reco} because for fixed E^{true} the distribution of the response is Gaussian, whereas this is not the case for fixed E^{reco}. The average response is then parametrised in terms of E^{reco} using a numerical inversion method, as detailed in Ref. [139]. Analogous to this jet energy correction, the reconstructed jet η is corrected by applying a scale factor of $(\eta^{\text{reco}} - \eta^{\text{true}})$, also parametrised in terms of E^{reco} and η_{det}. This absolute calibration, together with the pileup-based corrections, is collectively referred to as *MCJES*.

Following this, the jet response can still depend on the flavour and energy of the constituent particles, as well as the jet shape. The particle composition and shape of jets vary depending on their initiating particles, particularly between quark- and gluon-initiated jets. Quark-initiated jets tend to contain more high-p_{T} hadrons that travel far into the calorimeter, while gluon-initiated jets typically have more low-p_{T} particles, resulting in wider showers in the calorimeter. The next step in the calibration chain is the Global Sequential Calibration (GSC), which applies a series of multiplicative corrections to six observables in sequence to mitigate these effects and improve jet resolution. The six corrected variables, in order, are:

- f_{charged} - the fraction of jet p_{T} measured from charged particles;
- f_{Tile0} - the fraction of jet energy measured in the first layer of the hadronic tile calorimeter;
- f_{LAr3} - the fraction of jet energy measured in the third layer of the EM liquid argon calorimeter;
- n_{trk} - the number of tracks in the jet with $p_{\text{T}} > 1$ GeV;

- w_{trk} - the average p_{T}-weighted transverse distance in the $\eta - \phi$ plane between the jet axis and all tracks with $p_{\text{T}} > 1$ GeV, also known as the *track width*;
- n_{segments} - the number of muon track segments within the jet, also known as *punch-through*.

A correction is derived for each observable using the same method as MCJES, based on the inverse of the reconstructed jet p_{T} response, parametrised in terms of the truth jet p_{T} and $|\eta_{\text{det}}|$. The jet p_{T} response is used instead of the jet energy response, as it leads to improved jet resolution without altering the average jet energy response. The only exception is the n_{segments} correction, which is derived in terms of jet energy instead of p_{T}, as punch-through is associated with energy escaping the calorimeters. The order in which these corrections are applied does not affect the resolution of the average jet energy response.

Following MCJES and the GSC, the in situ calibration step is applied to data to correct for residual differences in jet response between data and simulation. This calibration adjusts the jet four-momenta based on the double ratio of the average jet response measured in data and simulation, defined as:

$$c = \frac{\mathcal{R}^{\text{data}}_{\text{in situ}}}{\mathcal{R}^{\text{MC}}_{\text{in situ}}}, \tag{5.8}$$

where the correction factor, c, is parametrised in terms of jet p_{T} and η. The in situ calibration is derived from three methods that determine jet response by balancing the p_{T} of a jet against that of a well-calibrated reference object. These methods are the η-intercalibration method, the Z/γ + jet Missing-E_{T} Projection Fraction (MPF) method, and the Multijet Balance (MJB) method, and are described below:

- The η-intercalibration method balances the energy scale of forward ($0.8 < |\eta| < 4.5$) jets with that of central ($|\eta| < 0.8$) jets by balancing p_{T} in dijet events.
- The Z/γ + jet MPF method balances the full hadronic recoil of an event against the p_{T} of a well-calibrated photon or a leptonically decaying Z boson ($Z \to e^+e^-$ or $Z \to \mu^+\mu^-$). Instead of using a single jet, this method relies on the full hadronic recoil, which is less sensitive to pileup, UE effects, and jet definitions. This technique applies only to central jets, but in combination with η-intercalibration, it provides a smooth calibration across a jet p_{T} range of 17 GeV $< p_{\text{T}} <$ 1.2 TeV.

- The MJB method calibrates a single high-p_{T} jet by using a system of well-calibrated low-p_{T} jets, extending the jet p_{T} calibration range up to 2.4 TeV. Jets with $p_{\mathrm{T}} > 2.4$ TeV are calibrated using studies of jet response to single high-energy hadrons [140, 141].

The final step in the JES calibration chain is the application of MC-to-MC scale factor corrections to alternative MC generators, ensuring that their jet response matches that of PYTHIA8 [142]. This process, known as *MC2MC*, prevents double-counting uncertainties related to the modelling of the detector response to jets, as these are already accounted for in the JES uncertainties. Additionally, these calibrations are necessary for physics analyses to incorporate the latest definition for the JES flavour dependence uncertainties (see Section 5.5.3). The MC2MC calibration is derived from the ratio of the jet response in the alternative generator to that in PYTHIA8, following a similar approach to the in situ calibration used to correct jet p_{T} in data.

Jet energy resolution

Precise determinations of SM parameters in processes involving jets, such as m_t extractions from top-quark decays to jets, depend on the accuracy of jet energy measurements. This accuracy is characterised by the Jet Energy Resolution (JER), which also affects the reconstruction of the $E_{\mathrm{T}}^{\mathrm{miss}}$. The JER must be determined to account for the limited accuracy of the detector, which is influenced by pileup and detector effects. It is parametrised by a noise term (N), a stochastic term (S), and a constant term (C), and is expressed as:

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{N}{p_{\mathrm{T}}} \oplus \frac{S}{\sqrt{p_{\mathrm{T}}}} \oplus C. \tag{5.9}$$

The noise term accounts for electronic noise from the front-end electronics and contributions from pileup, both of which are independent of the energy deposited in the calorimeters. As a result, this term scales with $1/p_{\mathrm{T}}$. The stochastic term arises from statistical fluctuations in the deposited energy and scales with $1/\sqrt{p_{\mathrm{T}}}$. The final constant term represents fluctuations due to jet shape variations and detector hardware effects, which contribute a fixed fraction of the jet p_{T}. This includes non-uniformities in the calorimeter response and variations in energy deposition within passive material. The noise term is expected to dominate at $p_{\mathrm{T}} < 30$ GeV, while the stochastic term

becomes the leading contribution at higher p_{T}, up to approximately $p_{\mathrm{T}} > 400$ GeV, where the constant term takes over.

The JER is determined using the same methods as the in situ JES calibration, by comparing jets within a well-measured dijet system. This process involves fitting the dijet balance measurements to the functional form of Equation 5.9, using a fixed noise term calculated separately via a random cones method, as described in Ref. [133]. To ensure that the average JER in simulation, σ_R^{MC}, matches that observed in data, σ_R^{data}, a smearing procedure is applied where necessary. In regions of jet p_{T} where $\sigma_R^{\mathrm{data}} > \sigma_R^{\mathrm{MC}}$, the simulated events are smeared so that the average resolution aligns with data. Conversely, if $\sigma_R^{\mathrm{MC}} > \sigma_R^{\mathrm{data}}$, no smearing is applied to preserve the data resolution. Instead, the difference is accounted for as an additional systematic uncertainty, defined as the full discrepancy in resolution between data and simulation (see Section 5.5.3).

5.5.3 Jet uncertainties

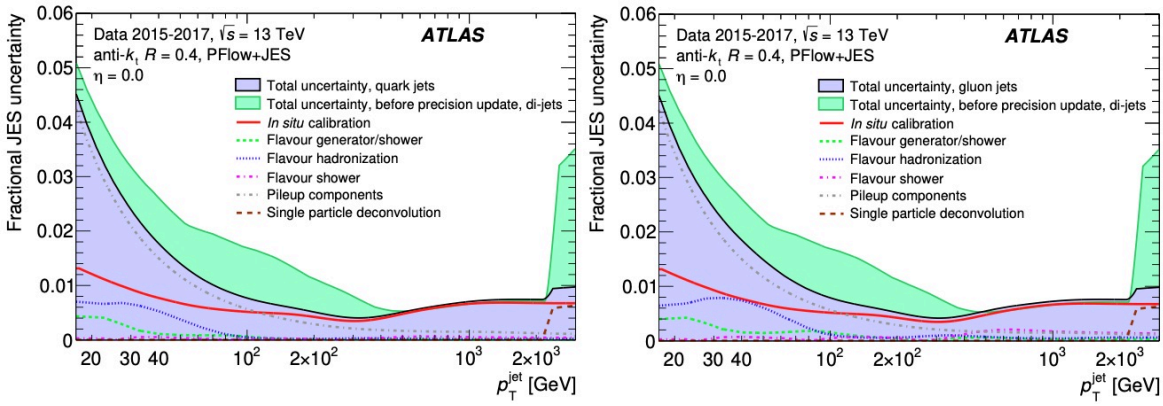
Systematic uncertainties are assigned to account for the various stages of the JES calibration chain. These include contributions from the η-intercalibration, MPF, and MJB methods used in the in situ calibration, as well as uncertainties related to pileup modelling and flavour dependence. Recently, the flavour uncertainty components, which account for variations in jet response based on parton flavour, were redefined, leading to a reduction in their magnitude by up to a factor of two compared to the previous approach [143]. Additionally, systematic uncertainties associated with the determination of the JER are incorporated [133].

Jet energy scale uncertainties

The uncertainty in the JES calibration consists of 130 individual terms, with 116 of these arising from the in situ calibration. The η-intercalibration method initially contributes 18 terms, which are reduced to 6 by combining multiple uncorrelated uncertainties. The remaining 98 terms are associated with the MPF and MJB analyses, accounting for effects such as analysis selection criteria, limited sample statistics, simulation mis-modelling, topology uncertainties, and uncertainties related to the calibration of leptons and photons. These 98 terms depend only on jet p_{T}, allowing for an eigenvector decomposition that reduces them to a smaller set of orthogonal terms while minimising the loss of correlation information between terms. This reduction is

achieved through the *category reduction scheme* [133], which condenses these terms into 15 orthogonal components. In addition to these in situ terms, 4 terms are included for pileup corrections, 1 term accounts for punch-through uncertainty related to jet mis-modelling in the GSC calibration stage, and 1 term covers high-p_{T} ($p_{\mathrm{T}} > 2.4$ TeV) jets, derived from single-hadron response using a deconvolution method [140, 141]. Furthermore, 8 terms describe flavour-dependence uncertainties [143]. In total, the uncertainties in the JES calibration are described by 35 independent components.

A summary of the fractional JES uncertainty, incorporating all 35 components, is presented in Figures 5.3a and 5.3b for quark-initiated and gluon-initiated PFlow jets, respectively [132]. These figures indicate that the pileup and in situ components of the fractional JES uncertainty are each slightly below 2% and 1%, respectively, for jets with $p_{\mathrm{T}} = 40$ GeV. Both components are reduced to around 0.4% at $p_{\mathrm{T}} = 200$ GeV. The total fractional uncertainty tends to decrease with increasing jet p_{T} and remains below 1% for jets with $p_{\mathrm{T}} > 100$ GeV. The figures also illustrate the improvement in the total JES uncertainty resulting from the recent changes to the evaluation of flavour-dependent uncertainty components [143], which are now discussed.



(a) **(b)**

Figure 5.3: The fractional JES uncertainty for (a) quark- and (b) gluon-initiated jets produced at $\eta = 0$. The new uncertainty is indicated by the filled blue region, and the filled green region shows the difference between the new total uncertainty and that obtained using the previous uncertainty definitions calculated in Ref. [133]. The various components contributing to the total uncertainty are included. This has been taken from Ref. [143].

Previously, uncertainties related to flavour dependence were accounted for using three terms [133]: the flavour composition and flavour response uncertainties, along with an additional term specifically applied to jets initiated by b-quarks. The flavour composition uncertainty quantified the difference in response between quark- and

gluon-initiated jets, given an uncertainty σ_g^f on the fraction of gluon-initiated jets in a sample, f_g. It was defined as:

$$\sigma_{\text{composition}} = \sigma_g^f \frac{|\mathcal{R}_q - \mathcal{R}_g|}{f_g \mathcal{R}_g + (1 - f_g)\mathcal{R}_q}, \tag{5.10}$$

where $\mathcal{R}_q$ and $\mathcal{R}_g$ represent the jet responses of quark- and gluon-initiated jets, respectively, as measured in PYTHIA8 [105]. The flavour response uncertainty accounted for variations in the response of gluon-initiated jets between different MC generators. It was defined by comparing the response measured using PYTHIA8, $\mathcal{R}_g^{\text{P}}$, with that from Herwig7 [113], $\mathcal{R}_g^{\text{H}}$, as follows:

$$\sigma_{\text{response}} = f_g \left(\mathcal{R}_g^{\text{P}} - \mathcal{R}_g^{\text{H}} \right). \tag{5.11}$$

Finally, the uncertainty term applied exclusively to b-quark initiated jets accounted for differences in jet response between light-flavoured (u, d, s) and heavy-flavoured (c, b) quark-initiated jets. This uncertainty replaced both $\sigma_{\text{composition}}$ and σ_{response} for heavy-flavour quark-initiated jets.

Since these flavour uncertainties were first defined and implemented, the underlying causes of variations in the ATLAS detector response to jets across different MC simulations have been studied in detail [144]. These studies examined the dependence of jet response on the particle content of a jet, leading to the conclusion that most differences arose from differences in the hadronisation models used in the simulations. In particular, a difference of up to 20% was observed in the simulated mean baryon energy fraction across the jet p_{T} range between PYTHIA8 and HERWIG7, whereas SHERPA [145] exhibited better agreement with PYTHIA8. Additionally, it was found that jets containing a higher fraction of baryons tend to have a lower response. Further studies [146] evaluated the agreement between hadronisation models in current MC generators and data from ATLAS and LEP. These analyses showed that the Lund string model (see Section 4.3), as implemented in PYTHIA8 and SHERPA, as well as the cluster model (see Section 4.3) in SHERPA (once tuned to LEP baryon content data), exhibited better agreement with experimental data. In contrast, the untuned cluster model in SHERPA and the cluster model in HERWIG7 showed poorer agreement. As a result of these findings, the flavour response uncertainty was redefined based on a more careful consideration of a jet's particle content. Instead of determining the

flavour response uncertainty by comparing PYTHIA8 and HERWIG7, the uncertainty was now decomposed into three distinct components [143]:

1. Flavour generator/shower uncertainty: Defined by comparing jet response between PYTHIA8 and SHERPA, both using the Lund string hadronisation model. This captures differences in momentum-ordered vs. dipole parton showers, as well as generator-related effects from UE and MPI.

2. Flavour hadronisation uncertainty: Defined by comparing SHERPA with the cluster hadronisation model to SHERPA with the Lund string model, isolating purely hadronisation effects.

3. Flavour shower uncertainty: Defined by comparing HERWIG7 with an angle-ordered parton shower to HERWIG7 with a dipole parton shower, isolating differences in shower modelling.

These three uncertainties are separately derived for light-flavoured quark- and gluon-initiated jets, with an additional three derived for heavy-flavour quark-initiated jets. This eliminates the need for the previously used heavy-flavour uncertainty component. Moreover, these 6 new flavour uncertainties are applied per jet, meaning that the flavour composition uncertainty is no longer required. However, because this work was completed after the in situ JES calibration, the previous uncertainties, $\sigma_{\text{composition}}$ and σ_{flavour}, are still propagated through the in situ calibration and included in physics analyses for the time being. Consequently, a total of 8 terms currently account for JES uncertainties related to flavour dependence.

Figures 5.3a and 5.3b illustrate that the recent precision improvements have reduced the total JES uncertainty by approximately a factor of two for jets with $p_{\text{T}} \approx 100$ GeV, due to the revised flavour uncertainty definitions. The significant reduction observed for high-p_{T} jets is largely attributed to the use of an updated single-particle uncertainty [143]. However, for the measurement discussed in Part III, the improved flavour uncertainties are the more critical factor. Not only are these new per-jet flavour uncertainties considerably smaller than the previous response and composition uncertainties, but they are also smaller than the variation in jet response across different MC generators. Consequently, to ensure their accurate application, MC2MC scale factor corrections must be applied to alternative generators to align the jet response with that of PYTHIA8 [142]. As previously discussed in Section 5.5.2, this correction is implemented as the final stage of the JES calibration chain.

Jet energy resolution uncertainties

Systematic uncertainties in the JER determination are necessary to account for the limited precision in JER. Additional uncertainty contributions arise from variations in the dijet selection criteria and differences in predictions from alternative MC generators. Furthermore, uncertainties in the random cones method, which is used to determine the noise term of the JER, are propagated by repeating the JER measurement with different values of the noise parameter, N. The final set of JER systematic uncertainties consists of 13 independent terms. Figure 5.4 presents the fractional JER uncertainty as a function of jet p_{T}.

JER uncertainties are incorporated into physics analyses by smearing jets with a Gaussian function of width σ_{smear}, defined as:

$$\sigma^2_{\mathrm{smear}} = \left(\sigma_{\mathrm{nominal}} + |\sigma_{\mathrm{syst}}|\right)^2 - \sigma^2_{\mathrm{nominal}}, \tag{5.12}$$

where $\sigma_{\mathrm{nominal}}$ represents the nominal[2] JER of a sample, and σ_{syst} corresponds to a variation of one standard deviation in the considered uncertainty component. To ensure proper propagation of JER systematic uncertainties in physics analyses, smearing is typically applied to the simulation when $\sigma_{\mathrm{syst}} > 0$ and to the data when $\sigma_{\mathrm{syst}} < 0$. This procedure affects only the systematic variation of the JER component while preserving the JER in data. In cases where the data sample lacks sufficient statistics, pseudo-data may be used for smearing instead. This approach, in which smearing is applied to both data and simulation based on the sign of σ_{syst}, ensures that anti-correlations are preserved when the uncertainty component crosses zero [133]. Additionally, JER uncertainties must account for discrepancies between JER in data and simulation. When the JER is lower in simulation than in data, this is already addressed by the initial smearing of the simulation (see Section 5.5.2). However, an additional systematic uncertainty is required to cover cases where the JER is lower in data than in simulation. This uncertainty is taken as the full difference between the JER in data and in simulation.

[2]The term "nominal" is used throughout this thesis to refer to a value, sample, or condition that represents the standard or expected case, based on theoretical predictions using the SM, without the application of any systematic uncertainty variations.

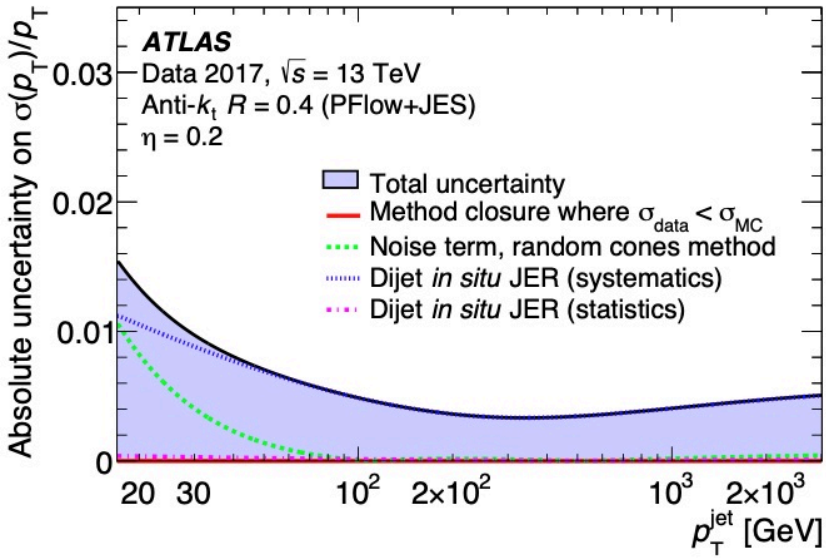


Figure 5.4: The fractional JER uncertainty for PFlow jets [132] with $\eta = 0.2$. The total uncertainty is indicated by the filled blue region. The various components contributing to the total uncertainty are included. This has been taken from Ref. [133].

5.6 *b*-tagged jets

As top quarks are reconstructed from their decay products, and their decay predominantly involves a b-quark, it is crucial to distinguish jets initiated by b-quarks, known as *b-jets*, from those initiated by light-flavoured quarks and c-quarks. This identification process is referred to as *b-tagging*. In the measurement discussed in Part III, the DL1r b-tagging algorithm [147] is employed. This algorithm utilises a recurrent neural network to differentiate b-jets from jets initiated by light-flavoured quarks and c-quarks by analysing the presence of secondary vertices and correlations between track impact parameters. The long lifetime of b-hadrons typically results in b-jets containing a secondary vertex and tracks with high impact parameters, which aids in their identification. Beyond distinguishing b-jets, the DL1r algorithm must also achieve effective background rejection for light-flavoured and c-quark jets, quantified by the inverse of their efficiencies. The algorithm was trained using a dataset in which 70% of the jets originated from $t\bar{t}$ events, while the remaining 30% were from $Z' \to q\bar{q}$ events [147]. For the measurement in Part III, a working point is chosen that corresponds to a b-tagging efficiency of 77% per b-jet, along with jet p_{T}-dependent rejection rates of up to 170 for light-flavoured quarks and 5 for c-quarks.

To ensure that the b-tagging efficiency and background *mistag* rates in simulation accurately reflect their performance in data, scale factors are applied. Systematic uncertainties are assigned to account for variations in b-tagging performance arising from the calibration [148]. These uncertainties are propagated by varying the b-tagging scale factors in simulation. Similar to the approach used for jets described in Section 5.5, an eigenvector decomposition is performed, resulting in 19 independent systematic uncertainty components: 9 for the b-tagging efficiency calibration, 4 for the c-quark mistag calibration, 4 for the light-flavoured quark mistag calibration, and 2 for the extrapolation of b-tagging results to high-p_{T} jets. The uncertainties exhibit a dependence on p_{T}, and this variation is accounted for through the eigenvector decomposition.

5.7 Reclustered large-radius jets

The m_t measurement in Part III focuses on hadronic top-quark decays with high transverse momentum ($p_{\mathrm{T}} \gtrsim 2m_t$). Since the angular separation of the jet decay products scales as $2m_t/p_{\mathrm{T}}$, these *boosted* top quarks are reconstructed using the anti-k_t algorithm with $R = 1.0$ [135], clustering the standard ATLAS calibrated PFlow small-R jets with $R = 0.4$ (see Section 5.5). This approach is beneficial in the high-p_{T} regime, where the particles from the hadronic top-quark decay may overlap in the detector, allowing the entire decay to be captured by a single jet object.

Once these Reclustered (RC) large-R jets are formed, a trimming procedure [149–151] removes any constituent jet carrying less than 5% of the large-R jet's p_{T}, as these are likely to originate from pileup. The RC large-R jets must contain at least two small-R constituent jets, accommodating cases where the hadronic decay of a high-p_{T} W-boson merges into a single small-R jet. Additionally, at least one constituent jet must be tagged as a b-jet, as described in Section 5.6.

Since RC large-R jets are built from already calibrated small-R jets, no further energy scale or resolution corrections are required. Likewise, no additional systematic uncertainties are introduced, as previous studies have shown that propagating the uncertainties from small-R jets to large-R jets is sufficient [152, 153].

5.8 Missing transverse momentum

In an ideal pp collision at the LHC, partons collide head-on along the beamline. Consequently, the total p_{T} of all objects must sum to zero due to momentum conservation. However, when reconstructing an event using detected objects, cases may arise where the total p_{T} does not sum to zero. This discrepancy can be attributed to undetected SM or BSM particles, as well as the mis-reconstruction of final-state particles. Since neutrinos are the only SM particles expected to traverse the entire ATLAS detector without leaving a track or energy deposit, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is typically associated with neutrinos in physics analyses that do not consider BSM processes.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ is determined as the negative vector sum of the transverse momenta of all reconstructed objects in an event, including leptons and jets. Additionally, $E_{\mathrm{T}}^{\mathrm{miss}}$ includes a term that accounts for the remaining soft radiation in the event. This soft term is constructed from tracks in the ID with $p_{\mathrm{T}} > 400$ MeV that are matched to the primary vertex and are not associated with any other reconstructed object. To prevent double counting of energy contributions between different objects in the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction, an overlap removal method is applied, as detailed in Ref. [154]. For the events used in the m_t measurement discussed in Part III, $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed using central ($|\eta| < 2.5$) jets with $p_{\mathrm{T}} > 20$ GeV and forward ($|\eta| > 2.5$) jets with $p_{\mathrm{T}} > 30$ GeV.

Since $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed from the negative vector sum of the transverse momenta of calibrated hard objects along with the additional soft term, systematic uncertainties are assigned to account for variations in reconstruction performance due to uncertainties in the p_{T} determination of both hard and soft objects. The energy scale and resolution uncertainties for leptons and jets are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ while preserving correlations between uncertainty components. Additionally, scale and resolution uncertainty terms for the soft term are included, derived from comparisons between $Z \to \mu^+\mu^-$ events in simulation and data [154].

PART II

Developing the Fast Track Finder particle track reconstruction algorithm for the ATLAS Phase-II upgrade

CHAPTER 6

The ATLAS Phase-II upgrade and ITk

The complete ATLAS Run 2 dataset provides an integrated luminosity of 140 fb^{-1} [56], representing a significant milestone in data collection. Similarly, at the time of writing, Run 3 of the LHC has already enabled ATLAS to record an unprecedented integrated luminosity of 183 fb^{-1} [61]. However, many analyses remain limited, or altogether unfeasible, given the current statistical constraints. These include measurements of Higgs boson properties, such as its couplings to fermions [155] and its self-interactions [156], which are crucial for a deeper understanding of the mechanism behind EW symmetry breaking. Additionally, a larger dataset would significantly enhance measurements of SM parameters, rare SM processes, and searches for BSM phenomena.

Enter the HL-LHC [157], a decade-long initiative beginning in 2030 designed to achieve a peak instantaneous luminosity of $7.5 \times 10^{34} cm^{-2} s^{-1}$, which is seven times the original LHC design value. This ambitious goal corresponds to operating the HL-LHC with up to 200 inelastic pp collisions per beam crossing, far surpassing the average of 54 collisions recorded in 2024 during Run 3 [61], resulting in an increase in the typical pileup density from approximately 0.5 vertices per mm to 1.5 vertices per mm [158]. Over its lifetime, the HL-LHC is expected to enable ATLAS to collect a dataset of up to 3000 fb^{-1}.

The high-luminosity environment presents significant challenges that the current ATLAS detector is not equipped to handle. Chief among these is the need for a substantially upgraded detector system to replace the current ID, which is already impacted by radiation damage [159, 160]. Key requirements for the new system include [161]:

- It must be capable of withstanding approximately 10 times more ionising radiation than the current ID.
- It must have a higher granularity to maintain a low particle hit occupancy of 1%.
- It must have minimal inactive material to preserve tracking performance.
- It should be able to handle a 10-fold increase in trigger rates with the output of the higher-level software-based trigger running at 10 kHz.

To address these challenges, ATLAS has developed the new all-silicon Inner Tracker (ITk). This is designed to deliver enhanced radiation-hardness, granularity, and readout speed, and reduced inactive material, ensuring the continuation of ATLAS's excellent tracking performance.

The integration of the ITk will be one of the most significant upgrades introduced during the ATLAS Phase-II upgrade, which is scheduled to precede Run 4 [162]. Other major enhancements include the addition of a high-granularity timing detector [163] to mitigate pileup using timing information, and new muon chambers [164] to improve trigger efficiency and momentum resolution while reducing fake signals from pileup.

Due to the increased complexity and event size caused by high-pileup conditions, the trigger rates for the ATLAS detector at the HL-LHC are expected to increase by a factor of ten. Consequently, the TDAQ system, as discussed in Section 3.3.6, requires a significant redesign [165, 166] to take advantage of the physics potential of the HL-LHC. The goal of the upgraded TDAQ system is to enhance trigger selectivity, increasing physics acceptance at the EW scale by ensuring that algorithms can manage the elevated pileup conditions while maintaining low energy thresholds.

This chapter opens with an overview of the ITk design and the upgraded ATLAS TDAQ system, and introduces the Event Filter (EF), a software-based trigger system intended to replace the existing HLT. This includes a discussion of the potential technological approaches for the EF, thereby establishing the context for Chapter 7, which focuses on the development and implementation of the CPU-based tracking solution.

6.1 The ATLAS inner tracker detector

6.1.1 Overview

The ITk detector will feature pixel [167] and strip [168] systems, enabling particle track reconstruction within $|\eta| < 4.0$. It will consist of approximately 27,000 sensor and readout modules, covering a total area of about 180 m^2 and incorporating over 5 billion readout channels. In comparison, the current ATLAS ID spans just 63 m^2 and contains roughly 50 times fewer channels. The inner pixel system and outer strip system are arranged in concentric cylindrical layers in the barrel region and ring-shaped layers in the end-caps. A schematic of one quadrant of the ITk, highlighting the pixel and strip systems, is shown in Figure 6.1. The ITk configuration is designed to ensure a minimum of 9 hit measurements per particle track in the barrel region and 13 in the end-caps[1], with the capability to reconstruct tracks from charged particles with $p_T > 1$ GeV.

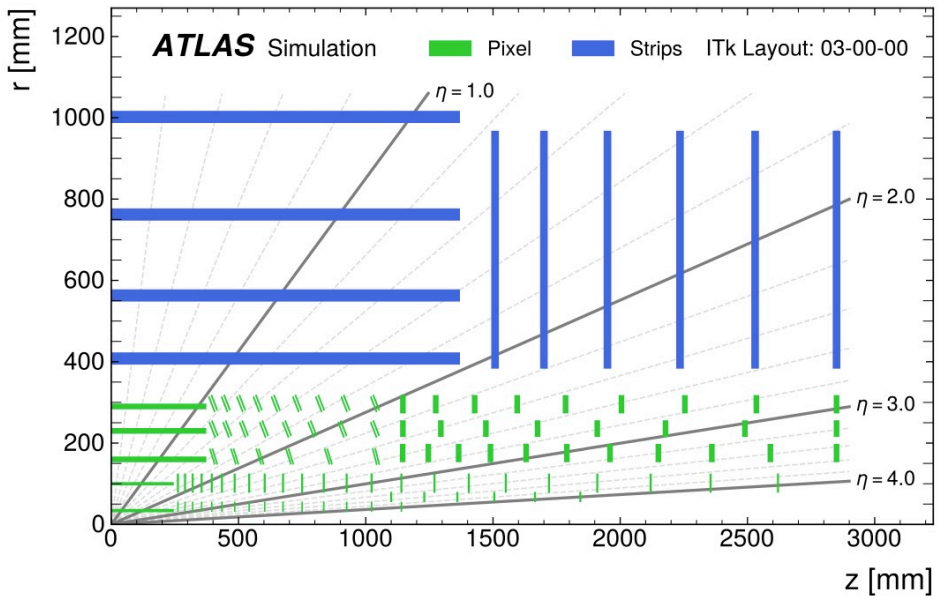


Figure 6.1: A schematic of the ITk detector layout taken from Ref. [169].

[1]The minimum requirement of 13 hits in the barrel region can be achieved due to the double-sided module mounting in the layers of the strip system, as described in Section 6.1.3.

6.1.2 Pixel system

The pixel system [167] is divided into three subsystems: the Inner System (IS), the Outer Barrel (OB), and the Outer End-Cap (OE). The layout of the pixel system is depicted in Figure 6.2. The IS, which includes the two innermost barrel layers and corresponding sets of rings, is designed for replacement after approximately 2000 fb^{-1} of recorded data due to the significant radiation exposure it will sustain. The OB consists of three flat layers and three sets of rings, covering the remainder of the barrel region. These rings are designed such that modules are mounted at inclined angles, thereby minimising the material traversed by particles travelling through this region. Lastly, the OE comprises three additional sets of rings mounted to the OB, extending coverage in the end-cap region.

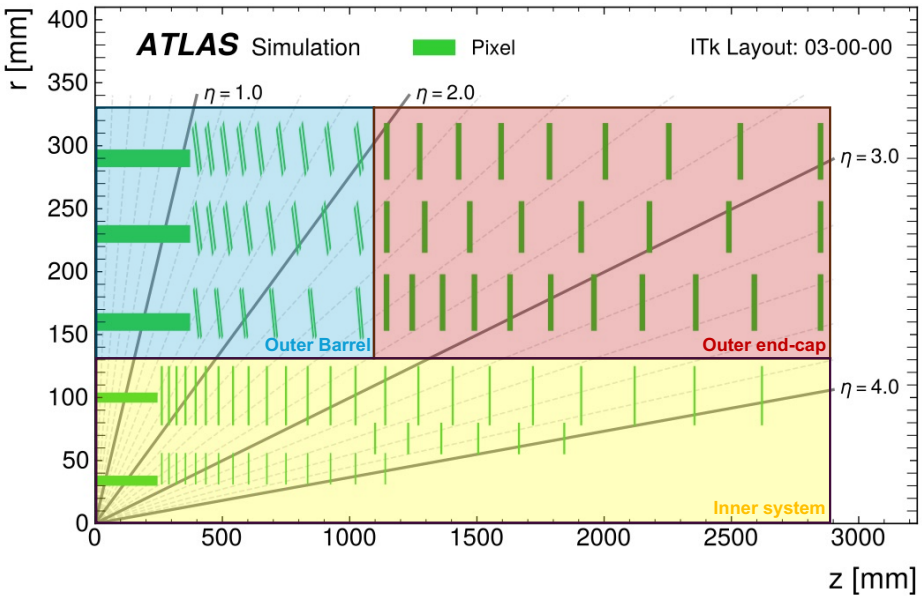


Figure 6.2: The layout of the pixel system in the ATLAS ITk taken from Ref. [169]. The different parts of the system have been highlighted.

Pixel modules are designed with a variety of architectures, thicknesses, and assembly configurations to achieve the high granularity and performance required by the detector. Each module is fabricated by bonding a sensor to one or more front-end readout chips, which are subsequently attached to a flexible printed circuit. The assembled modules are then mounted onto dedicated support structures: rectangular staves for the barrel region and rings for the end-caps.

To ensure the modules meet the stringent performance requirements of the ITk during HL-LHC operation, they undergo extensive validation. This includes achieving module efficiencies of up to 98.5% for un-irradiated modules and 97% for irradiated modules. Testing of irradiated modules is typically conducted in *test beam* campaigns. For example, the required performance has been demonstrated for both un-irradiated and irradiated pixel modules using a 120 GeV pion beam at the CERN Super Proton Synchrotron [170].

The pixel modules incorporate several advancements over the design used in the IBL [167]. Notable improvements include a reduction in pixel sensor size to 50 μm x 50 μm for most sensors, enhancing spatial resolution and particle track separation while maintaining a low hit occupancy of 1% and reducing detector material. The front-end readout chips have been optimised to handle the increased hit density, higher event rates, and elevated radiation levels expected during HL-LHC operation, all while reducing power consumption. This reduction in power consumption benefits the material budget by decreasing the need for extensive cabling and cooling infrastructure. Additionally, the readout chips feature a significantly increased output bandwidth, enabling efficient handling of higher hit rates. The pixel modules will also implement a serial powering scheme, minimising the number of required cables and further reducing the amount of inactive detector material.

6.1.3 Strip system

The strip system [168] comprises four barrel layers, each consisting of rectangular staves with modules (approximately 97 mm $\times$ 97 mm in size) installed on both sides. Similarly, end-cap modules are mounted on either side of trapezoidal petals, which are assembled to form six rings.

Each strip module consists of a sensor, one or two hybrid circuits, and a power board. Pre-production modules have been tested with a 6 GeV electron beam following irradiation to simulate the expected lifetime dose of the HL-LHC. These tests demonstrated excellent performance, with an efficiency of 99% and noise occupancy below 0.01% [162]. Furthermore, the modules are specifically designed to handle the 1 MHz trigger rates anticipated during HL-LHC operation [168].

6.1.4 Expected performance

The expected tracking performance of the ITk at the HL-LHC has been evaluated through offline reconstruction of simulated collision events [169]. This evaluation includes track reconstruction efficiency, the incidence of mis-reconstructed tracks, track parameter resolutions, and the reconstruction and identification of primary vertices. In all these areas, the ITk is anticipated to achieve excellent performance, even under the challenging conditions of an average of 200 inelastic interactions per bunch crossing, $\langle\mu\rangle = 200$.

The track reconstruction efficiency at $\langle\mu\rangle = 200$ in the barrel region of the ITk is expected to be slightly lower (by less than 5%) compared to the Run 3 detector when studied using $t\bar{t}$ events for particles with $p_{\mathrm{T}} > 1$ GeV. This performance is particularly impressive given the larger material budget of the ITk relative to the ID. In the $|\eta| > 2.5$ region, the efficiency is comparable to that achieved in the barrel, whereas the ID is unable to provide tracking information in this region. Moreover, the efficiency remains remarkably stable across different values of $\langle\mu\rangle$, with the performance at $\langle\mu\rangle = 200$ deviating by less than 0.5% from that expected in the absence of pileup interactions.

Mis-reconstructed tracks can be categorised into two types: fake tracks that do not correspond to any true particle trajectory and tracks associated with genuine trajectories but containing mis-attributed hits from pileup particles. At $\langle\mu\rangle = 200$, the rate of mis-reconstructed tracks is expected to be approximately 2%, a significant improvement compared to the 6.5% observed during Run 3. This substantial reduction is attributed to the optimised ITk detector layout, which provides a greater number of silicon hits (and therefore more detailed track information), enabling tighter selection criteria during track reconstruction to enhance performance. The fake track creation rate was determined using MC simulations, specifically examining cases where tracks have a matching probability of less than 50% with any true trajectory. This fake rate is predicted to be as low as 3×10^{-4} at $\langle\mu\rangle = 200$, highlighting the exceptional fake suppression capability of the ITk.

The smaller pixel sensor size used in the ITk, compared to the ID, leads to significant improvements in the resolution of track parameters. For example, the transverse and longitudinal impact parameter resolutions improve by up to 20% and a factor of 2, respectively, for 2 GeV muons. The improvement is even more pronounced for

100 GeV muons, which experience less scattering in the detector material, where the resolutions improve by factors of 2 and 4, respectively.

The expected performance of primary vertex reconstruction and identification has also been assessed. While the combined reconstruction and selection efficiency of the primary vertex decreases from approximately 99% for the ID under Run 3 conditions to 92% for the ITk at $\langle\mu\rangle = 200$, the resolution of the longitudinal position of the vertex improves by more than a factor of 2.

6.2 The upgraded ATLAS TDAQ system

6.2.1 Overview

Similar to the design of the current TDAQ system (see Section 3.3.6), the system for ATLAS at the HL-LHC will comprise a hardware-based, low-latency, real-time trigger operating at 40 MHz and a software-based trigger running at 1 MHz, employing offline-like algorithms on a large computing farm [165, 166]. This system could also have the potential to be augmented with commercial accelerators such as GPUs or Field Programmable Gate Arrays (FPGA). The upgraded system is designed to handle an output trigger rate of 10 kHz, an order of magnitude higher than the current system's rate. The two different triggers are the Level-0 (L0) trigger and the EF. The TDAQ architecture for Phase-II is shown in Figure 6.3, and is detailed in the following sections.

6.2.2 Level-0 trigger

The L0 trigger [165] is illustrated in Figure 6.3. It comprises the L0 calorimeter and muon systems (L0Calo and L0Muon), the Muon-to-CTP-Interface (MUCTPI), the Global Trigger, and the Central Trigger Processor (CTP). The latency of the L0 trigger will be increased to 10 μs, up from the 2.5 μs used for the current L1 hardware trigger, due to the larger event sizes anticipated at the HL-LHC. As the complexity of events increases, so does the combinatorics in the detector, leading to longer processing times for typical trigger algorithms.

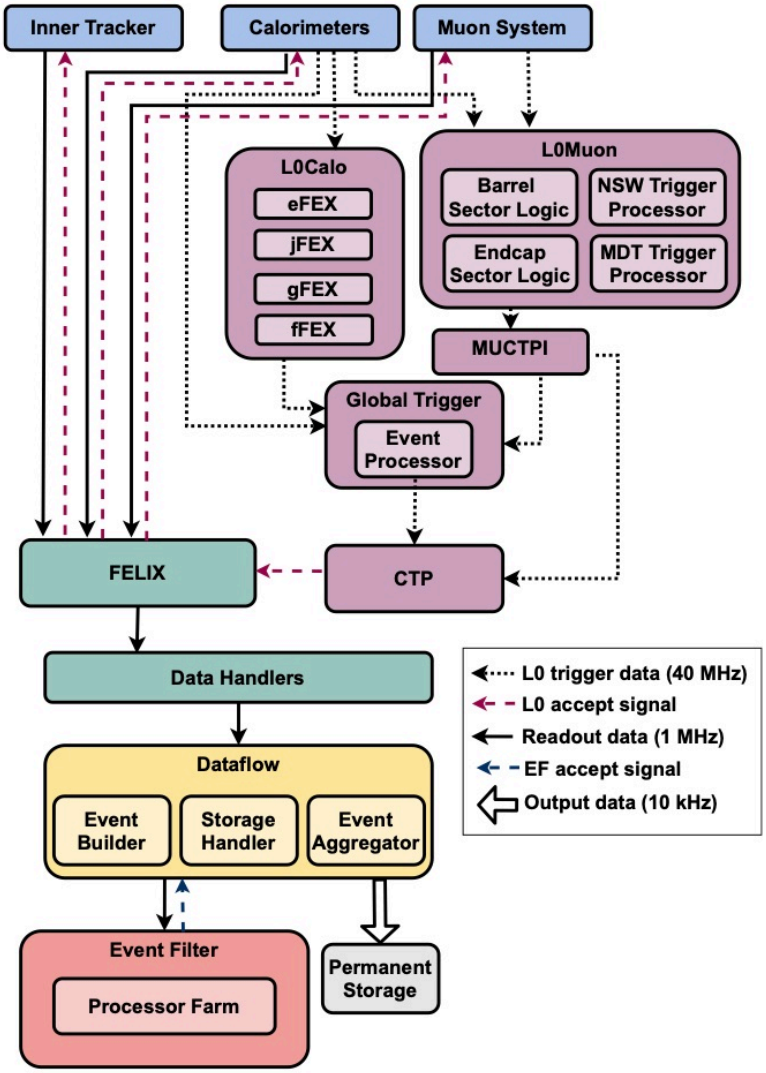


Figure 6.3: The architecture of the ATLAS TDAQ system for the Phase-II upgrade taken from Ref. [166]. The L0 trigger is shown in purple, the EF trigger in orange, the readout system in green, the dataflow in yellow, and the different detector systems in blue.

The L0Calo [165] utilises information from the EM liquid argon and hadronic tile calorimeter systems (see Section 3.3.4) to perform initial event selection and identify objects in RoIs. This information is sent to Feature Extractors (FEX), which are based on FPGA and identify electron, photon, and tau candidates (eFEX), tau, jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$ (jFEX), large-radius jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ (gFEX), as well as forward electrons with $|\eta| > 2.5$ and forward jets with $|\eta| > 3.3$ (fFEX). These objects are identified with improved resolution compared to the current TDAQ system, owing to the use of higher-granularity calorimeter information in the upgraded architecture, as detailed

in [165]. The generated trigger objects are then forwarded to the Global Trigger system for further processing.

The L0Muon system [165] uses information from all MS sub-detectors (see Section 3.3.5) to identify muon candidates in RoIs. This is a significant improvement on the current Run 3 detector, where only information from the RPC, TGC, and New Small Wheels[2] are used. Successful muon candidates are sent to the MUCTPI system.

The MUCTPI system [165] is designed to eliminate overlap between the final muon candidates created by the L0Muon system and calculate their multiplicities. The identified candidates are then transferred to the Global Trigger for further processing, while muon multiplicity information is sent to the CTP.

The Global Trigger [165] leverages high-granularity calorimeter information, along with data from the L0Calo and L0Muon systems, to enhance the identification of leptons, photons, and jets, while also performing pileup subtraction. This is achieved using offline-like algorithms that apply topological selections to the trigger objects. The results from the various algorithms are then sent to the CTP for further processing.

The CTP [165] gathers information from the Global Trigger, MUCTPI, forward detectors, and sub-detector calibration systems to make a final decision. If the event is selected for further processing, a L0 accept signal is sent to the readout system.

6.2.3 Readout system

The readout system, shown in Figure 6.3, consists of the Front-End Link Exchange (FELIX) board [172] and Data Handler systems. FELIX collects data fragments from the detector front-end electronics, a process triggered by the L0 accept signal from the CTP. If the decision is made to process the event further, FELIX extracts the data from the detector front-end and transfers it to the Data Handler. The Data Handler performs detector-specific processing such as data monitoring and formatting before passing the data to the Dataflow.

[2] These new detector systems were introduced in Run 3 to address issues related to reduced muon tracking acceptance and the increased production of high-p_{T} fake tracks in the high pseudorapidity region, which arise as a result of the increased luminosity of the LHC [171].

6.2.4 Dataflow

The Dataflow system [165] aggregates data fragments from the Data Handler to begin the initial full-event building process before sending the data to the EF. It is also responsible for transferring event data to permanent storage based on the final trigger decision made by the EF.

6.2.5 Event filter

The EF is designed as a purely software-based trigger for the Phase-II TDAQ system [165, 166]. Its purpose is to receive data from the Dataflow at a rate of 1 MHz and select the most important events for permanent storage at a rate of 10 kHz. This is accomplished using sophisticated offline-like algorithms that leverage all available detector information, similar to the HLT used during Run 2 and Run 3. However, the higher luminosity and pileup conditions of the HL-LHC introduce significant challenges for the event reconstruction algorithms. In particular, increased pileup leads to higher occupancy in the tracking detectors, making track pattern recognition slower and more susceptible to fakes, while the calorimeter energy resolution is reduced. As a result, trigger rates increase and efficiency decreases. To address these challenges, two strategies are employed. First, the hardware of the ATLAS detector must be upgraded to handle the higher pileup conditions, which is partially addressed by the ITk detector, discussed in Section 6.1. Second, the software for the trigger selection must be improved to ensure that algorithms are robust against the increased pileup, enabling efficient charged particle track reconstruction at a pileup of 200. This is the main focus of Chapter 7, which details my work on adapting the trigger software to the expected ITk geometry and optimising a specific trigger algorithm for the HL-LHC conditions.

As shown in Figure 6.3, the EF will rely on a processor farm for object and event reconstruction. This system is designed to be a flexible commercial setup using CPU cores, with the option to incorporate accelerators such as FPGAs and GPUs [166]. A final decision on the technology choice for the EF Processor Farm is expected to be made in mid-2025. To inform this decision, various *demonstrators* are being developed, but can be categorised by three groups: CPU-based, FPGA-based, and GPU-based. These demonstrators evaluate the performance of tracking algorithms on different accelerators in terms of tracking efficiency and resolution. The ultimate technology

choice will consider not only the trigger performance indicated by these demonstrators but also factors such as power consumption, physical space, cost, flexibility, scalability, and operational requirements. A heterogeneous approach may be taken where a mixture of technologies is used, allowing for the optimal selection of hardware to enhance various aspects of object and event reconstruction, thereby maximising the physics performance of the trigger. Accelerators also offer the potential for improved performance with lower power consumption and space requirements. Given the critical importance of the technology decision, having a demonstrator capable of assessing trigger performance is essential. Chapter 7 describes my contributions as a lead developer in integrating a functional charged particle track reconstruction chain into the trigger system, enabling it to run on CPU hardware using a simulated Run 4 dataset with the ITk detector geometry. The results of this implementation will provide key insights for the technology selection and, ultimately, the design of the ATLAS TDAQ system for the HL-LHC.

CHAPTER 7

CPU-based particle tracking solution for the Event Filter trigger

This chapter presents the development and successful implementation of a CPU-based particle tracking solution for the EF of the Phase-II ATLAS TDAQ system. This work represents the first instance in which the ATLAS trigger has been adapted to process simulated collision event data using the detector geometry of the ITk. Prior to this effort, the trigger was not capable of handling such data within the `main` branch of the GitLab repository [173] for the ATLAS Athena framework [174]. Athena is responsible for the majority of the core ATLAS production workflows, including event generation, simulation, and reconstruction. It is currently employed for the HLT and will also be used for the future EF.

As emphasised in Ref. [166], it is essential to simulate trigger track reconstruction on shared Grid and high-performance computing resources. These resources may not always have access to the same accelerator hardware as that available at ATLAS. One potential solution is to develop software to enable the simulation of track reconstruction with the ATLAS trigger on CPU hardware. Consequently, a CPU-based solution for the Phase-II EF is likely to be essential, underscoring the importance of the work presented in this chapter. Moreover, as mentioned in Chapter 6, such a solution is critical for providing the necessary information to guide the EF technology decision.

The topics, methods, and results discussed in this chapter reflect work I carried out between January 2022 and May 2024 as a lead developer of the CPU-based charged particle track reconstruction pipeline. During and after this period, significant contributions from other members of the collaboration also had a direct impact on the

pipeline outlined here. These contributions include optimisation studies and software restructuring.

The chapter is organised as follows. Section 7.1 provides an overview of the track reconstruction chain for the Phase-II EF. Section 7.2 discusses the Fast Track Finder (FTF) charged particle track reconstruction algorithm, covering its original design for Run 2 and the changes made for the HL-LHC. The main significant change is the use of a recently developed track seeding algorithm based on a Graph Neural Network (GNN), enabling the trigger to reconstruct tracks effectively in the expected high pileup environment at the HL-LHC. Section 7.3 presents optimisation studies for FTF, aiming to balance track reconstruction efficiency and CPU resource usage. Nightly build tests were introduced to provide daily performance monitoring for the CPU-based pipeline, and are described in Section 7.4. Finally, Section 7.5 summarises the work completed, highlights its impact, and outlines the latest performance of the FTF algorithm.

7.1 Overview of Event Filter track reconstruction

The reconstruction chain for charged particle tracks in the ATLAS EF trigger consists of five primary steps, following a methodology similar to that currently employed for track building (see Section 5.1):

- **Clustering and space-point formation:** Raw data from the ITk detector are processed to cluster detector hits and form space-points.
- **Track seeding:** Initial track candidates, consisting of a small number of space-points, are generated. These candidates act as seeds for the next stage, where the full track candidates are built. The seeding process produces initial estimates of potential tracks but can also include many fake seeds that do not correspond to real particle trajectories. Generating seeds reduces the unnecessary use of CPU resources on building unlikely full track candidates that extend through additional detector layers.
- **Track following:** Seeds generated in the previous step are extended to additional layers in the detector to construct full track candidates. Tracks with trajectories reconstructed in the barrel (end-caps) require at least nine (seven) hits across both pixel and strip layers.

- **Ambiguity and fake resolution:** Duplicate tracks are removed, fake tracks are rejected, and ambiguous candidates are resolved to improve the overall quality of the reconstructed tracks.
- **Track refit:** A final high-precision fit is applied to the track candidates to determine their parameters, such as p_{T} and η. These refined tracks are then used by trigger algorithms for tasks such as primary vertex determination, b-tagging, charged lepton identification, and PFlow [132].

The high pileup conditions at the HL-LHC necessitate that algorithms used for the track reconstruction chain are optimised for both efficiency and CPU time. To meet the performance demands posed by the increased track density, the EF will incorporate a combination of *regional tracking* at 1 MHz and *full-scan tracking* at 150 kHz [166].

Full-scan tracking involves utilising particle energy deposition and hit information from the entire ITk detector ($|\eta| < 4.0$) for track reconstruction. Its applications include b-tagging, track-based calibration for jets with $p_{\mathrm{T}} > 20$ GeV, pileup corrections, and the identification of additional soft jets. For Run 4, this tracking is designed to include tracks with $p_{\mathrm{T}} > 1$ GeV.

Regional tracking involves performing track reconstruction within RoIs defined by objects identified by the L0 trigger. Examples of these objects include potential track segments in the MS or high-energy clusters reconstructed in the calorimeters. The sizes of these RoIs are optimised for performance and tailored to the specific particle signatures such as electrons and muons. For the EF, regional tracking will operate at the full L0 throughput rate with 5% $\eta \times \phi$ coverage, focusing on identifying high-p_{T} tracks for single high-p_{T} lepton and multi-object triggers. Regional tracking is foreseen to reconstruct tracks with $p_{\mathrm{T}} > 2$ GeV.

The FTF algorithm can be used for the EF track reconstruction chain, incorporating all the steps mentioned, except for the hit clustering and space-point formation step, which is handled separately. Originally developed for the ATLAS Run 2 trigger, FTF is also currently in use for Run 3 data-taking. Unfortunately, to be used during the HL-LHC, the algorithm cannot be directly imported into the Run 4 trigger software to work "out-of-the-box" due to the evolving detector geometry. Consequently, it was necessary to modify the algorithm in the ATLAS Athena software framework to integrate it such that FTF could be used to process a Run 4 dataset and efficiently reconstruct tracks. This implementation also required the algorithm to maintain sufficient speed despite the increased pileup conditions. The work was carried out

using MC datasets that simulate the expected collision data from Run 4 with the ITk detector geometry.

7.2 The Fast Track Finder algorithm

The FTF algorithm was originally developed during Run 2 to reconstruct particle tracks in the ATLAS HLT, serving as a precursor to a precision tracking stage [121]. As a result, FTF prioritises track-finding efficiency over track purity.

The challenges posed by the HL-LHC require significant updates to the methods employed in the track reconstruction chain, particularly the track seeding step. To fully grasp the necessity of these changes, it is important to detail the algorithm and its role in the trigger, both in its original form prior to Run 4 and as adapted for Run 4.

7.2.1 Implementation before Run 4

When a particle traverses the detector, it interacts with the detector elements, enabling its trajectory to be reconstructed as a track (see Section 5.1). These interactions produce hits by depositing charge in the detector material. This charge may spread across multiple adjacent detector elements, forming clusters. Space-points are determined from these clusters using a charge-weighted approach that leverages the cluster's charge distribution.

During Run 2 and Run 3 of data-taking, FTF utilises space-points derived from particle hits in the pixel and SCT sub-detector layers of the ID [121, 175]. FTF does not make use of any information from the TRT to save CPU time.

Using these space-points, a track seeding step is performed by searching for triplets of space-points within RoIs provided by the L1 trigger. Track seed generation is carried out in approximately 50 regions of r and ϕ in the detector's cylindrical coordinate geometry [121]. When forming triplets, the middle space-point is selected first, followed by a search for the other two space-points: one at a smaller r (inner space-point) and one at a larger r (outer space-point). Space-points are not restricted to a single azimuthal region but can also come from adjacent regions with more positive or negative ϕ. Additionally, the inner and outer space-points must be consistent with the event's nominal interaction region along the beam line, as well as the z position of the

middle space-point. When RoIs from the L1 trigger are used, this region corresponds to the z width of the RoI. Illustrations of the r, ϕ, and z constraints are provided in Figure 7.1. The triplet's track parameters are approximated using a conformal transformation [176], using the middle space-point as the centre and applying cuts on d_0 and p_T [121].

For Run 3, machine learning techniques [177] were introduced to improve this seeding step by using a classifier to predict if a pair of hits belong to the same track given input hit features such as cluster width and track inclination.

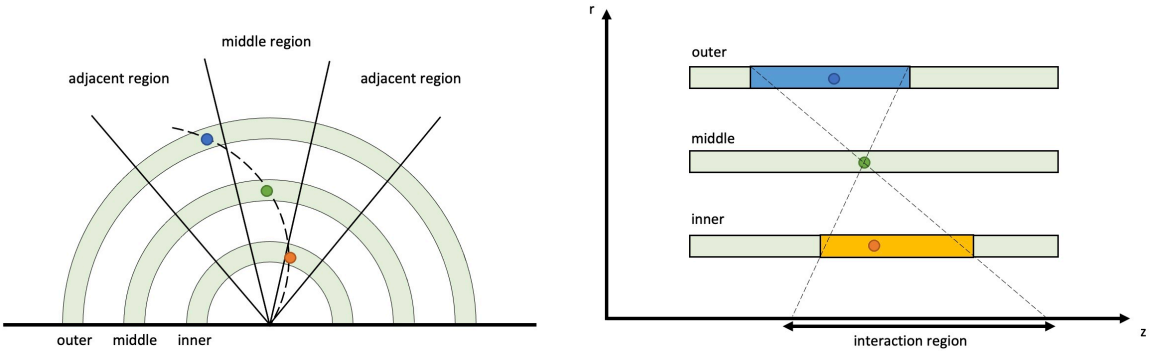


Figure 7.1: An illustration of the requirements on r, ϕ (a) and z (b) in the formation of triplets used as track seeds for FTF in Runs 2 and 3. The green shaded regions refer to layers of the detector. The coloured circles represent the clustered hits in the material. The blue and yellow shaded regions reflect how the extrapolated z-position of the triplet can be restricted a specific region.

Next, a track-following step is performed by using a combinatorial Kalman filter [118] to extend these seeds in search roads to find additional hits in subsequent detector layers, building track candidates (see Section 5.1.1). These search roads are not extended to the TRT to save CPU time.

A step to remove duplicate tracks sharing more than a certain number of hits is taken. This number of hits can be varied in the configuration of FTF to optimise performance.

Finally, a fast Kalman filter track fitter [119] is applied to produce the final tracks reconstructed by FTF. These tracks are then passed to a precision tracking stage [121, 175], which performs a full offline reconstruction to enhance track purity and quality. This precision tracking step reduces the number of fake tracks and extends tracks into the TRT, thereby improving track momentum resolution. The final track parameters are then determined.

7.2.2 New track seeding using a graph neural network

With the transition to the higher pileup environment of the HL-LHC, a challenge arises when using Kalman filter-based track following. The time required for track following per event, T, is modelled as:

$$T \approx T_{\text{seeds}} + [N_{\text{seeds}} \times T_{\text{KF}}]\,, \tag{7.1}$$

where T_{seeds} is the time taken to generate the set of track seeds, N_{seeds} is the total number of seeds and T_{KF} is the time taken by the Kalman filter for one seed. The number of seeds is related to the number of hits, N_{hits} via:

$$N_{\text{seeds}} \sim (N_{\text{hits}})^{2.5}, \tag{7.2}$$

as the number of possible connections between space-points increases combinatorially. Therefore, given Equation 7.1 and Equation 7.2, the processing time does not scale linearly with hit multiplicity.

From the current status of Run 3, it has already been concluded that seed generation is one of the most time-consuming aspects of tracking [175]. Given the anticipated increase in the number of hits in the detector during the HL-LHC, this will result in a significant rise in the number of seeds. However, not all generated seeds correspond to tracks resembling true particle trajectories. In fact, the track seeding stage used in the Run 3 implementation of FTF was tested on a Run 4 MC dataset using an offline reconstruction algorithm using the same seeding and found that approximately 85% of created seeds were eventually rejected during the track-following stage. Given this, and based on Equation 7.1, there would be a significant decline in tracking performance without improvements to the track-seed-formation stage. A more efficient seeding process would result in fewer incorrect seeds, reducing the CPU time spent rejecting these seeds during the subsequent track following stage.

A more efficient track seeding step for FTF has been developed, based on the FASTrack (Filter and Automaton for Silicon Tracking) algorithm [178], incorporating a GNN. This new seeding approach uses only space-points from the pixel layers of the ITk, where a space-point pair graph is constructed. The graph is specified using a probabilistic layer-linking scheme trained on MC simulated reference tracks for single muons with $p_{\text{T}} = 1$ GeV, uniformly distributed across $|\eta| < 4.0$. The layer-linking

scheme assigns a transition probability for connecting a hit in a given pixel layer, labelled A, to another layer, labelled B. This probability is estimated as the ratio of the number of reference tracks connecting layers A and B to the total number of reference tracks passing through A. The transition probability from A to B is independent of any preceding layers. Links with probabilities below a predefined threshold are removed by setting their probabilities to zero. This threshold is uniform across all potential links and can be optimised based on the trade-off between track reconstruction efficiency and CPU timing, with values up to 0.1% considered.

A pair of space-points formed from two connected hits is referred to as a *track segment*. Selection criteria is applied to these segments to reduce the creation of incorrect seeds before they are used for track following. These include a requirement on the curvature of the track segment (see Section 7.3), as well as a constraint on the azimuthal separation $\Delta\phi$ between the two space-points. An illustration of the resulting space-point pair graph, including potential track segments, is shown in Figure 7.2. Note that not all space-points are connected to each other, reflecting cases where the transition probability between layers is insufficient to justify a connection.

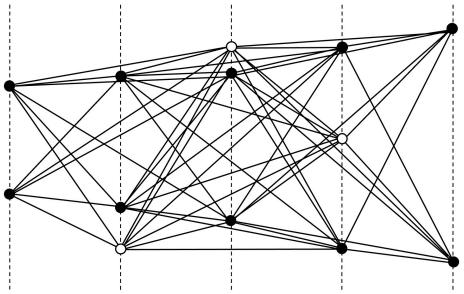


Figure 7.2: An illustration of a space-point pair graph. The dashed vertical lines represent a simplified view of the ITk pixel layers, with real particle hits corresponding to the black filled in circles. The circles with no filling represent hits in the detector not due to the particle's real trajectory.

To reduce the number of potential track segments, a machine learning classifier is employed, following an approach similar to the Run 3 implementation [177]. This classifier uses the length of pixel clusters and track inclination angles to predict compatible track segments.

An iterative cellular automaton approach [179] is used to identify track candidates. Using this method, track segments are referred to as *cells*. In the case where multiple cells share a space-point, the cells are called *neighbours*. Each cell is assigned a state,

labelled by a positive integer greater than or equal to one. This iterative method consists of two main stages: the *forward recursion* and the *backward pass*.

During the forward recursion, if a cell i with state s_i has a compatible neighbouring cell j with an equal state ($s_j = s_i$), the state of cell i is incremented by one in the next iteration. A neighbouring cell is considered compatible only if the pair satisfies specific selection criteria. These criteria closely resemble those applied to individual cells, and include requirements on the maximum difference in track azimuthal angle and the difference in curvature between the cells.

The incrementation of cell states is performed simultaneously for all cells, and the iterations stop when no neighbouring cells share the same state, indicating that no further state updates occur. The incrementation of cell states is applied in only one direction within the neighbourhood; otherwise, the states would increase indefinitely. An example of this forward recursion is shown in Figure 7.3, where the neighbourhood is constructed from left to right, meaning that each cell considers only its left-hand neighbour.[1]

The backward pass begins with the cell having the highest state S, and searches for its neighbour with state $S - 1$. From this neighbour, the cell with state $S - 2$ is found, and this process continues until the state of the last considered cell equals one. In cases where multiple neighbours share the same state, the neighbour with the smallest breaking angle to the previous one is chosen. This backward pass is repeated until all sequences are found. Since these cells represent track segments, the resulting sequences are track candidates formed from hits in the pixel layers of the ITk. It is worth noting that these track candidates are built during the seeding step, and do not refer to those candidates built in the track following stage.

From the produced track candidates, triplets of space-points are sampled to form the track seeds. Each triplet consists of two consecutive track segments that share a common space-point.

An illustration of how this CA approach works on the example space-point pair graph shown in Figure 7.2 is given in Figure 7.4.

[1] For the geometry of the ATLAS ITk, the neighbourhood is constructed from the outermost layers inward toward the origin. Consequently, each cell searches for a neighbouring cell that is located closer to the origin.

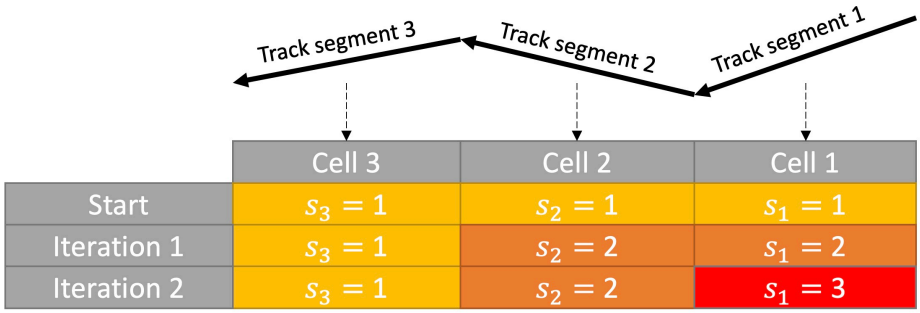
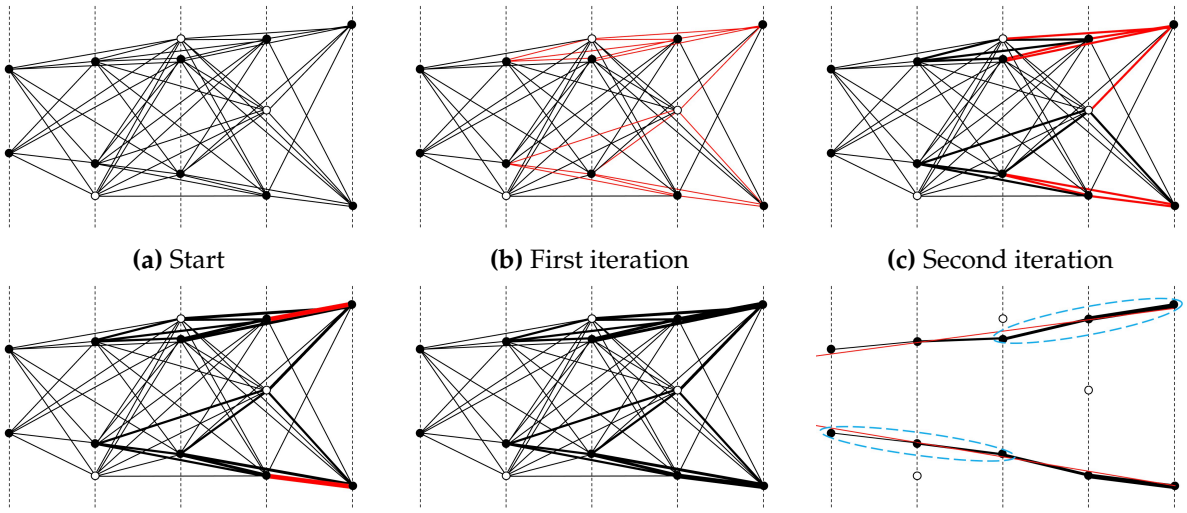


Figure 7.3: An illustrated example of how the forward recursive stage of the cellular automaton approach updates cell states. The neighbourhood is built from left to right, meaning that a cell will look for neighbour to its left.



(a) Start **(b)** First iteration **(c)** Second iteration

(d) Third iteration **(e)** Fourth iteration **(f)** Backward pass

Figure 7.4: An illustration demonstrating the iterative cellular automaton starting from (a). Iterations (b) one, (c) two, (d) three and (e) four of the forward recursion are illustrated. The neighbourhood of track segments is built from left to right. The width of the track segments represent the value of the cell's state, where a thicker segment refers to a higher cell state. At each iteration, cells with states increasing by one are shown in red. The (f) backward pass forming the track candidate is also provided, in which triplet sampling is represented by the blue dashed ovals and the red line represents a straight line fit to a track candidate's trajectory.

7.2.3 Run 4 implementation with new track seeding

Within Athena, trigger algorithms are implemented in C++. To execute algorithms using an Athena job [174], Python configuration files are used to steer the job to specify the input data, the algorithms to be run, and the desired output format. For FTF, the

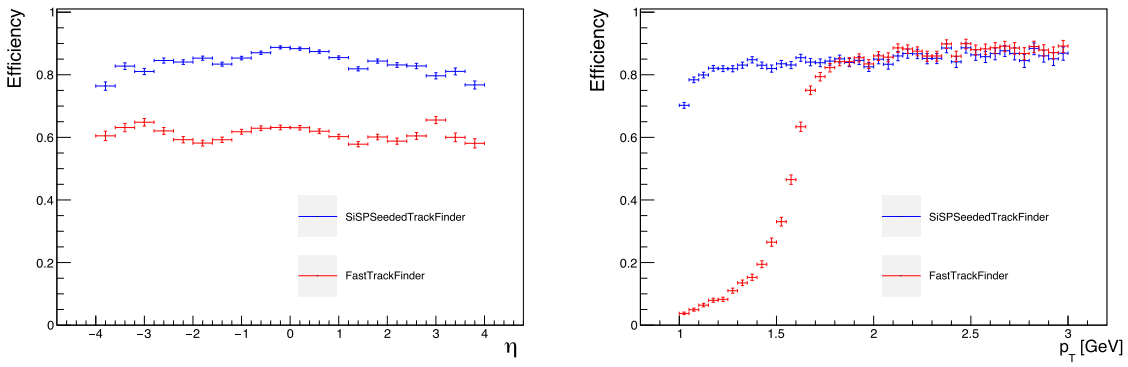
input consists of a file containing hit data from the ATLAS detector with ITk geometry. The tracks reconstructed by FTF are then stored in an output file.

FTF has already been implemented to work using the new GNN-based seeding with an older version of the ITk geometry in an older release of the Athena framework. However, these changes needed to be integrated into the `main` branch, requiring a complete rewrite of the Python configuration code to ensure it worked with other software developments introduced for the Phase-II upgrade. This integration has been completed [180, 181], enabling the first implementation of executing FTF as part of the trigger using regional tracking with RoIs. An additional implementation for FTF was also created such that it could be executed as a full-scan tracking algorithm in the offline reconstruction.

The implementation of FTF to execute as part of the EF trigger involved adding new Python configuration code to Athena, as well as modifications to the algorithm's C++ code enabling it to use the new GNN-based seeding. The new Python code included setting up the configurations for FTF, as well as additional settings for the algorithms and tools required for its operation. It also leveraged pre-existing configurations for offline reconstruction algorithms set up to work with ITk geometry, importing them into the trigger software where applicable. In cases where an offline configuration could not be directly imported for use in the trigger, new configurations were created by combining those used in Run 3 with the pre-existing configurations for the offline reconstruction algorithms designed for Run 4. As a result of these changes, FTF can be executed as part of the trigger and successfully reconstruct tracks using input events simulated with the ITk geometry.

Unfortunately, when this task was completed, the track reconstruction efficiency of FTF could not be measured due to the need for a large migration of the trigger software to a new modular framework [182]. As a result, an additional FTF implementation was introduced to enable the reconstruction efficiency to be measured. In this approach, FTF was integrated into the offline reconstruction framework, meaning it did not execute as part of the trigger and only used full-scan tracking. This allowed the use of offline reconstruction monitoring tools to create track reconstruction efficiency histograms for FTF. Furthermore, the performance of FTF could be compared directly with the standard offline track reconstruction algorithm, SiSPSeededTrackFinder, without involving other aspects of the online trigger reconstruction. Integrating FTF in this manner required fewer changes than adapting it to work within the trigger for an input ITk MC dataset. This was because most of the pre-existing configurations

for offline tools and algorithms working with ITk geometry could be reused for FTF. Similar to the above approach, these configurations were incorporated into Athena using new Python configurations, ensuring that the output of the offline track reconstruction job included tracks created by both FTF and SiSPSeededTrackFinder. This enabled the comparison of reconstruction efficiency between the two algorithms in a single Athena job. Efficiency histograms, which display the performance of FTF in reconstructing truth tracks and compare it with the offline algorithm using MC simulated $t\bar{t}$ events with no pileup, are shown in Figure 7.5. The offline reconstruction was run in the fast-tracking configuration for ITk [183], which omits the ambiguity solving stage and applies a tighter track selection, ensuring a fair comparison with FTF.[2]



(a) **(b)**

Figure 7.5: Histograms obtained from the Athena offline track reconstruction job comparing the efficiency of FTF with that of SiSPSeededTrackFinder in the fast-tracking configuration as functions of η (a) and p_{T} (b). Efficiency in each bin is calculated as the ratio of the number of reconstructed tracks found matched to true tracks to the total number of true tracks. A MC sample corresponding to 1,000 $t\bar{t}$ events with no pileup was used. The vertical error bars represent the statistical uncertainty of the input sample.

From Figure 7.5, it is evident that FTF was not effectively reconstructing tracks using the ATLAS Run 4 dataset, even with no pileup. Specifically, FTF struggled to efficiently reconstruct tracks with $p_{\mathrm{T}} < 2$ GeV. As mentioned in Section 7.1, full-scan tracking is designed to be efficient for tracks with $p_{\mathrm{T}} > 1$ GeV. Therefore, optimisation studies were necessary to address this issue.

[2]This fast-tracking implementation is still a work-in-progress with more optimisations required to increase the overall efficiency.

7.3 Optimisations

Full-scan tracking for Run 4 is intended to efficiently reconstruct tracks with $p_{\mathrm{T}} > 1$ GeV [165]. However, the FTF performance shown in Figure 7.5 does not meet this requirement. As discussed in Section 7.2.2, a maximum curvature constraint is applied to the track segments used during the triplet formation process. Specifically, track segments must satisfy a condition on the curvature, κ, of:

$$\kappa^2 < \frac{D}{(r_{\mathrm{min}})^2}, \tag{7.3}$$

where D is a constant and r_{min} is the minimum radial distance for the track segment, defined by its radius of curvature. This minimum distance is set by the minimum p_{T} requirement of 1 GeV assuming a uniform magnetic field [5].

It was found that increasing D from its default value of 0.4 improved efficiency for low-p_{T} tracks, as shown in Figure 7.6. The results in Figure 7.6 indicate that a value of $D \geq 1.2$ is necessary to recover efficiency for tracks with $p_{\mathrm{T}} < 2$ GeV. However, processing time measurements for this study, provided in Table 7.1, reveal a negative impact on FTF processing time when D is increased. This behaviour is expected, as raising the curvature threshold allows more seeds to be generated, thereby increasing the duration of the track-following stage. Consequently, while increasing D was essential to regain efficiency for low-p_{T} tracks, additional modifications were required to mitigate the adverse impact on FTF processing time.

Table 7.1: The impact of increasing D on the total processing time and the number of seeds generated per event for FTF is shown. These results were derived from 200 $t\bar{t}$ events simulated with $\langle\mu\rangle = 200$, using an Intel(R) Xeon(R) CPU E5-2620 v4 running at 2.10 GHz.

D	Number of input seeds per event	Total FTF processing time per event (s)
0.4	11175	1.22
0.6	17334	1.47
0.8	23006	1.63
1.0	28156	1.82
1.2	32801	2.06
1.4	37021	2.10

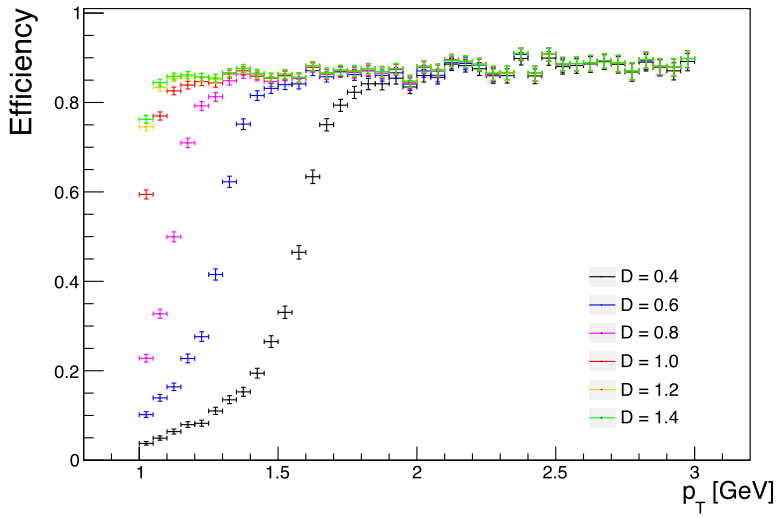
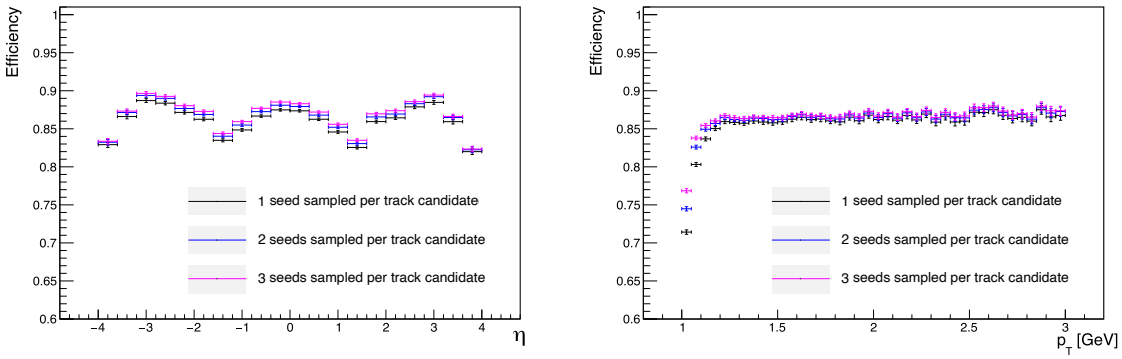


Figure 7.6: Histograms produced from executing FTF as an offline full-scan tracking algorithm illustrate the impact of varying the value of D in Equation 7.3 on the FTF track reconstruction efficiency for low-p_{T} tracks. Efficiency in each bin is computed as the ratio of reconstructed tracks matched to true tracks to the total number of true tracks. A MC sample of 1,000 $t\bar{t}$ events simulated with no pileup was used. The vertical error bars represent the statistical uncertainty of the input sample.

Using the results of these studies, the seeding algorithm for FTF was updated to enhance both efficiency and processing time [184,185]. Efficiency improvements stemmed from replacing the fixed cut on the track segment's $\Delta\phi$ (mentioned in Section 7.2.2) with a dynamic cut that depends on the radial distance between the space-points of the track segment and the minimum p_{T} used for track reconstruction. Additionally, Equation 7.3 was applied with $D \approx 1.6$.

To mitigate the negative timing impact of using this larger D value, the number of track seeds sampled per track candidate during the seed generation step was reduced from three to one. While Figure 7.7 shows a minor decrease in efficiency due to this adjustment, the timing results presented in Table 7.2 demonstrate a significant improvement in FTF processing time.

The trade-off in efficiency to reduce CPU time was deemed acceptable. Following this optimisation study, updated performance plots were generated and are presented in Figure 7.8. These results were obtained using 10,000 Monte Carlo simulated $t\bar{t}$ events, each with $\langle\mu\rangle = 200$. Figure 7.8 demonstrates that the poor efficiency observed in Figure 7.5 was successfully recovered. Moreover, FTF achieved higher efficiency than



(a) **(b)**

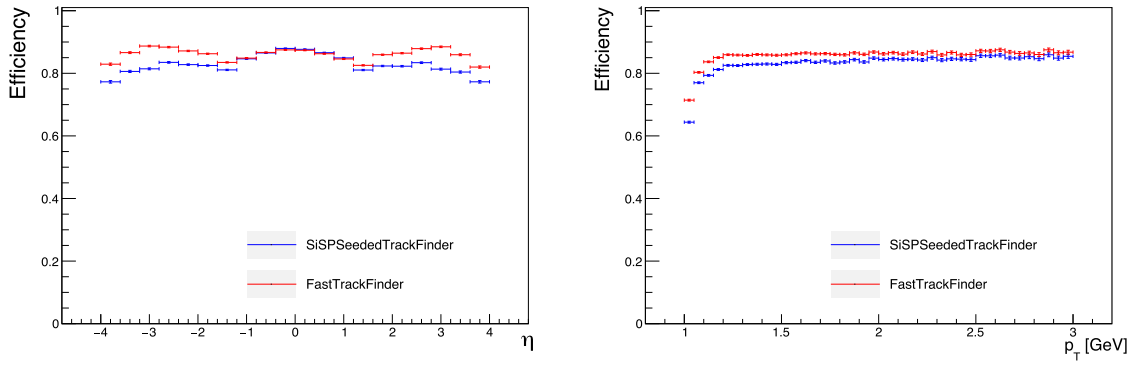
Figure 7.7: Histograms produced from executing FTF as an offline full-scan tracking algorithm showing the effect of changing the number of seeds sampled per track candidate on FTF efficiency for low-p_{T} tracks as functions of η (a) and p_{T} (b). Efficiency in each bin is computed as the ratio of reconstructed tracks matched to true tracks to the total number of true tracks. A MC sample of 1,000 $t\bar{t}$ events simulated with no pileup was used. The vertical error bars represent the statistical uncertainty of the input sample.

Table 7.2: The impact of varying the number of seeds sampled per track candidate on the total processing time per event for FTF is shown. The parameter D was set to 0.8. These results were obtained from 200 $t\bar{t}$ events with $\langle\mu\rangle = 200$, using an Intel(R) Xeon(R) CPU E5-2620 v4 operating at 2.10 GHz.

Number of seeds sampled	Total FTF processing time per event (s)
1	1.52
2	1.63
3	1.86

the offline SiSPSeededTrackFinder algorithm in the $|\eta| > 1.5$ region. This improvement may be linked to FTF's enhanced efficiency for low-p_{T} tracks. A detailed breakdown of FTF's CPU time at this stage of development is provided in Table 7.3.

In comparison to the results in Table 7.3, the total processing time per event for SiSPSeededTrackFinder was calculated to be 1.35 seconds, using the same 200 $t\bar{t}$ events simulated with $\langle\mu\rangle = 200$. Given FTF's higher efficiency compared to the offline method, this difference in processing time was deemed acceptable. The new GNN-based seeding method resulted in a 77% reduction in the number of input seeds compared to the Run 3 technique. This indicates that the updated seed generation step



(a) **(b)**

Figure 7.8: Histograms produced from executing FTF as an offline full-scan tracking algorithm comparing the efficiency of FTF to that of SiSPSeededTrackFinder in the fast-tracking configuration after adjustments to the number of seeds sampled per track candidate and the value of D. Efficiency is shown as functions of η (a) and p_{T} (b). In each bin, efficiency is calculated as the ratio of reconstructed tracks matched to true tracks to the total number of true tracks. A Monte Carlo sample of 10,000 $t\bar{t}$ events simulated with $\langle\mu\rangle = 200$ was used. The vertical error bars represent the statistical uncertainty of the input sample.

Table 7.3: Breakdown of the FTF processing time per event, obtained using 200 $t\bar{t}$ events with $\langle\mu\rangle = 200$. Results were obtained using the Intel(R) Xeon(R) CPU E5-2620 v4 at 2.10 GHz.

Process	Processing time per event (s)
Track seeding	0.68
Track following and ambiguity resolution	0.71
Final track fit	0.13
Total	1.52

is now more efficient, producing fewer incorrect seeds that would subsequently be rejected during the track-following stage.

Additional optimisations to FTF's track reconstruction efficiency include modifying the minimum p_{T} requirement to depend on the track's η position [186]. This allows tracks in the end-caps of the ITk detector to have lower p_{T}, resulting in a 5% efficiency improvement in these regions. Another optimisation involved reducing the width in z used when executing FTF as an offline algorithm with full-scan tracking, aligning it with the size used by the default offline reconstruction algorithm [187]. This change led to a further 4% efficiency increase in the barrel region of the ITk detector.

Building on the success of the new GNN-based track seeding, the developer of this method proceeded to enhance the track-following stage to further optimise the CPU processing time of FTF for Run 4. This involved the development of a new non-combinatorial track-following algorithm capable of handling track seeds with more than three connected space points, along with an improved track extrapolation method. These enhancements were integrated into the Python configurations for executing FTF both as an offline algorithm and as part of the trigger [188]. The new approach resulted in halving the time spent during the track-following stage, leading to a reduction of up to 30% in FTF processing time. However, during the implementation of this new track-following method, the focus was solely on improving CPU timing. Dedicated measurements and optimisations for track reconstruction efficiency are still ongoing.

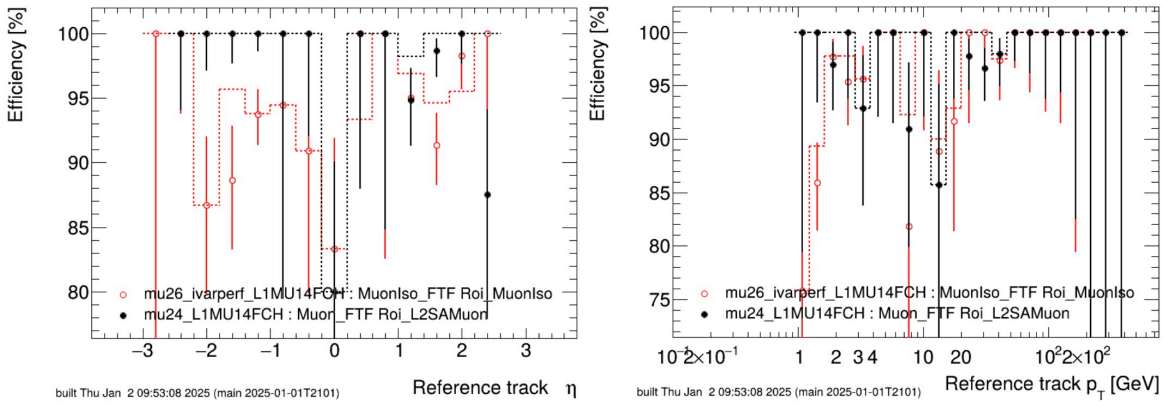
7.4 Performance monitoring

During the development of the CPU-based tracking solution for the EF using FTF, monitoring the algorithm's performance, particularly track reconstruction efficiency, is crucial. To facilitate this, tests have been integrated into Athena to simulate the track reconstruction of FTF each night when the Athena framework is built and a new nightly branch in the repository is created [180, 189–193]. This enables continuous monitoring of the algorithm's performance over time, allowing for the detection of any changes in the software that affect its performance. The introduction of these tests has proven essential for development and optimisation studies.

Currently, five distinct tests are available for executing FTF: two for running FTF as an offline algorithm using full-scan tracking, and three for running FTF as part of the trigger using regional tracking. For each implementation, the tests vary only by the physics signature of the input simulated Run 4 MC sample. For example, the three trigger tests correspond to executing FTF with Run 4 samples based on single electron, single muon, and $t\bar{t}$ signatures. It is worth noting that the lepton objects used in these samples are formed using the same definitions as those discussed in Section 5.4[3]. In the single lepton tests, samples with no pileup are used, while the $t\bar{t}$ test uses a sample simulated with $\langle\mu\rangle = 200$.

[3] In practice, the definitions of electron and muon candidates will evolve for Run 4 with the introduction of the new ITk detector. However, the underlying principles used to define these candidates will remain consistent. Therefore, it is valid to use the current simulated objects in these tests to predict the expected performance of the ITk.

The output of these tests includes histograms of track reconstruction efficiency, track parameter resolutions, and the track parameters themselves. Example output histograms for the Run 4 trigger and offline implementations of FTF are shown in Figure 7.9 and Figure 7.10, respectively. Note that these results include developments and optimisations to the software that are not covered in this chapter.



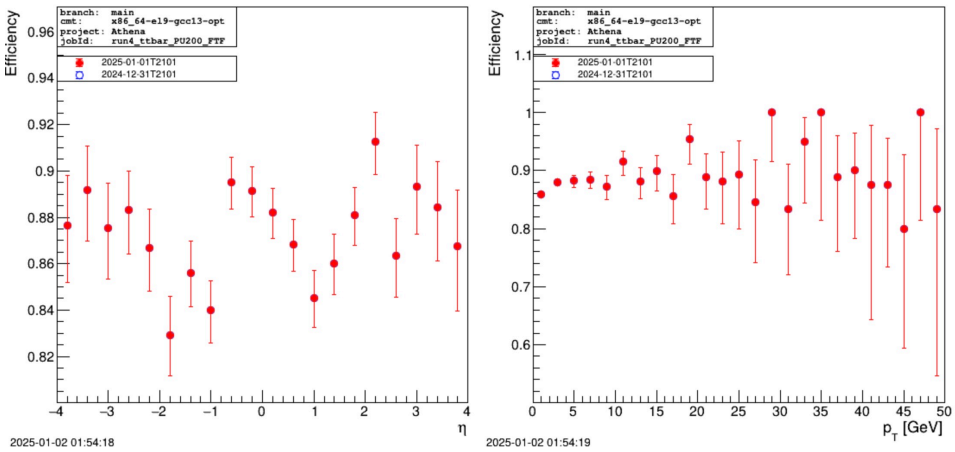
(a) **(b)**

Figure 7.9: Example output histograms for the nightly reconstruction tests for the trigger FTF implementations for Run 4. The results were obtained using a $t\bar{t}$ MC sample simulated with $\langle\mu\rangle = 200$. Efficiency is calculated using offline tracks as the reference for the trigger test. These output histograms correspond to the nightly test performed on January 1st, 2025. The different coloured points refer to the results obtained using two different trigger chains.

7.5 Summary

Given the increased pileup environment of the HL-LHC, the Phase-II upgrade of the ATLAS detector will include the new ITk detector. This brings significant challenges to the TDAQ system, which must handle a tenfold increase in throughput compared to Run 3, while maintaining excellent track reconstruction performance despite the higher number of particle interactions in the detector. To address this, a new EF trigger step is defined, utilising information from the L0 trigger to perform fast reconstruction of particle track trajectories. This will allow a final decision on whether to store the data for permanent storage and offline reconstruction.

In Run 2 and Run 3, the FTF algorithm using CPU hardware has been key to making this decision. For Run 4, a choice will be made regarding the hardware technology to use for the EF farm, whether CPU, GPU, FPGA, or a combination. The work presented



(a) **(b)**

Figure 7.10: Example output histograms for the nightly reconstruction tests for the offline FTF implementations for Run 4. The results were obtained using a $t\bar{t}$ MC sample simulated with $\langle\mu\rangle = 200$. Efficiency is calculated using truth tracks as the reference for the offline test. These output histograms correspond to the nightly test performed on January 1st, 2025. The results are shown in red, with the results of the previous nightly test shown in blue. In this case, the results of the test did not change since the last nightly build and cannot be distinguished on the histograms.

in this chapter has focused on implementing the first functional pipeline for the CPU-based solution for EF trigger track reconstruction. FTF is intended to be used for this CPU-based solution in Run 4, but required a new Python configuration to adapt it to work with an input dataset representing simulated data obtained using the new ITk geometry.

This was achieved by enabling FTF to be executed as part of the trigger for regional tracking, as well as an offline algorithm for full-scan tracking. These implementations also incorporate a newly developed GNN-based track seeding method, designed to reduce the number of incorrect seeds generated. This approach reduces combinatorics, allowing the algorithm to efficiently and quickly reconstruct tracks despite the increased pileup conditions. Optimisation studies were carried out, and nightly tests were included into Athena. These tests allow developers to monitor FTF's track reconstruction performance during development.

Continued work has resulted in impressive track reconstruction efficiencies, as shown in Figure 7.9 and Figure 7.10, and timing improvements, with FTF taking only 350 ms per event for full-scan tracking as of January 2025, using $t\bar{t}$ events simulated with $\langle\mu\rangle = 200$ on CPU hardware. Recent developments from other collaboration

members have introduced a new track-following method for FTF, which shows significant CPU speed-up. Future work will focus on optimising this new track-following procedure to achieve the best balance between track reconstruction efficiency and CPU timing.

PART III

Measurement of the top-quark mass using $t\bar{t}$ events with a boosted top quark

CHAPTER 8

Analysis strategy

The analysis presented in this thesis employs the kinematic reconstruction of the final-state decay products from the lepton + jets decay channel of $t\bar{t}$ events to determine the top-quark mass. The analysis is calibrated using MC simulation and, as discussed in Section 2.3, the precise definition of the top-quark mass used in the MC in terms of a renormalisation scheme is the subject of theoretical studies. Consequently, the mass parameter measured in this analysis corresponds to the generator-level mass. For simplicity, it will be referred to as m_t throughout the remainder of this thesis.

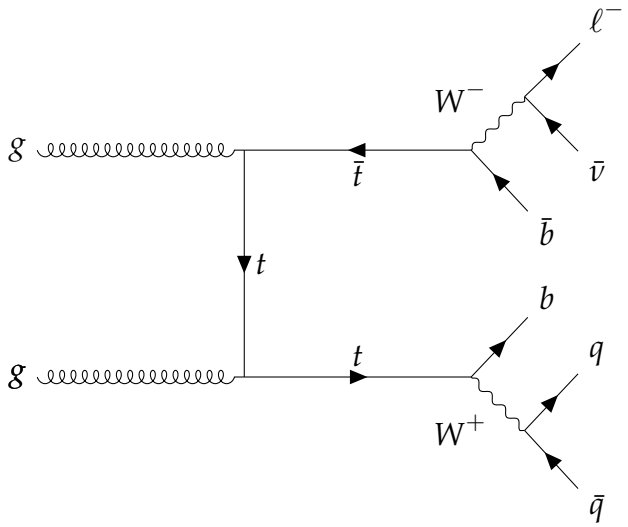


Figure 8.1: LO Feynman diagram for $t\bar{t}$ production and decay via the lepton + jets decay channel.

Figure 8.1 illustrates a LO Feynman diagram for the $t\bar{t}$ production and decay process. In this analysis, m_t is determined exclusively from events in which the hadronically decaying top quark exhibits high p_{T}. In this *boosted* regime, the decay

products of the top quark become highly collimated within the detector, allowing them to be reconstructed as a single jet object. This approach simplifies the reconstruction of boosted hadronic top-quark decays compared to conventional techniques, where the decay products are resolved as separate jets, introducing greater ambiguity in correctly assigning jets to the decaying top quarks. Consequently, this method has the potential to reduce systematic uncertainties in the measurement of m_t, thereby improving the overall precision.

The subsequent chapters detail the m_t measurement in the lepton + jets channel of $t\bar{t}$ events with a boosted top quark, recently published by ATLAS [1]. This analysis is conducted within a phase space closely aligned with a previous boosted differential $t\bar{t}$ cross-section measurement by ATLAS [151], employing many of the same reconstruction techniques and object definitions. Notably, this represents the first m_t measurement by ATLAS utilising boosted top quarks. CMS has previously performed three m_t measurements using similar boosted event topologies, with the most recent achieving a precision of 0.81 GeV [34, 194, 195].

The remainder of this thesis is organised as follows. Section 8.1 investigates the background contributions to the $t\bar{t}$ signal process. Section 8.2 describes the data and MC simulated samples used in the analysis. These simulations are employed to model the $t\bar{t}$ signal, account for background processes, and evaluate systematic uncertainties. Section 8.3 outlines the selection criteria for $t\bar{t}$ events in the boosted regime. Sources of systematic uncertainty, arising from both theoretical predictions and experimental limitations, are detailed in Section 8.4. Section 8.5 introduces the key observables used in the measurement, including the mean value of the mass distribution of the RC large-R jets used to reconstruct the boosted top quark, denoted as $\overline{m_J}$, which is sensitive to m_t. In addition, two observables are defined to control and mitigate systematic uncertainties that limit the precision of the m_t measurement. The measurement of m_t is performed using a profile likelihood fit, described in Chapter 9. This fit combines likelihood functions for the defined observables and incorporates systematic uncertainties, each represented by a Nuisance Parameter (NP) in the model. A profile likelihood fit is first applied to a dataset corresponding to the SM prediction, in order to assess the expected precision of the m_t measurement. Validation studies are conducted to ensure the robustness of the fit model prior to its application to ATLAS collision data. Chapter 10 presents the fitted values and associated uncertainties of the NPs obtained using data, before ultimately reporting the measured value of m_t. This result is then compared with previous measurements and extrapolated to evaluate

the expected improvement in precision at the HL-LHC as additional collision data become available.

8.1 Background estimation

Many processes can lead to the same final-state products as the lepton + jets $t\bar{t}$ decay channel. These backgrounds must be accounted for prior to extracting m_t from the data. The dominant background is anticipated to come from single top-quark production (known as *single-top*), which consists of three production processes: s-channel, t-channel, and tW production. Representative LO Feynman diagrams for these processes are shown in Figure 8.2.

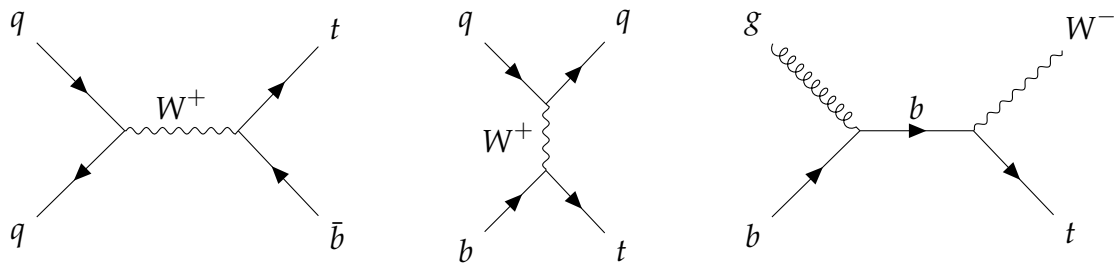


Figure 8.2: Feynman diagrams for single-top production at LO in QCD including the three production modes: s-channel (left), t-channel (centre) and tW production (right).

Among the three single-top production channels, the tW process is expected to yield the largest background contribution, as the W boson from the associated production can decay leptonically, producing a neutrino and charged lepton, while the top quark decays hadronically. This configuration closely resembles the topology of the $t\bar{t}$ signal. In contrast, for the s-channel and t-channel to produce a lepton-neutrino signature, the top quark must decay leptonically. Consequently, the boosted hadronic top quark signal must be mis-reconstructed from the remaining jets.

Another notable background contribution arises from processes in which a $t\bar{t}$ pair is produced in association with a boson ($t\bar{t}X$), where X denotes a W, Z, or Higgs boson. Representative tree-level Feynman diagrams for these processes are shown in Figure 8.3. This background will include events with a boosted top quark produced in association with a boson. Furthermore, a fraction of events will not include a boosted top quark, but the additional boson can decay in ways that can mimic the boosted $t\bar{t}$ topology. The production cross-sections for $t\bar{t}X$ processes are significantly lower

than that of single-top production, particularly tW production, meaning their overall contribution to the background is expected to be lower in comparison.

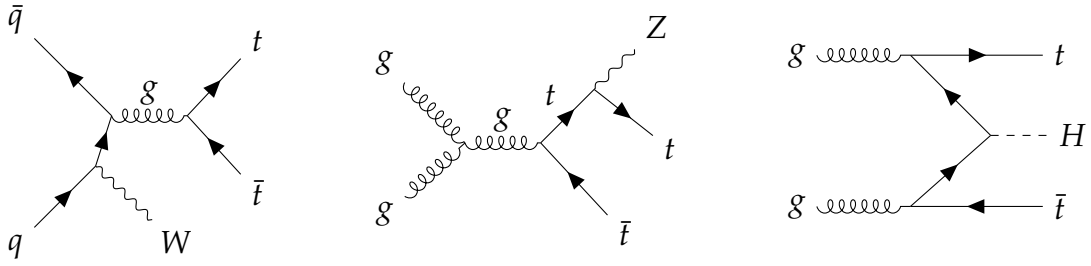


Figure 8.3: Example tree-level Feynman diagrams for the production of $t\bar{t}W$ (left), $t\bar{t}Z$ (middle), and $t\bar{t}H$ (right). These processes are collectively referred to as the $t\bar{t}X$ background to the signal process.

Processes that do not produce top quarks in the final state can still contribute as backgrounds to the signal. These include events where a vector boson (W or Z) is produced in association with QCD jets (V+jets), and diboson production (VV). Representative tree-level Feynman diagrams for these processes are shown in Figure 8.4. Since the events used in this measurement must have a lepton, Z+jets and ZZ processes are required to include the leptonic decay of a Z-boson ($Z \to \ell\bar{\ell}$). This decay mode has the lowest branching fraction among all Z-boson decay channels, and it further necessitates that the additional lepton be mis-reconstructed to produce a final state resembling the boosted topology. Consequently, these processes are expected to be significantly suppressed relative to those involving a W boson.

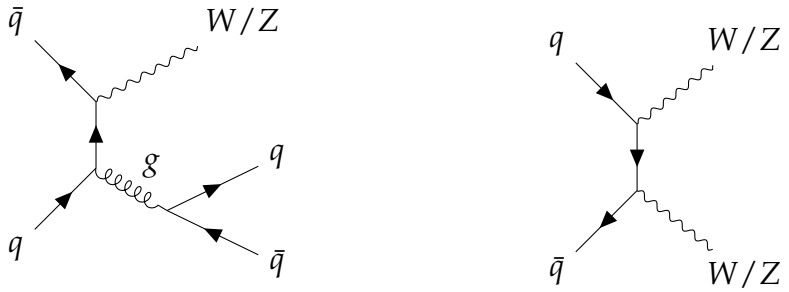


Figure 8.4: Example tree-level Feynman diagrams for processes not including a top quark in the final state: V+jets (left) and VV (right).

The final background contribution arises from multijet events, which do not contain a lepton originating from a boson decay but can still contribute if a lepton is produced in the hadronic decay of a heavy-flavour quark and is sufficiently isolated, or if constituents of a jet are mis-reconstructed as a lepton. Unlike the other back-

grounds discussed in this section, which are modelled using dedicated MC samples (see Section 8.2), multijet events are estimated using a data-driven matrix method [196]. This approach is employed because multijet processes are challenging to simulate reliably with MC generators due to the multitude of contributing processes and their strong dependence on detector performance. The matrix method has been previously applied in the ATLAS $t\bar{t}$ differential cross-section measurement using events with a boosted hadronically decaying top quark [151]. It estimates the multijet background by comparing two datasets with different lepton selection criteria, referred to as the *Tight* and *Loose* selections. The *Tight* selection uses the *Tight* identification and isolation requirements (see Section 5.4), just like the nominal $t\bar{t}$ signal selection (see Section 8.3). The *Loose* selection does not include any isolation requirements and uses the *Medium* identification requirements for both electrons and muons (see Section 5.4). As a result, the dataset using the *Loose* selection is expected to include a higher fraction of multijet events compared to the dataset using the *Tight* selection. Using the matrix method, the number of multijet events present in the *Tight* selection is given by:

$$N_{\text{multijet}}^{Tight} = \frac{\epsilon_f}{\epsilon_r - \epsilon_f} \left(\epsilon_r N^{Loose} - N^{Tight} \right), \tag{8.1}$$

where N^{Tight} and N^{Loose} represent the total number of events present in the *Tight* and *Loose* selections, respectively, while ϵ_r and ϵ_f denote the fractions of real and fake (non-prompt) leptons in the *Loose* selection that are also present in the *Tight* selection, respectively. These values depend on lepton kinematics and the characteristics of events in the *Loose* selection. They are determined using standard techniques in control regions enriched in real or fake leptons, as discussed in Ref. [197]. To account for these dependencies, the estimate for the multijet background to the signal selection is derived using event weights, w_i, calculated for each event, i, in the *Loose* selection by:

$$w_i = \frac{\epsilon_f}{\epsilon_r - \epsilon_f} \left(\epsilon_r - \delta_i \right), \tag{8.2}$$

where δ_i is defined as one if an event passes the requirements of both the *Loose* and *Tight* selections, and zero if the event only satisfies the *Loose* selection requirements. For a binned histogram of a given observable, the sum of event weights in each bin, corresponding to events containing a *Loose* lepton, provides an estimate of the total number of multijet events in that bin.

8.2 Data and Monte-Carlo simulated samples

8.2.1 Data sample

This analysis utilises data collected by the ATLAS detector from pp collisions at $\sqrt{s}$ = 13 TeV during the 2015-2018 data-taking period, corresponding to a total integrated luminosity of 140 fb^{-1}. The luminosity is measured and calibrated using the LUCID-2 detector [198] during dedicated low-luminosity runs, complemented by measurements using the ID sub-detectors and the calorimeters [56]. A small fraction of the dataset is excluded from the analysis due to detector hardware being offline or experiencing issues during specific periods [199]. Events included in the analysis are required to pass at least one of the standard single-lepton triggers for electrons or muons [128, 131].

8.2.2 Monte-Carlo samples

Simulated MC event samples play a crucial role in this analysis. They model the $t\bar{t}$ signal process and estimate the majority of background contributions. Additionally, they are used to parametrise the sensitivity of $\overline{m_J}$ to different values of m_t. These samples are also employed to construct template distributions for the observables in the profile likelihood fit framework, which extracts the best value of m_t to describe a given dataset (see Chapter 9). Furthermore, they define the NP variations, quantifying the impact of systematic uncertainties on these observables.

Wherever possible, the MC samples outlined in this section are processed using the full simulation of the ATLAS detector [200] using GEANT4 (see Section 4.6). Where this is not possible the alternative AFII simulation is used.

The effect of pileup interactions is simulated by overlaying additional inelastic pp collisions, generated with the PYTHIA8 event generator [105] using the A3 tuned parameter set [201] and the NNPDF2.3LO PDF set [202], onto the primary simulated events. For all MC samples, events are re-weighted to match the observed distribution of the average pileup in data. Additionally, the value of the average pileup is scaled by a factor of 1/1.03, based on differences between data and simulation in an inelastic pp cross-section measurement [203]. Unless stated otherwise, all MC samples use

$m_t = 172.5$ GeV. Events are processed using the same reconstruction pipeline as the data.

Unless otherwise stated all $t\bar{t}$ samples are normalised to an inclusive cross-section prediction at NNLO in QCD including the resummation of NNLL soft-gluon terms calculated using TOP++2.0 [88–91, 93, 94, 204]. This cross-section corresponds to $\sigma(t\bar{t})_{\text{NNLO+NNLL}} = 832 \pm 49$ pb using a m_t value of 172.5 GeV for pp collisions at $\sqrt{s} = 13$ TeV. This total uncertainty on the cross-section is calculated by combining the uncertainties due to the choice of PDF, α_s, m_t, μ_r, and μ_f in quadrature.

Nominal $t\bar{t}$ sample

The production of $t\bar{t}$ events is modelled using the POWHEG BOXv2 generator [107, 109, 205, 206], which generates MEs at NLO in α_s and employs the NNPDF3.0NLO PDF set [207]. The h_{damp} parameter, which regulates the high-p_T radiation against which the $t\bar{t}$ system recoils, is set to $1.5 \times m_t$ [208]. The μ_r and μ_f scales are set using the default functional form, $\sqrt{m_t^2 + {p_T}^2}$. Parton showering, hadronisation, and the UE are simulated by interfacing the generated events with PYTHIA8.230 [105], using the A14 set of tuned parameters [209] and the NNPDF2.3LO PDF set [202]. The ${p_T}^{\text{hard}}$ parameter, which controls the matching between the ME calculations of the hard scatter and the parton shower, is set to its default value of zero. This setting corresponds to POWHEG providing the p_T of its hardest emission to PYTHIA in order to define the matching between the two components. The parameter deciding which recoil scheme to use for secondary gluon emissions from the b-quark in the $t \to Wb$ vertex is set to use the `recoilToColored` scheme such that the the b-quark is the chosen recoil target. The early resonance decay parameter is set such that resonance decays can only occur after CR takes place. The decays of bottom and charm hadrons are simulated using the EVTGEN1.6.0 program [114]. This configuration, including the full simulation of the ATLAS detector, is referred to as the *nominal sample*. Additionally, a version of the nominal sample using the AFII detector simulation is available, and is used for comparisons with alternative MC samples lacking the full GEANT4 simulation.

Alternative $t\bar{t}$ samples

Four MC samples are used that span a range of m_t values. These supplementary samples are generated using the same settings as the nominal sample, with m_t set

to 171, 172, 173, and 174 GeV, respectively. Each of these samples is normalised to the NNLO + NNLL cross-section corresponding to the specific m_t value used for that sample.

Variations in the Initial-State Radiation (ISR) are assessed by re-weighting the nominal sample to mimic variations in the μ_r and μ_f scales. Additional re-weighting is performed to reproduce the effect of varying the Var3c up/down variants of the A14 tune, altering the α_s scale of the ISR, as described in Ref. [208]. Another sample is also produced with identical settings to that of the nominal, but with the h_{damp} parameter set to $3 \times m_t$.

To evaluate the effect of variations in the Final-State Radiation (FSR), two dedicated samples are used with identical settings to the nominal, but each with a different value of α_s used in the parton shower. This parameter is changed such that it relates to modifying the μ_r scale used in the parton shower by factors of 0.5 and 2.0 with respect to the nominal.

A sample identical to the nominal, but with ${p_T}^{hard}$ set to one is included to evaluate the uncertainty in the matching between the ME calculations of the hard scatter and the parton shower, as described in Ref. [210]. This alternative setting of ${p_T}^{hard}$ relates to comparing the p_T of the hardest emission provided by POWHEG against all other incoming and outgoing partons, with the minimal value chosen to define the matching scale between POWHEG and PYTHIA.

A sample with POWHEG interfaced with HERWIG7.2.1 [104, 113, 211, 212] is generated to assess the impact of the choice of hadronisation model and parton shower. This sample employs the default set of the HERWIG tuned parameters [104] and the MMHT2014LO PDF set [213]. The POWHEG settings for this sample are identical to those used for the nominal sample. The decays of heavy-flavoured quarks are simulated using the EVTGEN1.7.0 program [114].

The effects of varying settings related to MPI and CR in the A14 tune of PYTHIA8 are evaluated using samples identical to the nominal, but with modified tunes based on alternative CR models and adjusted MPI and beam remnants settings. Three distinct samples are considered, each tuned to ATLAS UE data [214], and corresponding to one of the CR0, CR1, or CR2 models (see Section 4.4). These samples are otherwise identical to the nominal sample, with the only differences being the modified MPI and CR settings in each case. Additionally, two further samples are included to assess the full range of effects related to the UE. These samples maintain the same settings as the

nominal but vary the value of α_s used in the MPI and the CR range of beam remnants, following the VAR1 eigentune from the PYTHIA8 A14 tune [209].

An ambiguity exists in the choice of recoil target for secondary and subsequent gluon emission from the b-quark in the $t \to Wb$ vertex. This is evaluated by using an alternative $t\bar{t}$ sample identical to the nominal but using the alternative `recoil-to-top` recoil scheme that models the top quark itself as the recoil target.

Two additional $t\bar{t}$ samples are used to assess the robustness of the fit setup (see Section 9.5). The first sample is generated using the same configuration as the nominal sample, but with the early resonance decay parameter adjusted such that resonance decays occur prior to the CR in the simulation. The second sample is simulated with the SHERPA2.2.1 generator [145], utilising NLO MEs for up to two partons and LO MEs for up to four partons, calculated with the Comix [215] and OPENLOOPS [216–218] libraries. These MEs are then matched to the SHERPA parton shower [106] using the MEPS@LO prescription [111, 219–221], with the set of tuned parameters developed by the SHERPA authors and the NNPDF3.0NNLO PDF set [207].

For the nominal sample, the simulation of $t\bar{t}$ events includes up to NLO QCD corrections in the ME calculation, which is matched to a parton shower. Higher-order QCD corrections matched to a parton shower are not yet available. To assess the impact of these missing corrections, an estimate is produced using three-step iterative re-weighting procedure to match the QCD prediction at NNLO. This re-weighting is based on the distributions of top-quark transverse momentum ($p_{\mathrm{T}}(t)$), $t\bar{t}$ mass ($m(t\bar{t})$), and $t\bar{t}$ transverse momentum ($p_{\mathrm{T}}(t\bar{t})$). The re-weighting process employs the NNLO $t\bar{t}$ re-weighting tool [222], which incorporates predictions at NNLO in QCD and NLO EW corrections from [92] for $p_{\mathrm{T}}(t)$ and $m(t\bar{t})$, utilising the NNPDF3.0QED PDF set [207,223]. For $p_{\mathrm{T}}(t\bar{t})$, the NNLO QCD predictions from Ref. [224, 225], using the NNPDF3.0 PDF set [207], are applied. In all predictions, m_t is set to 173.3 GeV, and dynamic μ_{r} and μ_{f} scales are used with functional forms: $m_{\mathrm{T}}(t)/2$ for $p_{\mathrm{T}}(t)$, and $H_{\mathrm{T}}(t\bar{t})/4$ for $m(t\bar{t})$ and $p_{\mathrm{T}}(t\bar{t})$, where $H_{\mathrm{T}}(t\bar{t})$ is the scalar sum of the transverse momenta of the two top quarks in the $t\bar{t}$ pair.

Background samples

As discussed in Section 8.1, various processes contribute as backgrounds to the $t\bar{t}$ signal, including those where at least one W or Z boson decays into leptons. The majority of these processes are simulated using MC generators.

Single-top tW production is modelled using the POWHEG generator [226, 227] at NLO in QCD using the five-flavour scheme[1], utilising the NNPDF3.0NLO PDF set [207]. To remove overlap between tW and $t\bar{t}$ production, the diagram removal scheme [228] is applied. Additionally, an alternative sample is generated using the diagram subtraction scheme [208, 228] to assess the uncertainty associated with this choice. The functional forms of the μ_r and μ_f scales are set to dynamic values of $\mu_r = \mu_f = H_T/2$ [229]. These events are interfaced with PYTHIA8.230 [105] using the A14 tune [209] and the NNPDF2.3LO PDF set [202]. To assess the uncertainty due to the parton shower and hadronisation model, an additional tW sample is generated using the POWHEG BOX v2 generator [107, 109, 206, 226] interfaced with HERWIG7.04 [104, 113], employing the H7UE set of tuned parameters [104] and the MMHT2014LO PDF set [213]. Samples of tW with varied m_t are not included, as the impact of this variation is found to be small enough to be accounted for by the closure uncertainty in the profile likelihood fit method (see Section 9.5). Single-top s-channel production is simulated using the same setup as the tW sample, without any interference considerations with $t\bar{t}$ processes. The single-top t-channel production uses the same setup as the s-channel but is generated using the POWHEG generator in the four-flavour scheme. All single-top samples are normalised using higher-order cross-section predictions: NNLO in QCD for the s-channel and t-channel [230], and NLO + NNLL in QCD for the tW process [231].

Events in which either a W or Z boson is produced in association with a $t\bar{t}$ pair are modelled using the MADGRAPH5_AMC@NLO2.3.3 generator [232] at NLO in QCD, utilising the NNPDF3.0NLO PDF set [207]. These events are interfaced with PYTHIA8.210 [105] using the A14 tune [209] and the NNPDF2.3LO PDF set [202]. The decays of heavy-flavour hadrons are simulated with the EVTGEN1.2.0 program [114].

[1] The two schemes typically considered are the "four-flavour scheme" and the "five-flavour scheme". These differ in the treatment of the b-quark within the hard-scatter process. In the four-flavour scheme, the b-quark is treated as a massive particle and is excluded from the PDFs, which prevents b-quarks from appearing as initial-state partons. In contrast, the five-flavour scheme allows b-quarks to participate as initial partons.

Events where a $t\bar{t}$ pair is produced in association with a Higgs boson are modelled using the same setup as the nominal sample.

The production of V+jets, where V is either a W or Z boson, is simulated using the SHERPA2.2.1 generator [145] in the same way as the SHERPA $t\bar{t}$ sample, normalised to an NNLO prediction [233].

The production of diboson final states (VV) is simulated using the SHERPA2.2.1 or 2.2.2 generator [145], depending on the specific process. Off-shell effects and contributions from the Higgs boson are also incorporated. Fully leptonic and semi-leptonic final states are generated using MEs at NLO in QCD for up to one additional parton, and at LO for up to three additional partons. Dibosons produced by loop-induced processes ($gg \to VV$) are generated using MEs at LO for up to one additional parton emission, for both fully leptonic and semi-leptonic final states. The ME calculations are matched and merged in the same manner as for the V+jets samples, using the same set of PDFs and the same dedicated set of tuned parton-shower parameters.

As discussed in Section 8.1, the QCD multijet background for the signal process is estimated using a data-driven matrix method rather than a dedicated MC sample.

8.3 Event selection and top-quark reconstruction

The events used to determine m_t must pass selection criteria based on the number and kinematic reconstruction of various physics objects. These criteria are designed to ensure a high signal purity of $t\bar{t}$ events and to ensure that the events have topologies that closely match the idealised boosted scenario, as illustrated in Figure 8.5. A summary of the event selection criteria is given in Table 8.1.

Each selected event must contain exactly one lepton candidate (see Section 5.4) responsible for triggering one of the standard ATLAS electron or muon triggers [128, 131]. Electrons must have $p_{\mathrm{T}} > 27$ GeV and $|\eta| < 2.47$, and must be located in the active regions of the EM calorimeter, avoiding the transition region between the barrel and end-cap ($1.37 < |\eta| < 1.52$). The *Tight* identification and isolation criteria are applied to ensure that the electrons are well separated and isolated from other physics objects. Muon candidates must have $p_{\mathrm{T}} > 27$ GeV, $|\eta| < 2.5$, and must satisfy the *Medium* identification and *Tight* isolation criteria.

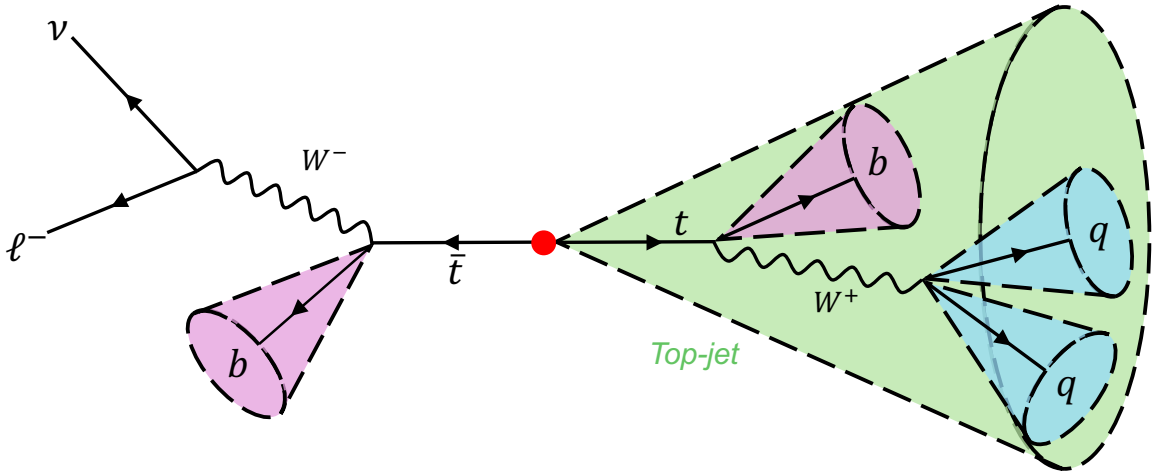


Figure 8.5: Diagram illustrating the ideal topology of a $t\bar{t}$ event decaying via the lepton + jets channel containing a boosted top quark. The top-jet (green) is a RC large-R jet that contains two small-R jets (blue) originating from the hadronic decay of the W boson, along with a small-R jet that is b-tagged (purple). It is important to note that the top-jet is only required to contain at least two jets, with at least one being b-tagged. The event also includes a second b-tagged jet and the decay products of the leptonically decaying W boson associated with the decay of the other top quark.

Small-R jets are required to have $p_{\mathrm{T}} > 26$ GeV and $|\eta| < 2.5$ to ensure that the jets are within ID acceptance. Jets that are tagged as originating from b-quarks are identified using the DL1r multivariate b-tagging algorithm [147] (see Section 5.6).

The boosted hadronically decaying top quark is reconstructed using an RC large-R jet (see Section 5.7). This jet is required to have $p_{\mathrm{T}} > 355$ GeV to ensure it is consistent with a boosted top quark, and must satisfy $|\eta| < 2$. To be considered consistent with a top-quark decay, the invariant mass of the jet must be close to m_t, falling within the range $120 < m_J < 220$ GeV. Additionally, the jet must contain at least two small-R jet constituents, with at least one of them being b-tagged. A requirement of $\Delta R(e,t) > 1.0$ ensures that the RC large-R jet is not originating from a high-p_{T} electron. Each event must contain at least one RC large-R jet that satisfies these conditions. If multiple jets pass these criteria, the jet with the highest p_{T} is selected and referred to as the *top-jet*.

The event must also contain at least one b-tagged jet that is not associated with the top-jet and is located near the selected lepton, with $\Delta R(\ell, b) < 2.0$.

To reduce the contribution from the multijet background, events are required to have a minimum amount of missing energy given by $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} + m_{\mathrm{T}}^{W} > 60$ GeV, where m_{T}^{W} is the transverse mass of the leptonically decaying W-boson[2].

A requirement is applied that the invariant mass of the selected lepton and the nearest b-tagged jet not associated with the top-jet, denoted as $m_{\ell b}$, must be less than 180 GeV. This helps reject events originating from tW production while preserving on-shell $t\bar{t}$ signal events.

The selected objects are used to approximate the top-quark kinematics. The kinematics of the hadronically decaying top quark are estimated from the selected top-jet. The kinematics of the leptonically decaying top quark are reconstructed by forming the four-vector sum of the lepton, the neutrino, and the closest b-tagged jet to the lepton that is not associated with the hadronically decaying top quark. The neutrino is not directly measured in the detector; instead, the x and y components of the $E_{\mathrm{T}}^{\mathrm{miss}}$ are used to estimate the corresponding components of the neutrino four-momentum, $p_{x,\nu}$ and $p_{y,\nu}$, respectively. The z-component is calculated by assuming the lepton-neutrino system has an invariant mass equal to the W-boson mass [6] and that the neutrino is massless. This results in a quadratic equation for the z-component of the neutrino momentum, given by:

$$p_{z,\nu} = \frac{k p_{z,\ell} \pm \sqrt{k^2 p_{z,\ell}^2 - (E_\ell^2 - p_{z,\ell}^2)(E_\ell^2 p_{T,\nu}^2 - k^2)}}{E_\ell^2 - p_{z,\ell}^2}, \tag{8.3}$$

with

$$k = \frac{m_W^2 - m_\ell^2}{2} + p_{x,\ell} p_{x,\nu} + p_{y,\ell} p_{y,\nu},$$

where $p_{X,\ell}$ represents the X-component of the lepton three-momentum, E_ℓ is the energy of the lepton, $p_{T,\nu}$ is the component of the neutrino momentum transverse to the beam axis, and m_ℓ and m_W are the rest masses of the lepton and the W boson, respectively. If the solution to Equation 8.3 is complex, the imaginary component is discarded, leaving only the real part. If there are two real solutions, the one that

[2] Defined as $m_{\mathrm{T}}^{\mathrm{W}} = \sqrt{2 p_{\mathrm{T}}^{\ell} E_{\mathrm{T}}^{\mathrm{miss}} [1 - \cos(\Delta\phi)]}$, where p_{T}^{ℓ} is the transverse momentum of the charged lepton coming from the W-boson decay, and $\Delta\phi$ is the azimuthal angle between the charged lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector.

Table 8.1: The criteria used to select events with the boosted $t\bar{t}$ lepton + jets decay channel topology.

Object	Selection criteria	
Leptons	Exactly 1 lepton in event	
	Electrons	Muons
	$p_{\mathrm{T}} > 27$ GeV	$p_{\mathrm{T}} > 27$ GeV
	$\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.47$	$\lvert\eta\rvert < 2.5$
Small-R jets ($R = 0.4$)	$p_{\mathrm{T}} > 26$ GeV	
	$\lvert\eta\rvert < 2.5$	
b-tagged jets ($R = 0.4$)	DL1r multivariate tagger at 77% efficiency	
	≥ 1 b-tagged jet is constituent of top-jet	
	≥ 1 b-tagged jet near lepton: $\Delta R(\ell, b) < 2.0$	
Hadronic top-jet (t, $R = 1.0$)	≥ 1 large-R reclustered jet candidate	
($R = 0.4$ jets as input)	$p_{\mathrm{T}} > 355$ GeV	
	$\lvert\eta\rvert < 2.0$	
	120 GeV $< m_J < 220$ GeV	
	≥ 2 small-R jet constituents	
	≥ 1 b-tagged jet constituent	
$E_{\mathrm{T}}^{\mathrm{miss}}$ & m_{T}^{W}	$E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV	
	$E_{\mathrm{T}}^{\mathrm{miss}} + m_{\mathrm{T}}^{W} > 60$ GeV	
Electron isolation	$\Delta R(e, t) > 1.0$	
m_{lb}	$m_{lb} < 180$ GeV	

minimises the mass of the leptonically decaying top quark is chosen, following the strategy used in the ATLAS differential $t\bar{t}$ cross-section measurement using boosted top quarks [151]. Small-R jets that were not used in the reconstruction of the top-jet or the leptonically decaying top quark are referred to as *additional jets*.

8.4 Uncertainties

Several sources of uncertainty arise from the limited precision of theoretical predictions and experimental methods used in the m_t measurement. These uncertainties include the reconstruction and calibration of physics objects in the detector, the modelling of the $t\bar{t}$ signal events, and the estimation of background contributions.

8.4.1 Detector uncertainties

Uncertainties associated with the reconstruction and identification of leptons, the performance and calibration of the b-tagging algorithm, the measurement of the JES and JER for PFlow jets, and the determination of $E_{\mathrm{T}}^{\mathrm{miss}}$ are collectively referred to as *detector uncertainties*. These uncertainties have been discussed in Chapter 5, and are summarised below.

Leptons

Both electrons and muons have associated uncertainties arising from the efficiency of their reconstruction and identification, which are propagated through variations in the corresponding lepton scale factors. Additionally, uncertainties in the trigger efficiencies for electrons and muons are accounted for by varying the respective trigger efficiency scale factors. Further contributions include uncertainties in the electron and muon energy/momentum scales and resolutions, which are incorporated by directly varying the momentum four-vectors of the selected leptons.

Jets

The uncertainty in the JES of $R = 0.4$ PFlow jets is parametrised using the Category Reduction scheme (see Section 5.5.3). This provides 35 independent variations, each modifying the energy of the jets. These variations include the latest precision flavour uncertainty recommendations for ATLAS, which are smaller than the previous flavour response and composition uncertainties [143]. The uncertainty in the JER is comprised of 13 independent variations, each modifying the energy of the small-R PFlow jets. As these jets are used to reconstruct the RC large-R jets, any changes in jet energy

associated with the JES and JER uncertainties are propagated to the RC large-R jets. Consequently, no additional uncertainties are required for these jets. As discussed in Section 5.7, propagating the uncertainties from the small-R jets to the large-R jets is sufficient [152, 153]. An uncertainty in the efficiency of the JVT requirement, used for pileup suppression, is included. This uncertainty is evaluated by varying the scale factor associated with the JVT requirement. Uncertainties corresponding to the calibrations of the b-tagging algorithm in data are propagated by varying the scale factor associated with the b-tagging, where the uncertainty model includes 19 independent variations.

Missing transverse momentum

Uncertainties in the $E_{\mathrm{T}}^{\mathrm{miss}}$ arise from potential mis-reconstruction of the p_{T} of the hard and soft objects used in the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction. These are accounted for by four independent components, which directly affect both the magnitude and direction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector. Additionally, uncertainties in the energy scales and resolutions of the leptons and jets are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ in a correlated manner.

8.4.2 $t\bar{t}$ modelling uncertainties

The alternative MC samples described in Section 8.2.2 are used to assess systematic uncertainties in the modelling of signal $t\bar{t}$ events. The uncertainty is estimated by comparing these alternative predictions with that of the nominal, unless stated otherwise.

Parton-shower and hadronisation

The uncertainty due to the choice of parton shower algorithm and hadronisation model is assessed using the POWHEG+HERWIG7 sample.

Matching of matrix elements and parton-shower

The uncertainty related to the matching of the NLO hard-scatter MEs to the parton shower is evaluated using the POWHEG+PYTHIA8 sample with ${p_{\mathrm{T}}}^{\mathrm{hard}} = 1$.

Initial-state and final-state radiation

Multiple sources of uncertainty are considered in the modelling of ISR. The uncertainty associated with the setting of the h_{damp} parameter is evaluated using a dedicated POWHEG+PYTHIA8 sample, where h_{damp} is scaled by a factor of two relative to the nominal value. In addition, an extra component of the uncertainty on ISR is evaluated through the comparison with the re-weighted events reproducing the effect of varying the Var3c up/down variants of the A14 tune [208].

The uncertainty associated with the simulation of FSR is evaluated using dedicated POWHEG+PYTHIA8 samples, with direct variations of the shower α_{s} value. This is equivalent to modifying the μ_{r} scale in the shower by factors of 0.5 and 2.0.

Higher-order effects

An uncertainty is included for the re-weighting procedure performed to account for missing higher order QCD corrections (see Section 8.2.2). This is evaluated by comparing the nominal prediction with that obtained with the re-weighting applied. Additionally, uncertainties related to the choice of μ_{r} and μ_{f} scales used in the hard-scatter calculation are assessed by varying both scales by factors of 0.5 and 2.0 independently in POWHEG+PYTHIA8.

Colour reconnection and underlying event

Assumptions regarding CR between the hard-scatter final state and the UE influence the kinematics of the final-state particles. Two uncertainty contributions for the modelling of CR are included by independently comparing the samples using the CR1 and CR2 models with that using CR0. Additionally, an uncertainty on the modelling of the UE is considered by using two samples obtained from applying the VAR1 PYTHIA8 A14 eigentune variations.

Recoil scheme

The parton shower algorithm simulates the emission of secondary and subsequent gluon radiation from the b-quark in top-quark decay. Two possible recoil schemes

determine which particle is chosen to recoil from this radiation: `recoilToColored` and `recoil-to-top`. In the nominal $t\bar{t}$ sample, the `recoilToColored` scheme is used, where the b-quark recoils against the emitted radiation. This influences the radiation pattern from the b-quark, affecting the shape of b-jets and leading to narrower reconstructed m_t distributions [234]. In contrast, the `recoil-to-top` scheme assigns the the top quark itself to recoil from this radiation, resulting in more wide-angle radiation, leading to more events where the b-quark is reconstructed as two separate jets in the detector. This introduces an additional modelling uncertainty due to the choice of recoil scheme. To assess this uncertainty, a dedicated POWHEG+PYTHIA8 sample using the alternative `recoil-to-top` scheme is studied, as described in Section 8.2.2. Since this ambiguity arises from approximations in the parton shower, neither scheme is expected to perfectly represent real data.

PDF set

The uncertainty due to the choice of PDF set used is evaluated using the 30 eigenvectors of PDF4LHC [235], in which each eigenvector setup is compared with that using eigenvector 0, instead of the nominal $t\bar{t}$ prediction.

8.4.3 Background modelling uncertainties

Similar uncertainties to those associated with the modelling of $t\bar{t}$ events are also applied to the modelling of the single-top background. These are considered uncorrelated with those associated to the signal. The uncertainties related to the settings of the μ_r and μ_f scales used in both the hard scatter and showering processes are evaluated using different POWHEG+PYTHIA8 re-weighted samples. In these samples, the μ_r and μ_f scales in the hard scatter are varied by factors of 0.5 and 2.0, independently, and the scales in the showering process are varied using the Var3c eigentune of the A14 tune [208]. The uncertainty associated with showering and hadronisation is evaluated using the alternative POWHEG+HERWIG7 sample. The uncertainty due to the modelling of FSR is evaluated using a POWHEG+PYTHIA8 sample where the nominal setting of the μ_r scale in the final-state shower is varied by factors of 0.5 and 2. The uncertainty arising from the method of removing the interference between the tW and $t\bar{t}$ processes is evaluated by comparing the predictions between two samples generated using the two available schemes. Finally, the uncertainty on the NNLO

cross-section used to normalise the single-top processes includes uncertainties due to scale settings, the choice of PDF set, and the value of α_s. These uncertainties are 1.9%, 3.7%, and 3.9% for the t-channel, tW production, and s-channel, respectively.

The uncertainties associated with the W+jets background are evaluated by re-weighting events to correspond to changing the μ_r and μ_f scales simultaneously by factors of 0.5 and 2. Since very few Z+jets and diboson background events pass the selection criteria, a conservative 50% cross-section uncertainty is applied to these processes.

The $t\bar{t}Z$, $t\bar{t}W$, and $t\bar{t}H$ processes that contribute to the $t\bar{t}X$ sample have associated cross-section uncertainties of 13.3%, 18.1%, and 9.9%, respectively [236]. A conservative normalization uncertainty of 18.1% is applied to the overall $t\bar{t}X$ cross-section.

For the multijet background, a normalisation uncertainty of 65% is used, as derived for the ATLAS differential $t\bar{t}$ cross-section measurement using boosted top quarks [151].

8.4.4 Other uncertainties

The uncertainty on the calibrated luminosity is 0.83% [56]. The uncertainty in the modelling of pileup in the simulation is evaluated by varying the event weights used in the pileup re-weighting procedure [237]. The uncertainty associated with limited MC statistics is addressed in Section 9.3.

8.5 Definition of observables

Average top-jet mass

To measure m_t, an observable must be chosen that is sensitive to its value. Figure 8.6 shows the simulated m_J distribution for the nominal $t\bar{t}$ signal prediction, corresponding to events that satisfy the selection criteria outlined in Section 8.3, with $m_t = 172.5$ GeV. This distribution is compared with two alternative predictions generated using different values of m_t. As is evident from Figure 8.6, an increase in m_t shifts the peak of the distribution toward higher values. Consequently, the mean value

of the m_J distribution, $\overline{m_J}$, serves as a sensitive observable for m_t. This mean value is calculated using the range of m_J used in Figure 8.6.

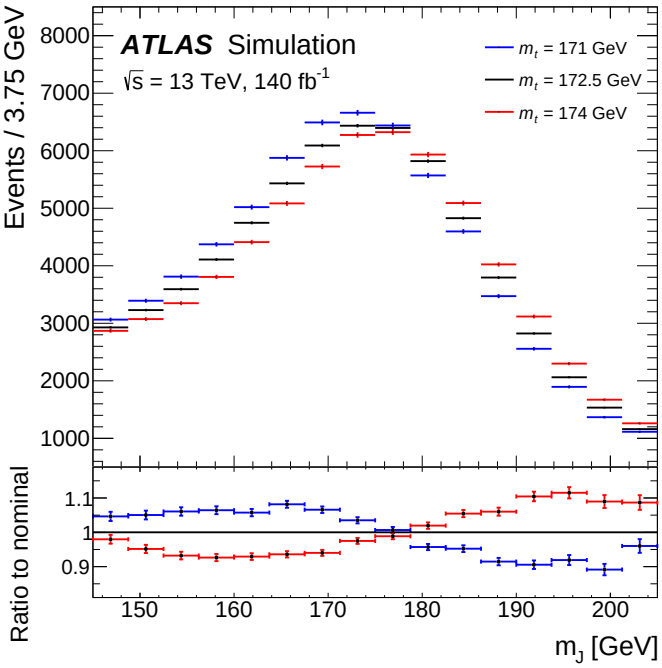
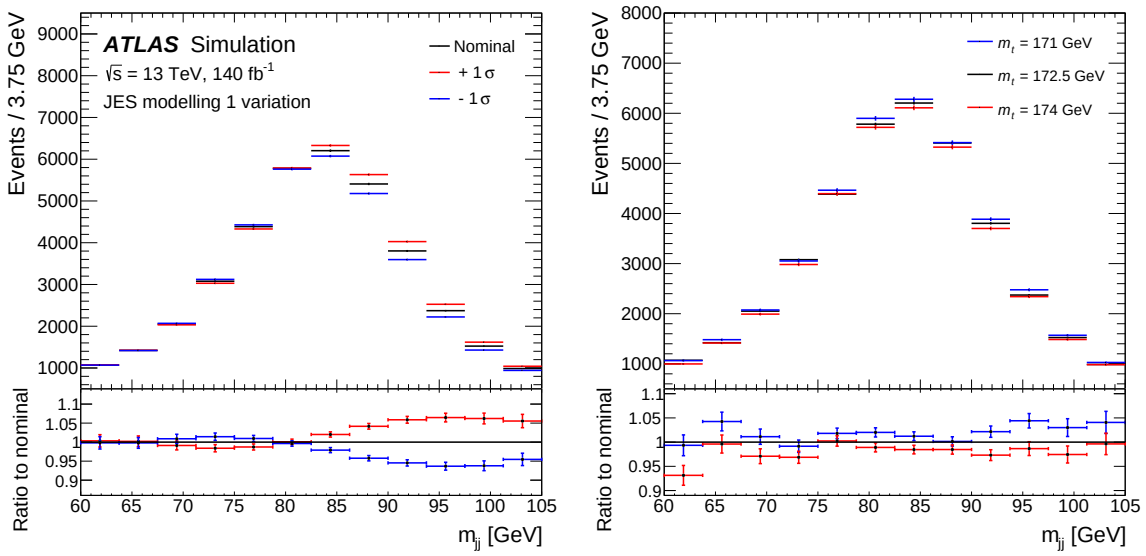


Figure 8.6: Distributions of m_J for simulated MC samples generated using three different values of m_t.

Invariant mass of the hadronic *W*-boson

This measurement relies heavily on jets, where the mass of the top-jet is directly driven by the energy of the constituent small-R jets. Therefore, the systematic uncertainties related to the JES calibration are expected to play a significant role in the m_t measurement. To mitigate the impact of these uncertainties, the hadronically decaying W boson is reconstructed using the invariant mass of the two not b-tagged jets with the highest p_{T} inside the top-jet, denoted as m_{jj}. This approach allows for the separation of effects that influence the JES (which affects both $\overline{m_J}$ and m_{jj}) from those that only affect the value of m_t (which impact $\overline{m_J}$ alone). As a result, the JES uncertainties in this measurement can be constrained when using a profile likelihood fit (see Chapter 9), reducing their contribution to the total uncertainty of the m_t result. The events used to determine m_{jj} must meet the requirements outlined in Section 8.3 and satisfy $60 < m_{jj} < 105$ GeV. Figure 8.7a shows the histogram of m_{jj} for the nominal $t\bar{t}$ signal prediction, alongside variations corresponding to the first eigenvector component of

the JES systematic uncertainties associated with the modelling of the simulation used for in-situ calibration. The distribution of m_{jj} exhibits sensitivity to the JES, with an increase in the JES resulting in a shift of the peak toward higher values. Figure 8.7b compares the m_{jj} distributions for $t\bar{t}$ signal events generated with three different values of m_t, demonstrating no clear dependence of m_{jj} on m_t.



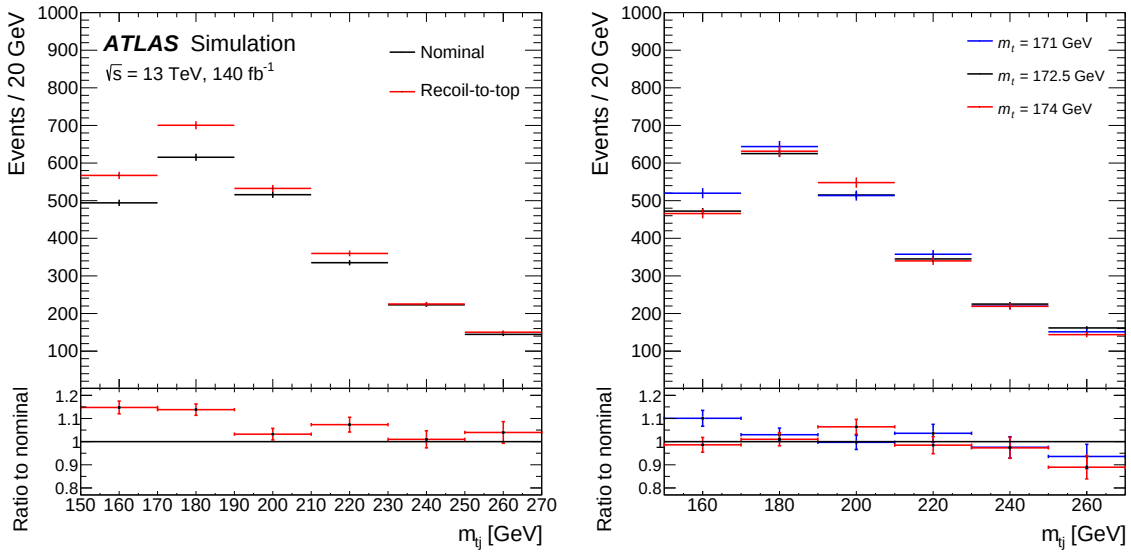
(a) **(b)**

Figure 8.7: The impact of varying (a) the first eigenvector component of the JES systematic uncertainties relating to the modelling of the simulation used for the in-situ JES calibration, and (b) the value of m_t on the (black) nominal $t\bar{t}$ prediction for m_{jj}. Both the impacts of (red) increasing and (blue) decreasing the variations are included. The lower panels show the ratio of the variations to the nominal. The nominal $t\bar{t}$ sample has $m_t = 172.5$ GeV.

Invariant mass of the leptonically decaying top quark and closest additional jet

The third and final observable included in the measurement of m_t is the invariant mass of the leptonically decaying top quark and the closest additional jet in ΔR, denoted as m_{tj}. This observable is introduced to reduce the uncertainty associated with the choice of recoil scheme used in the MC simulation. As discussed in Section 8.4, changing the recoil scheme changes the amount of wide-angle QCD radiation from the b-quark in the top-quark decay. As a result, for the case of more wide-angle radiation when using the alternative `recoil-to-top` recoil scheme, a larger number of events are expected to have an additional jet reconstructed close to the b-tagged jet originating from the decay of the leptonic top quark. Therefore, events used to construct this observable

are required to include at least one additional jet within $\Delta R < 0.5$ of the leptonically decaying top quark, and are required to satisfy $150 < m_{tj} < 270$ GeV. If multiple jets satisfy this condition, the jet with the highest p_{T} is used to calculate m_{tj}. Figure 8.8a shows the distribution of m_{tj} for the nominal $t\bar{t}$ signal prediction, compared with the prediction using the alternative recoil scheme. From Figure 8.8a, it is clear that there is an enhancement in the number of events around the m_t peak, as expected. The leptonic top-quark decay is used to minimise the correlation with the hadronic side of the event in the m_t measurement. Figure 8.8b compares the m_{tj} distributions for signal $t\bar{t}$ events generated with three different values of m_t, illustrating that there is no clear dependence of m_{tj} on m_t. Including m_{tj} in the fit model used to determine m_t (see Chapter 9) allows the model to differentiate between changes caused by changing the recoil scheme in the simulation and those that only affect m_t. As a result, the uncertainty arising from the choice of recoil scheme can be reduced.



(a) **(b)**

Figure 8.8: Comparison between distributions of m_{tj} for different signal $t\bar{t}$ simulated events generated using (a) the `recoilToColored` (nominal) and `recoil-to-top` recoil schemes, and (b) the `recoilToColored` scheme for predictions using different values of m_t. The nominal $t\bar{t}$ sample has $m_t = 172.5$ GeV.

Observed event yields

A summary of the additional requirements applied when using each observable is provided in Table 8.2. The event yields for the three selections in the measurement, for

both signal and background processes, are shown in Table 8.3. The data yields for the three selections are also presented, and overall the yields demonstrate that the selected events exhibit very high signal purity with minimal background contributions. As expected, single-top is the largest background source. The approximately 15% offset between the total prediction and data is consistent with the results observed in the ATLAS $t\bar{t}$ differential cross-section analysis using events with boosted top quarks [151]. Furthermore, this is also consistent with the fact that NNLO corrections produce a softer top-quark p_{T} spectrum, resulting in fewer events with top-quarks exhibiting high-p_{T} compared to the NLO prediction used in the MC simulation [151, 238, 239].

Table 8.2: Summarising the additional criteria to that in Table 8.1 when using $\overline{m_J}$, m_{jj}, and m_{tj}.

Observable	Additional criteria
$\overline{m_J}$	$145 < m_J < 205$ GeV
m_{jj}	Number of non-b-tagged constituent jets in top-jet ≥ 2 $60 < m_{jj} < 105$ GeV
m_{tj}	$\Delta R(\mathrm{t}^{\mathrm{lep}}, \mathrm{j}) < 0.5$ $150 < m_{tj} < 270$ GeV

It is important to note that the events used to construct $\overline{m_J}$, m_{jj}, and m_{tj} are not orthogonal nor completely overlapping. For example, events that cannot reconstruct the W boson inside the top-jet can still be used to compute $\overline{m_J}$, provided they fall within the mass window used for the calculation of the distribution's mean value. Similarly, events may fall outside this mass window for the $\overline{m_J}$ calculation but still pass the criteria for the m_{jj} selection. The requirement for an additional jet in the construction of m_{tj} means that the sample of events used to build m_{tj} is smaller compared to the $\overline{m_J}$ and m_{jj} selections, however, statistical overlap is still possible. Details of the overlap between events passing these observable selections are shown in Figure 8.9. Approximately 53% of the events used in the $\overline{m_J}$ observable are also included in the m_{jj} selection, whereas only about 3% of the events are shared with the m_{tj} observable.

Table 8.3: The observed number of data events is compared with the expectation from the signal and background processes for the selections for the m_J, m_{jj} and m_{tj} observables. The uncertainties on the expectations include all systematic uncertainties described in Section 8.4.

Process	m_J	m_{jj}	m_{tj}
$t\bar{t}$	65000 ± 9500	38100 ± 5800	2340 ± 430
Single-top	1000 ± 170	400 ± 130	50 ± 9
$t\bar{t}V$	700 ± 130	360 ± 67	54 ± 10
Multijet	400 ± 260	170 ± 110	23 ± 15
W+jets	250 ± 100	59 ± 22	21 ± 9
Z + jets	49 ± 24	10 ± 5	5 ± 3
Diboson	22 ± 11	5 ± 3	2 ± 1
Total prediction	67400 ± 9500	39100 ± 5800	2500 ± 430
Data	57459	32722	2312

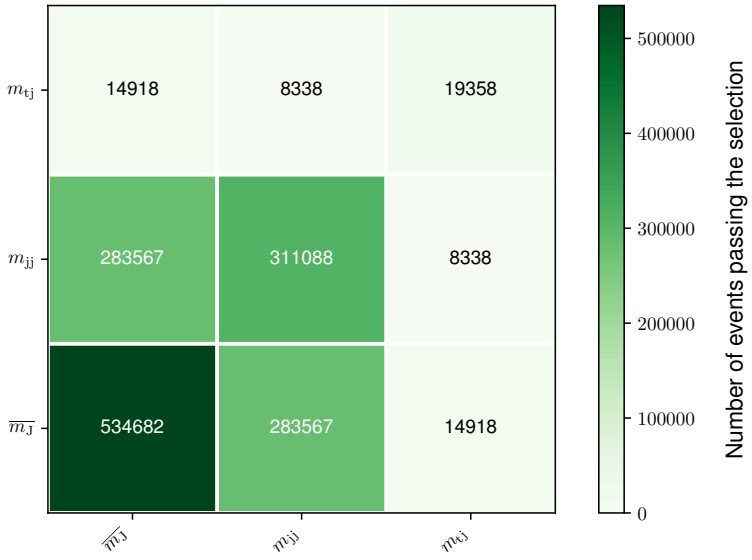


Figure 8.9: Number of nominal $t\bar{t}$ MC events that pass the additional criteria to build the observables used in this measurement. The diagonal terms correspond to the total number of events used to construct each observable, and off-diagonal terms denote the overlap of events.

CHAPTER 9

Profile likelihood fit

This chapter introduces the profile likelihood fitting method and explains how it is used to measure the value of m_t in data, utilising events that pass the requirements discussed in Section 8.3, while incorporating the sources of uncertainty outlined in Section 8.4. The contributions of these systematic uncertainties to the precision of the m_t result are quantified using the covariance matrix from the profile likelihood fit. The quality of the fit is then assessed through goodness-of-fit tests, using either the χ^2 test statistic or a saturated model. Validation studies are conducted to ensure the profile likelihood fit methodology is valid for measuring m_t in real ATLAS data. A simulated dataset built from the SM prediction is then used to investigate the expected precision of the m_t measurement.

9.1 Generic fit method

Consider the goal of determining m_t from an observable x. Suppose a dataset of N independent and identically distributed measurements of x is available, denoted as $D = (x_1, x_2, \ldots, x_N)$, where the measurements are distributed according to a known probability density function, $p(x|\, m_t, \vec{\theta})$. This function depends on the Parameter of Interest (POI), which is m_t, as well as additional NPs, denoted by $\vec{\theta}$, that are introduced to model unknown aspects of the dataset. These NPs account for systematic uncertainties arising from limitations in the precision of both experimental procedures and theoretical predictions. Each NP, θ_s, corresponds to a specific source of systematic uncertainty, s, and is associated with a measured value β_s. Each measurement of each NP follows a probability density function, $p\left(\beta_s \mid \theta_s, 1\right)$, which is assumed to be a

Gaussian defined such that β_s is equal to zero and the Gaussian has a width equal to one. Accordingly, each NP, has a *pre-fit value* equal to zero and *pre-fit uncertainty* equal to one.

The *likelihood* is then defined as the product of the probability density functions described above, and quantifies how well the statistical model describes the observed dataset. The values of the POI and the NPs are estimated by maximising the likelihood, expressed as [240]:

$$L\left(D\middle|m_t,\vec{\theta}\right) = \prod_i p\left(x_i\middle|m_t,\vec{\theta}\right) \times \prod_s G\left(\beta_s \mid \theta_s, 1\right). \tag{9.1}$$

The likelihood function is often referred to as a *model*. Maximising the likelihood yields the values of m_t and $\vec{\theta}$ that provide the best agreement between the model and the dataset. These values are denoted by $\hat{m}_t$ and $\hat{\vec{\theta}}$, respectively, and are referred to as the *best-fit* values of the parameters.

Rather than maximising the likelihood directly, it is often more practical to work with the negative logarithm of the likelihood, $-\log L$. This transformation converts the product of probability density function terms into a summation of logarithmic terms. In this formulation, maximising L is equivalent to minimising $-\log L$. While this minimisation can be performed analytically in simple cases, in most realistic scenarios, numerical techniques are required to minimise $-\log L$ effectively.

When performing this minimisation for a given dataset, the final *post-fit value* of a NP may differ from its pre-fit value. In such cases, the NP is said to exhibit a *pull*. This typically occurs when a non-zero value of the NP improves the agreement between the model and the observed data. The model can constrain the NPs because the individual measurements, x_i, depend on the set of NPs. This is only achievable if the dataset contains more data points than the number of POIs in the model. For instance, if the dataset consists of only a single data point, the value of m_t may still be estimated through the minimisation of $-\log L$. However, the model lacks sufficient information to constrain the NPs in this scenario. To effectively constrain the NPs, additional measurements need to be incorporated into the model

In many cases in particle physics, the analytical form of the probability density function for a given observable is not known. A common approach in such scenarios is to construct a binned histogram of the observable, transforming the dataset into $D = (n_1, n_2, \ldots, n_N)$, where n_i denotes the number of events in bin i of a histogram

consisting of N bins. The number of observed events in each bin is assumed to follow a Poisson distribution, with a mean that depends on the model parameters. The likelihood function is then constructed using template histograms that represent the nominal distribution of the observable (i.e., without any systematic variations) as well as the corresponding $\pm 1\sigma$ variations arising from systematic effects. Here, σ denotes the pre-fit uncertainty associated with a given NP representing a specific source of systematic uncertainty. The influence of a systematic uncertainty can affect the normalisation of the distribution, its shape, or both.

A useful quantity to define is the *profile likelihood ratio* [240] defined via:

$$\lambda(D \mid m_t) = \frac{L\left(D \middle| m_t, \hat{\hat{\vec{\theta}}}\right)}{L\left(D \middle| \hat{m}_t, \hat{\vec{\theta}}\right)}, \tag{9.2}$$

where $\hat{\hat{\vec{\theta}}}$ represents the conditional maximum likelihood estimate of $\vec{\theta}$ for a fixed value of m_t, i.e., the values of the NPs that maximise the likelihood for a given m_t. The denominator of Equation 9.2 corresponds to the maximum of the likelihood function, evaluated at the best-fit values of the model parameters. By construction, the profile likelihood ratio satisfies $0 \leq \lambda(D|\, m_t) \leq 1$, with values close to one indicating that the model with a given m_t provides a good description of the data. A commonly used transformation of this quantity is $-2\ln\lambda(D|\, m_t)$, which defines a function with a minimum at the point of best agreement between the model and the data. According to Wilks' theorem [240], the distribution of $-2\ln\lambda(D|\, m_t)$ asymptotically follows a χ^2 distribution in terms of the POI. The minimum of this distribution occurs at $m_t = \hat{m}_t$, where $-2\ln\lambda(D|\, m_t) = 0$. This property enables one to determine both the best-fit value of m_t and its associated uncertainty by identifying the values of m_t at which the distribution intersects the horizontal line defined by $-2\ln\lambda(D|\, m_t) = 1$. A fit based on this approach is referred to as a *profile likelihood fit*. As discussed in Section 9.4, the uncertainties on the model parameters used to measure m_t are derived from the covariance matrix obtained from this fit.

9.2 Goodness-of-fit tests

The agreement between two distributions can be quantified using a *goodness-of-fit* test. For the m_t measurement discussed in this thesis, goodness-of-fit tests are performed using either the χ^2 test statistic or the saturated model.

9.2.1 Using the χ^2 statistic

The χ^2 statistic is employed to assess the agreement between data and prediction in validation regions using the model described in Section 9.3, as discussed in Section 10.2. The general definition for the χ^2 statistic can be written in terms of the covariance matrix, C, as:

$$\chi^2 = \sum_{ij} (d_i - \nu_i)\, C_{ij}^{-1} \left(d_j - \nu_j\right), \tag{9.3}$$

where d_i is the measured value in data, ν_i is the predicted value, and C_{ij}^{-1} is (i,j)-th element of the inverse of the covariance matrix that encodes the uncertainties on the measurements. A small χ^2 value indicates that the measured and true distributions are in good agreement relative to the uncertainties. A large value indicates that the distributions do not agree, possibly because the uncertainties are underestimated.

Using the value of χ^2 and the number of degrees of freedom, N_{dof}, which is the difference between the number of measurements (or data points) and the number of free parameters in the model, the probability that the data is consistent with the prediction can be calculated as:

$$p = 1 - F\left(\chi^2, N_{\text{dof}}\right), \tag{9.4}$$

where $F\left(\chi^2, N_{\text{dof}}\right)$ is the cumulative distribution function of the χ^2 distribution.

9.2.2 Using the saturated model

A complication arises when attempting to use the χ^2 test statistic, as described in Section 9.2.1, to assess the quality of a profile likelihood fit to data. Specifically, N_{dof} is

not well defined because the NPs are neither entirely free parameters nor strictly fixed in the model. Consequently, an alternative method is required.

Another way to assess the quality of a fit to data is by using a goodness-of-fit test based on the *saturated model* [241]. This saturated model allows for perfect description of the data by introducing an additional free parameter for each data point in the dataset. The test statistic, $\lambda_{\text{saturated model}}$, is then defined as the ratio of the likelihood obtained from the minimisation using the original *nominal model* to that obtained using the saturated model, as follows:

$$\lambda_{\text{saturated model}} = -2\ln\left(\frac{L_{\text{nominal model}}}{L_{\text{saturated model}}}\right). \tag{9.5}$$

This ratio asymptotically follows a χ^2 distribution, as stated by Wilks' theorem [242], and can be used analogously to a standard χ^2 test to evaluate the fit probability (see Section 9.2.1). In this context, N_{dof} is defined as the difference in the number of free parameters between the saturated model and the nominal model.

9.3 Profile likelihood fit to measure the top-quark mass

As described in Section 8.5, $\overline{m_J}$ is chosen as the observable to measure m_t. Before constructing the likelihood function using $\overline{m_J}$, it is important to parametrise how $\overline{m_J}$ depends on m_t. From Figure 8.6, it is clear that a larger value of m_t shifts the peak of the distribution to higher values. Therefore, the simplest functional form for $\overline{m_J}$ in terms of m_t is the linear relationship:

$$\overline{m_J}(m_t) = A + B\,(m_t - 172.5)\,, \tag{9.6}$$

where the constant coefficients A and B are determined through a fit to $\overline{m_J}$ values obtained from five pseudo-data samples, constructed from the m_t-varied $t\bar{t}$ samples and the background samples discussed in Section 8.2. This fit is shown in Figure 9.1, where A and B are found to have values of (172.45 ± 0.01) GeV and 0.46 ± 0.02, respectively. The uncertainties on these coefficients refer to the statistical uncertainties returned from the fit. From Figure 9.1, it is clear that a linear relationship is sufficient to describe the dependence of $\overline{m_J}$ on m_t.

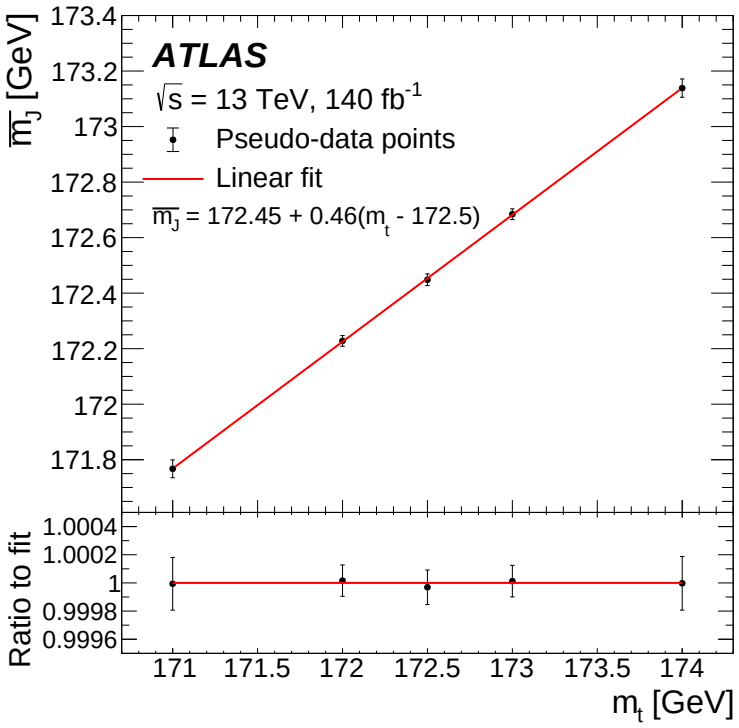


Figure 9.1: The fit of Equation 9.6 to the $\overline{m_J}$ values for the $t\bar{t}$ MC samples with varied values of m_t. Simulations of background processes are included in the determination of $\overline{m_J}$ at each data point. The error bars relate to the MC statistical error.

To account for the impact of the systematic uncertainties on $\overline{m_J}$, an additional term is included that incorporates the NPs and their impact given a one standard deviation change of a given NP, denoted by Δ_s. An additional term is also included to allow for a dependence on μ, a normalisation scale factor defined as the ratio of the measured cross-section to the SM prediction for the $t\bar{t}$ production cross-section used to normalise the signal $t\bar{t}$ simulation (see Section 8.2.2). Introducing this dependence allows for μ to be included as an additional POI in the model. This approach ensures that the value of m_t extracted from the profile likelihood fit remains independent of the SM prediction for the $t\bar{t}$ production cross-section. Such treatment is necessary due to the presence of background events in the dataset. Putting all of this together, $\overline{m_J}$ is parametrised as:

$$\overline{m_J}\left(m_t, \mu, \vec{\theta}\right) = A + B\left(m_t - 172.5\right) + C\left(\mu - 1\right) + \prod_s \theta_s \Delta_s, \tag{9.7}$$

where the value of the constant coefficient C is determined from the change in $\overline{m_J}$ observed when μ is varied by $\pm 10\%$ in the simulation. A value of $C = 0.02$ GeV is

obtained[1]. A value of $C = 0$ would be observed if no background contributions were present in the analysis, and the small value obtained demonstrates how the selected events have a very high signal purity, as reflected in Table 8.3.

It is assumed that measurements of $\overline{m_J}$ follow a Gaussian probability density function, allowing the likelihood to be constructed as a single Gaussian term, as the dataset consists of a single data point, $\overline{m_J}^{\text{data}}$. The probability density function describing this observable also depends on a set of NPs that account for the systematic uncertainties outlined in Section 8.4. However, with only one data point, there is insufficient information to constrain or pull the NPs. While this is not inherently problematic, achieving a precise measurement of m_t requires NPs associated to leading systematic uncertainties to be constrained. To improve the precision, the likelihood is extended by incorporating the two additional observables defined in Section 8.5: m_{jj} and m_{tj}. Measurements of these observables are represented as binned histograms, where the number of events in each bin follows a Poisson distribution with a mean that depends on the model parameters. The likelihood that is used to measure m_t in $t\bar{t}$ events with a boosted top quark is given by:

$$\begin{aligned} L\left(\overline{m_J}^{\text{data}}, \vec{n}_{m_{jj}}, \vec{n}_{m_{tj}} \middle| m_t, \mu, \vec{\theta}\right) = \; & L_{\overline{m_J}}(\overline{m_J}^{\text{data}} \mid m_t, \mu, \vec{\theta}) \\ & \times L_{m_{jj}}(\vec{n}_{m_{jj}} \mid \mu, \vec{\theta}) \\ & \times L_{m_{tj}}(\vec{n}_{m_{tj}} \mid \mu, \vec{\theta}) \\ & \times \prod_s G(\beta_s \mid \theta_s, 1), \end{aligned} \tag{9.8}$$

where $L_{\overline{m_J}}$, $L_{m_{jj}}$, $L_{m_{tj}}$ denote the likelihood functions for the $\overline{m_J}$, m_{jj}, and m_{tj} observables, respectively, given by:

$$L_{\overline{m_J}}(\overline{m_J}^{\text{data}} \mid \mu, \vec{\theta}) = G\left[\overline{m_J}^{\text{data}} \middle| \overline{m_J}(m_t, \mu, \vec{\theta}), \sigma_{\overline{m_J}}\right], \tag{9.9}$$

$$L_{m_{jj}}(\vec{n}_{m_{jj}} \mid \mu, \vec{\theta}) = \prod_j P(n_{m_{jj},j} \mid \nu_j(\mu, \vec{\theta})), \text{ and} \tag{9.10}$$

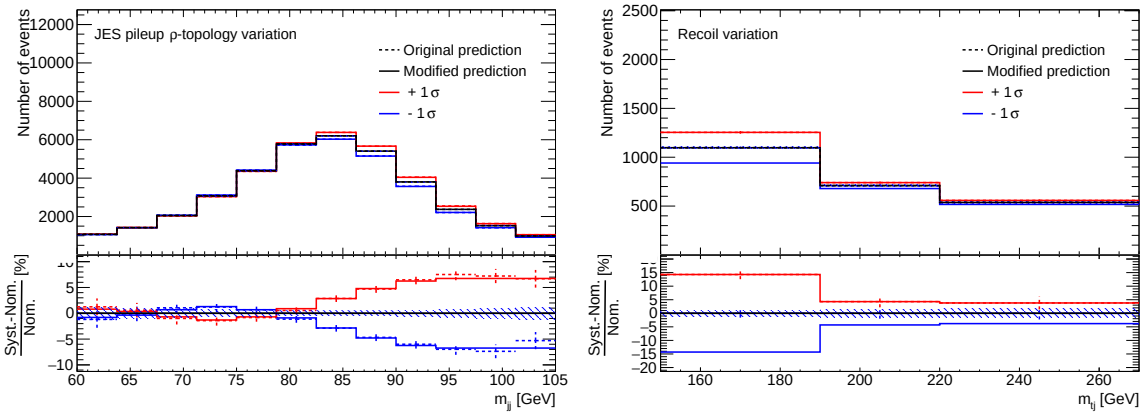
[1]The same events are present in both distributions used to compare the impact of making a 10% change in μ, meaning they are 100% correlated. As a consequence, no uncertainty is quoted on the value of C.

$$L_{m_{tj}}(\vec{n}_{m_{tj}} \mid \mu, \vec{\theta}) = \prod_k P(n_{m_{tj},k} \mid \rho_k(\mu, \vec{\theta})), \tag{9.11}$$

where $\sigma_{\overline{m_J}}$ is the statistical uncertainty on $\overline{m_J}$, $n_{m_{jj},i}$ denotes the number of entries in bin i of the m_{jj} distribution for data, ν_i is the expected number of entries in that bin, $n_{m_{tj},k}$ denotes the number of entries in bin k of the m_{tj} distribution for data, and ρ_k is the expected number of entries in that bin. Twelve bins are chosen to describe the m_{jj} distribution and three bins are used for m_{tj}. In this way, the distributions are sufficiently sensitive to the systematic uncertainties, with enough events in each bin to avoid issues from statistical uncertainties in the simulation. From Figures 8.7b and 8.8b, there is little visual dependence of m_{jj} or m_{tj} on m_t. Furthermore, the use of fewer bins with larger bin widths ensures m_{tj} is sufficiently insensitive to m_t. Consequently, ν_i and ρ_k depend on μ and the NPs, but not on m_t. This allows these additional observables to constrain the NPs, thereby improving the precision of the fitted m_t. This likelihood function assumes that the bins in the m_{jj} and m_{tj} distributions and the value of $\overline{m_J}$ can be treated as statistically uncorrelated, even though the same events can contribute to all three observables as mentioned in Section 8.5. A validation study confirms that this assumption is appropriate (see Section 9.5).

The impacts of the NPs on ν_{j} and ρ_{k} are evaluated using histogram templates corresponding to $\pm 1\sigma$, where σ represents the pre-fit uncertainty for a given NP. These are referred to as the *up* and *down* variations of a NP, respectively. The impact of the systematic uncertainties is assumed to be symmetric in each bin, and the up and down variations of each NP are correspondingly symmetrised. This symmetrisation is performed by averaging the difference between the up and down variations, keeping track of the sign of the effect. In rare cases where both up and down variation have the same sign, the uncertainty is taken as the maximum magnitude of the variations. The sign is positive if the up variation has the largest magnitude and positive sign, or if the down variation has the largest magnitude and negative sign. The sign is negative for the other two possibilities. In some cases, explicit up and down variations for a NP do not exist. Instead they are assessed by comparing the nominal prediction with an alternative prediction. For this situation, the variation from this comparison is defined as the up variation and is then symmetrised to create a down variation. An automatic smoothing algorithm [243] is applied to the systematic variations of ν_{j}, such that the model is less sensitive to statistical fluctuations in the m_{jj} templates. No smoothing algorithm is needed for the three bins of the m_{tj} distribution. As discussed in Section 8.5, m_{jj} and m_{tj} are sensitive to the effects of changing the JES

and recoil scheme, respectively. Figure 9.2 shows the impact of varying one of the NPs associated to the JES systematic uncertainties on m_{jj}, as well as varying the NP associated with the uncertainty due to the choice of recoil scheme in the simulation on m_{tj}. From Figure 9.2, it is clear that these systematic uncertainties have large impacts on these observables for the signal $t\bar{t}$ events in simulation, and the automatic smoothing algorithm is working to provide a smooth shape in the variation for m_{jj}. The effects of the up and down NP variations are symmetrised for the impact of the chosen JES NP on m_{jj}, and the up variation of the recoil uncertainty NP on m_{tj} is symmetrised to define the down variation.



(a) **(b)**

Figure 9.2: The effect of varying (a) one of the NPs relating to the systematic uncertainties in the pileup correction of the JES calibration on m_{jj}, and (b) the NP associated with the systematic uncertainty due to the choice of recoil scheme on m_{tj} for (black) the signal $t\bar{t}$ prediction and the associated NP varied by (red) $+1\sigma$ and (blue) -1σ, where σ represents the pre-fit error for that parameter. Smoothing is applied to the original prediction for m_{jj}. The original variations are shown by the dashed lines, and the solid lines represent these variations following smoothing and/or symmetrisation.

An automatic procedure is used to remove the impact of systematic uncertainties which have an effect below a given threshold. The effect of a NP is split into a shape effect and a normalisation effect. If for all bins the magnitude of the difference between the nominal yield and that of the systematic variation divided by the nominal yield is less than 0.01%, the shape effect is dropped. If the nominal and systematic variation yields are within 0.1% of each other, the normalisation effect is dropped. In some cases, both the shape and normalisation effects are dropped and the NP is said to be *pruned*, and it is removed from the model. Following this procedure, there are 133 NPs in the model using all three observables.

Every NP is implemented such that they each have effects correlated between the three observables used in the model, except for that corresponding to the uncertainty due to the choice of parton shower and hadronisation model. This uncertainty is chosen to be decorrelated between the observables, meaning it is quantified using three NPs, one for each observable. This is done to prevent a large constraint on the NP that is assessed by the comparison of the m_{jj} distributions built using POWHEG+PYTHIA8 and POWHEG+HERWIG7 from being applied to the other observables (see Section 9.6).

The complete fit model includes terms for the limited number of MC events. For the m_{jj} and m_{tj} distributions, these are an additional NP for each bin, where the constraint term is a Poisson distribution corresponding to the total MC statistical uncertainty in that bin. The uncertainty due to the limited number of MC events on the $\overline{m_J}$ is assumed to be negligible. Any impact from the limited event yields in the samples with alternate m_t values used to derive the parametrisation of Equation 9.6 is addressed through the linearity test described in Section 9.5.

An important aspect of the model defined in Equation 9.8 is that the fitted values of the NPs and μ are determined entirely by the m_{jj} and m_{tj} observables. Since there is only a single measurement of $\overline{m_J}$ in a given dataset, and this is the sole component of the model that depends on m_t, the structure of the fit remains largely unchanged whether m_{jj} and m_{tj} are fitted alone or in combination with $\overline{m_J}$. The inclusion of the $\overline{m_J}$ observable introduces just one additional data point and one additional free parameter, m_t. Consequently, the fitted values of the NPs and μ remain unaffected by the inclusion of $\overline{m_J}$. This allows for the evaluation of the NPs independently of the measurement of $\overline{m_J}$ or m_t in data, as discussed in Section 10.1.

The profile likelihood fit results presented in this thesis are obtained by minimising $-2\ln\lambda$, where λ is the profile likelihood ratio defined in Equation 9.2, and the likelihood is given by Equation 9.8. The model is constructed using the `TRExFitter` package [244] for the m_{jj} and m_{tj} observables, that is then combined with $\overline{m_J}$ constructed using the RooFit framework [245]. RooFit then performs the profile likelihood fit, utilising the `Minuit` numerical minimisation algorithms [246]. For goodness-of-fit tests, the `TRExFitter` package offers the functionality to compute both the χ^2 and saturated model test statistics, as well as the fit probability.

9.4 Quantifying uncertainties

To evaluate the total uncertainty on the measured m_t value, and quantify the contributions from the systematic uncertainties, it is possible to directly use elements of the covariance matrix determined from the profile likelihood fit, as discussed in Ref. [247]. This matrix encodes the covariance between the individual NPs, as well as the covariance between the NPs and the POIs, m_t and μ. The elements of the covariance matrix, C_{ij}, are numerically estimated using the `Minuit` program [246] as the inverse of the second-order partial derivative matrix of the negative logarithm of the likelihood function [242], such that:

$$C_{ij} = \left[-\frac{\partial^2 \log L}{\partial\theta_i \partial\theta_j}\Big|_{\vec{\theta}=\hat{\vec{\theta}}} \right]^{-1} . \tag{9.12}$$

For the model described in Section 9.3, the effect of each NP on m_t is obtained directly from the off-diagonal elements of the covariance matrix, $\Delta_i = C_{i,m}$, where i represents the considered NP and m refers to the row/column of the matrix corresponding to m_t. The total systematic uncertainty on m_t is given by the quadratic sum of the individual contributions. The total uncertainty on m_t is identical to $C_{m,m}$, while the statistical uncertainty on m_t is determined by the quadratic difference between the the total uncertainty and the total systematic uncertainty. To evaluate the impact of groups of systematics (e.g. JER systematics), the quadratic sum of all sources within the group is calculated.

9.5 Validation of the profile likelihood fit model

A series of tests are conducted using simulated events to validate the assumptions made during the development of the model described in Section 9.3. These tests include verifying the assumption that treating the observables used in the model as statistically independent is valid, even though there is some degree of statistical correlation as these observables are not built using orthogonal sets of data (see Section 8.5). Another test is to evaluate the level of closure in the fit model, checking the linear dependence of $\overline{m_J}$ on m_t, as well as any potential dependence of m_{jj} and m_{tj} on m_t. The final tests consisted of injecting pseudo-data, built from $t\bar{t}$ MC samples with alternative generator setups, to assess the robustness of the model.

9.5.1 Test of statistical correlations between observables

The fit model described in Section 9.3 includes the product of three likelihoods for the selected events: one representing $\overline{m_J}$ and two representing the binned event yields of m_{jj} and m_{tj}. This setup implicitly assumes that these observables are independent; however, Figure 8.9 shows that there is statistical overlap between them.

To compute the statistical correlations between the observables, a 3D histogram is created for the nominal signal + background prediction of m_J, m_{jj}, and m_{tj}. Pseudo-data is then generated by fluctuating the yield in each bin of this histogram according to a Poisson distribution, with the mean equal to the nominal prediction for each bin. This process is repeated to generate 100,000 sets of pseudo-data. Using these pseudo-data sets, the statistical correlations between each bin are calculated, resulting in the matrix shown in Figure 9.3. From Figure 9.3, it is clear that the statistical correlations between the observables are small. Furthermore, the sign of the correlation between $\overline{m_J}$ and the bins of the m_{jj} distribution is as expected, with higher mass bins showing a positive correlation and lower mass bins showing a negative correlation. As expected, the correlation between the bins of m_{tj} and $\overline{m_J}$ is very small, as these observables do not target the same top quark in the $t\bar{t}$ decay and the overlap in the selected events is small. To verify the validity of the assumption that these observables can be treated as statistically independent, it must be tested whether accounting for these statistical correlations would affect the m_t result.

To test whether the small statistical correlations are significant, pseudo-experiments are performed. In the first set of pseudo-experiments, 1D histograms of m_J, m_{jj}, and m_{tj} are constructed from the nominal signal + background prediction, and the yield in each bin of the m_J, m_{jj}, and m_{tj} distributions is fluctuated according to a Poisson distribution based on the expected number of events. In the second set of pseudo-experiments, a 3D histogram of m_J, m_{jj}, and m_{tj} is created from the nominal signal + background prediction, and each bin of the 3D distribution is fluctuated according to a Poisson distribution for the expected number of events in each bin. This second set of pseudo-experiments properly accounts for the correlations between the observables, while the first set neglects these correlations. Each pseudo-experiment is treated as a data sample, and the profile likelihood fit is performed, yielding a measured m_t for each pseudo-experiment. Any effect due to statistical correlations would manifest as a difference in the standard deviation of the measured m_t between the two sets of pseudo-experiments. Figure 9.4 shows the results of the two sets

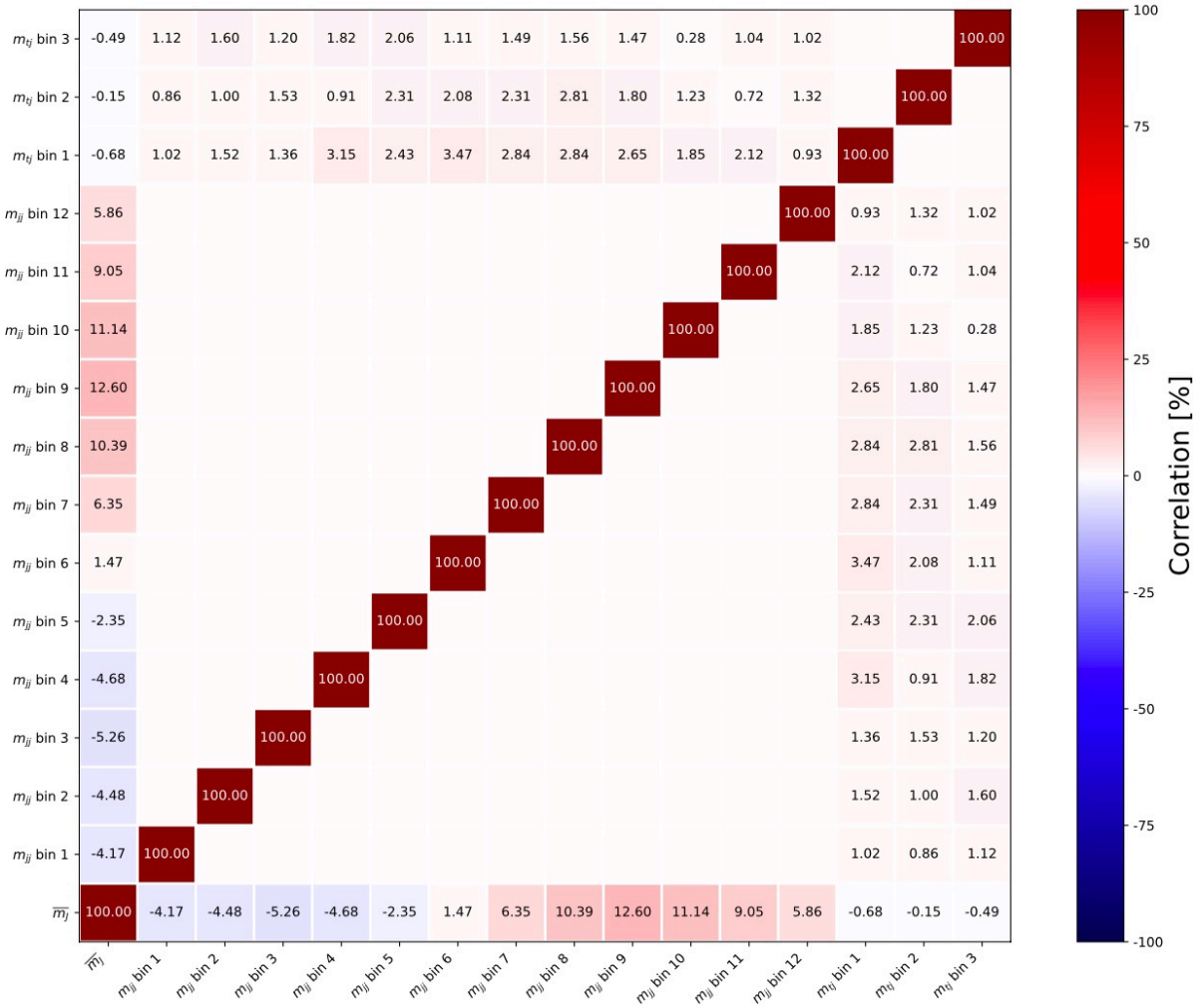


Figure 9.3: Statistical correlations between the observables used in the profile likelihood fit.

of pseudo-experiments, each consisting of 200 pseudo-experiments. The standard deviations of the two sets are found to match within the statistical precision of the test. Therefore, it is concluded that treating the observables as uncorrelated is a good approximation.

9.5.2 Test of fit linearity and closure

The linearity and closure of the profile likelihood fit are tested by injecting pseudo-data into the model and assessing the m_t result. Five pseudo-data samples are constructed from the m_t-varied $t\bar{t}$ samples, which are then added to the background samples. Each pseudo-data sample is treated as data, and the profile likelihood fit is performed. The extracted m_t value is compared to the simulated value. Injecting the pseudo-data

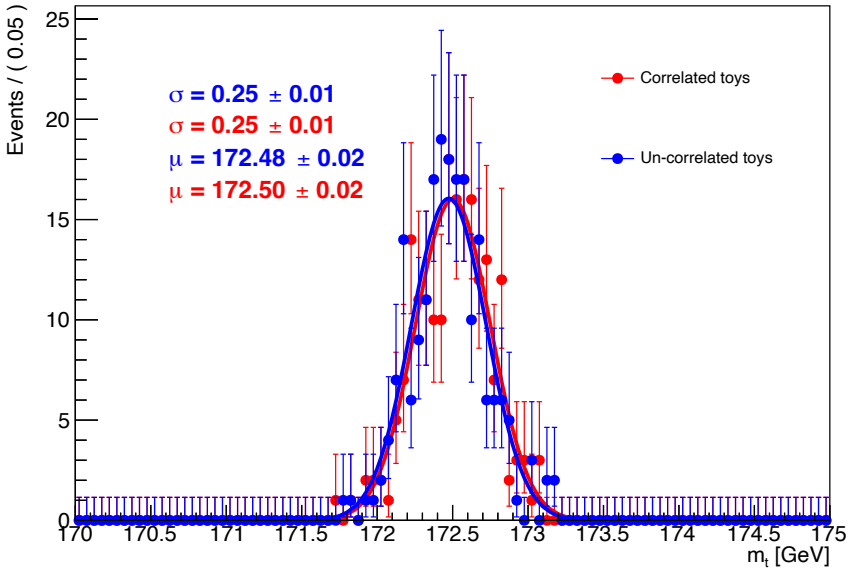


Figure 9.4: Distribution of the fitted m_t values extracted from the two sets of pseudo-experiments. The red (blue) points are the results from the pseudo-experiments that include (exclude) the effect of the statistical correlations. The lines are fits of Gaussian distributions to the two sets of points and the fitted means and standard deviations are displayed.

into the model setup tests whether m_{jj} and m_{tj} have negligible dependence on m_t and if there is any bias in the fit. Small deviations from perfect linearity are expected due to the limited number of events in the simulation. The results of the test are shown in Figure 9.5, where the fit is observed to have excellent linearity. A χ^2 test comparing the fit to the linear expectation of perfect linearity yields $\chi^2 = 4.8$, with a corresponding fit probability of 30%. Performing a linear fit to the residuals (r) results in $r = a + bm_t$, with $a = 40$ MeV and $b = 1.8 \times 10^{-4}$. Since negligible dependence on m_t is observed, a systematic uncertainty of 70 MeV, corresponding to the observed bias at $m_t = 172.5$ GeV, is incorporated into the fit model to account for potential non-closure.

An additional test is performed to verify that the fit model can recover m_t independently of the $t\bar{t}$ cross-section. For this test, pseudo-data samples are generated from the nominal signal + background prediction, with μ varied from 0.8 to 1.1. The factor of 0.856 is also tested, corresponding to the ratio of the measured to predicted production cross-sections observed in the ATLAS $t\bar{t}$ differential cross-section measurement using events with boosted top quarks [151]. Each sample is treated as data, and the profile likelihood fit is performed to extract the measured m_t and μ. The results, shown in

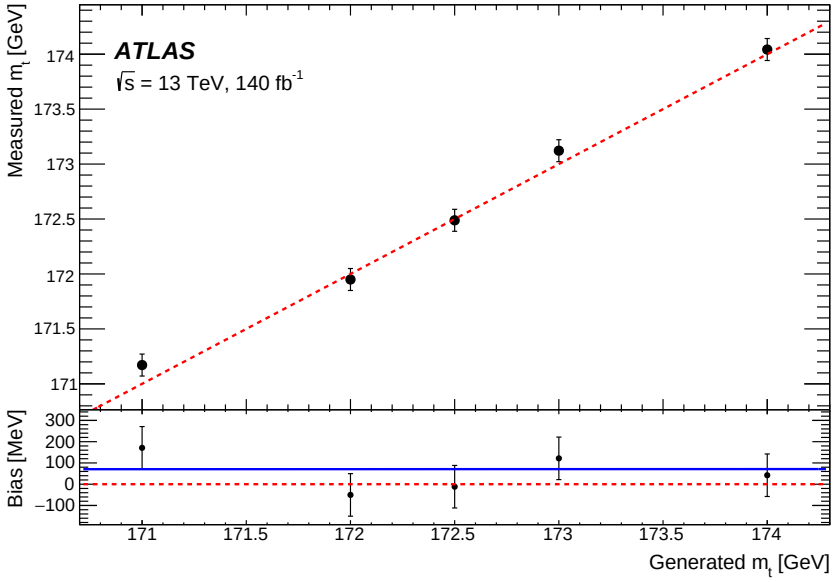
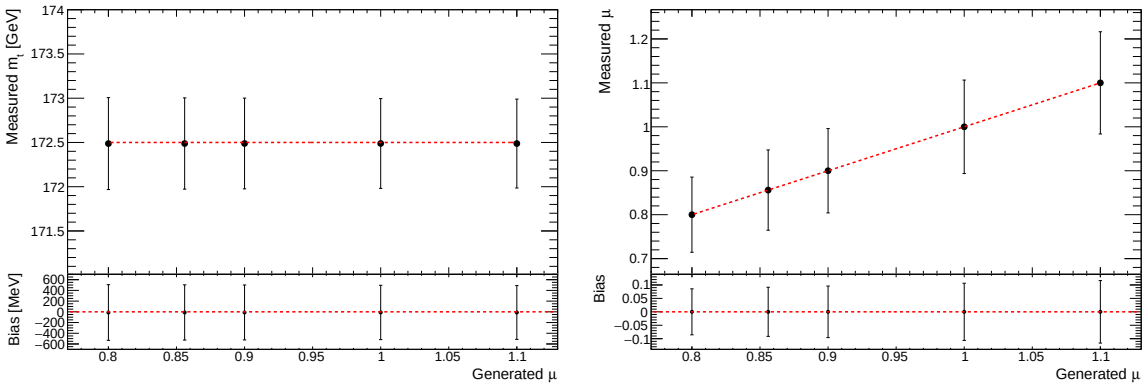


Figure 9.5: The measured m_t value is shown as a function of the input m_t value for the test. The uncertainty bars on each point show the uncertainty from MC statistics. The dashed red line shows perfect linearity and the blue line shows a linear fit to the points (excluding m_t = 172.5 GeV).

Figure 9.6, demonstrate that the extracted m_t is independent of the $t\bar{t}$ cross-section, and the fit is also able to measure μ in an unbiased way. Given these observations, no additional uncertainty is required to account for any dependence on the SM $t\bar{t}$ cross-section prediction.



(a) **(b)**

Figure 9.6: (a) The measured m_t value as a function of the input $t\bar{t}$ cross-section relative to the SM prediction for m_t = 172.5 GeV, and (b) the extracted cross-section relative to the SM. The uncertainty bars on each point show the expected uncertainties from the fit. The dashed red line shows perfect linearity.

9.5.3 Checking the robustness of the fit

One method to evaluate the robustness of the fit is to inject pseudo-data generated from simulated signal events produced with alternative MC setups not explicitly included in the fit model. If any observed bias in the fitted result is covered by the expected $t\bar{t}$ modelling uncertainty, which is defined as the quadrature sum of all individual modelling uncertainties evaluated in the 3D Asimov fit described in Section 9.6, this demonstrates that the fit can adapt to models of $t\bar{t}$ simulation that are not directly included. This is tested by injecting pseudo-data into the fit and comparing the extracted value of m_t to its known input. Two such tests are performed, both using pseudo-data constructed with the nominal background prediction. One test uses the SHERPA $t\bar{t}$ sample, while the other uses an alternative POWHEG+PYTHIA8 $t\bar{t}$ sample that includes resonance decays prior to CR (see Section 8.2.2). In both cases, the injected value of m_t is 172.5 GeV. The observed biases in the fitted m_t are 0.04 and 0.08 GeV for the pseudo-data based on the SHERPA and alternative POWHEG+PYTHIA8 samples, respectively. These deviations are covered by the total expected modelling uncertainty of 0.26 GeV, as determined from the profile likelihood fit using a dataset corresponding to the SM prediction (see Section 9.6). It can therefore be concluded that the fit model is sufficiently robust to handle datasets generated with signal models not explicitly included in the fit configuration.

9.6 Profile likelihood fit using an Asimov data set

An *Asimov dataset* is one such that, when used to evaluate the best-fit values of all model parameters, the generated parameter values are obtained [240]. For this study, an Asimov dataset is generated using the model defined in Section 9.3, based on the nominal $t\bar{t}$ and background samples described in Section 8.2.2. Consequently, the Asimov dataset represents the SM prediction for $m_t = 172.5$ GeV for all processes contributing to the m_t measurement. By applying the profile likelihood fit to this dataset, the expected precision on m_t can be evaluated. This procedure is referred to as an *Asimov fit*.

By construction, all NPs are expected to have post-fit values of zero in an Asimov fit, although their post-fit uncertainties may be smaller than their pre-fit values, reflecting expected constraints from the data. If the fitted central value of m_t does not agree with

the value used in the simulation, this would indicate a problem with the model setup or the fit minimisation procedure. The contributions of individual uncertainty sources to the total uncertainty on m_t are determined using elements of the covariance matrix from the nominal fit, as discussed in Section 9.4. For this study, the pre-fit value of μ is set to one.

In addition to evaluating the expected precision on m_t using the profile likelihood fit with the full model, which includes $\overline{m_J}$, m_{jj}, and m_{tj}, it is also important to assess the improvement in precision gained by incorporating the m_{jj} and m_{tj} observables. To do this, three separate Asimov fits are defined:

- 1D Asimov fit: Profile likelihood fit using the likelihood defined by Equation 9.8 with the $L_{m_{jj}}$ and $L_{m_{tj}}$ terms removed. This fit assesses the expected m_t precision when only fitting $\overline{m_J}$.
- 2D Asimov fit: Profile likelihood fit using the likelihood defined by Equation 9.8 with the $L_{m_{tj}}$ term removed. This fit assesses the expected m_t precision when fitting $\overline{m_J}$ and m_{jj}. This fit model is intended to have constrained NPs relating to the JES systematic uncertainties, with an improvement to the m_t precision over the 1D Asimov fit.
- 3D Asimov fit: Profile likelihood fit using the likelihood defined by Equation 9.8. This fit assesses the expected m_t precision when fitting all three observables. This fit model is intended to have constrained NPs relating to the systematic uncertainties associated with the JES calibration and the choice of recoil scheme in the simulation. Consequently, this is expected to give an improved m_t precision over the 2D Asimov fit.

Performing the three Asimov fits results in a fitted central m_t value of 172.49 GeV for all three fits with total expected uncertainties of 1.66, 0.63 and 0.51 GeV for the 1D, 2D, and 3D Asimov fits, respectively. The 3D Asimov fit result indicates that this measurement has the capability of reaching a higher precision than any other single direct m_t measurement performed by ATLAS. The breakdowns of the expected uncertainties contributing to the total uncertainties obtained from these three Asimov fits are shown in Table 9.1.

As discussed in Section 9.5, the results of the linearity validation test meant that an additional 70 MeV uncertainty is included. This uncertainty is shown in Table 9.1 for the 3D Asimov fit. Fit closure uncertainties for the 1D and 2D Asimov fit models are

Table 9.1: Breakdown of the categories of uncertainty sources contributing to the total expected uncertainty in m_t for the 1D, 2D, and 3D profile likelihood fit setups using Asimov data. The results were obtained using an m_t value of 172.5 GeV and $\mu = 1$. The contribution of each individual uncertainty component is obtained from the covariance of the corresponding NP and m_t. Each category's contribution is equal to the sum in quadrature of the individual uncertainty contributions. 'Radiation' refers to the sources relating to the modelling of ISR and FSR for $t\bar{t}$ events. Those associated with ISR are the setting of h_{damp} and the use of the Var3c eigentune of the PYTHIA8 A14 tune. 'Higher-order corrections' refers to sources corresponding to the μ_f and μ_r scales and the NNLO reweighting. Systematic uncertainties related to the modelling of background processes are represented by the 'Background modelling' category.

Source	Uncertainty (GeV)		
	1D fit	2D fit	3D fit
JES	± 1.41	± 0.30	± 0.26
Colour reconnection (CR1 and CR2)	± 0.05	± 0.18	± 0.15
Radiation (ISR and FSR)	± 0.82	± 0.08	± 0.14
JES heavy flavour	± 0.08	± 0.15	± 0.14
Parton shower and hadronisation model	± 0.10	± 0.10	± 0.13
JER	± 0.14	± 0.09	± 0.10
MC statistics	< 0.01	± 0.09	± 0.09
Underlying event	± 0.07	± 0.07	± 0.08
Recoil	± 0.10	± 0.36	± 0.08
Fit closure	< 0.01	± 0.07	± 0.07
Background modelling	± 0.05	± 0.05	± 0.05
Matrix element matching ($p_{\text{T}}^{\text{hard}} = 1$)	± 0.05	± 0.02	± 0.05
b-tagging	± 0.02	± 0.04	± 0.04
$E_{\text{T}}^{\text{miss}}$	± 0.01	± 0.01	± 0.02
Higher-order corrections	± 0.02	± 0.03	± 0.01
Pileup	± 0.15	± 0.02	± 0.01
JVT	± 0.01	± 0.01	± 0.01
PDF	± 0.01	± 0.01	± 0.01
Luminosity	< 0.01	< 0.01	< 0.01
Leptons	< 0.01	< 0.01	± 0.01
Total statistical	± 0.11	± 0.27	± 0.25
Total systematic	± 1.66	± 0.57	± 0.44
Total	± 1.66	± 0.63	± 0.51

also defined with values obtained from similar validation tests using their respective likelihood functions. These are also included in Table 9.1.

From Table 9.1, it is evident that the expected precision in m_t improves significantly due to the substantial reductions in the contributions relating to the JES and modelling of ISR and FSR. By including m_{jj} in the fit model, the JES contribution is reduced by 1.11 GeV, and the combined ISR and FSR uncertainty reduces by 0.74 GeV. Figure 9.7 illustrates the fitted values of the NPs and their post-fit uncertainties for those NPs included in the fit models (i.e. those that are not pruned) for the 2D and 3D Asimov fits. A similar overview of the NPs included in the 1D Asimov fit is not included, given that the model used for the 1D Asimov fit does not have sufficient information to constrain any of the NPs. Hence all NPs for the 1D Asimov fit must have a fitted value of zero and a post-fit uncertainty of one.

Figure 9.7 clearly shows that several NPs are constrained when the m_{jj} and m_{tj} observables are included in the model. This results in a reduction of the total uncertainty on m_t. Before proceeding to examine the profile likelihood fit using real ATLAS data, it is important to investigate the origin of these constraints.

One constrained NP is that associated to the JES uncertainty component relating to the uncertainty in the per-event p_{T} density modelling in the simulation. Figure 9.8 illustrates the impact of this NP on the nominal predictions for m_{jj} and m_{tj}. As expected, m_{jj} is sensitive to this NP, whereas the distribution of m_{tj} is less sensitive. This is reflected in the fact that the NP is constrained when including m_{jj} in the model, as shown in Figure 9.7. Note that the fit has no way of precisely distinguishing between the many different JES NPs, so the JES NPs themselves are not tightly constrained, but the overall JES effect is constrained.

From Table 9.1, it is clear that when performing the 2D Asimov fit, the leading uncertainty is no longer the JES contribution. Instead, the recoil uncertainty becomes dominant and limits the precision on m_t. This contribution is reduced by 0.28 GeV by introducing m_{tj} into the model. Given the post-fit uncertainty for the associated NP in Figure 9.7, this NP is significantly constrained by the fit, where the large constraint is only present when m_{tj} is included in the fit model. Figure 9.9 shows the impact of this NP variation on the m_{jj} and m_{tj} distributions for the nominal prediction. The figure shows a large shape impact on m_{tj}, whereas m_{jj} shows less sensitivity. However, the shape of the variation of this NP on m_{jj} is similar to other variations, in particular the JES with an impact in the high m_{jj} tail (see Figure 9.8). Consequently, just as the individual JES NPs are slightly constrained in the 2D Asimov fit, so is the recoil scheme NP. However, the very large and distinct shape of the recoil scheme NP variation in

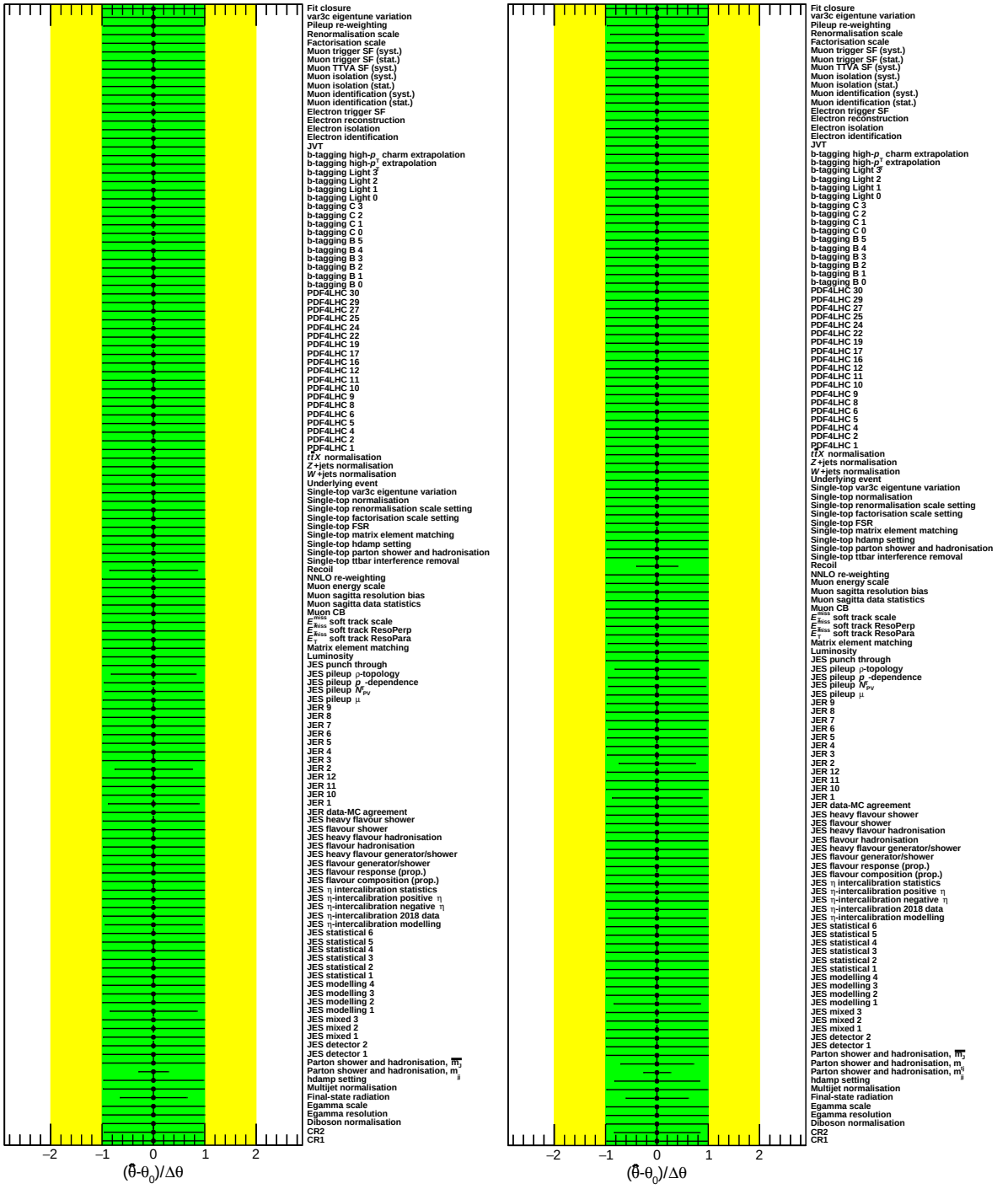
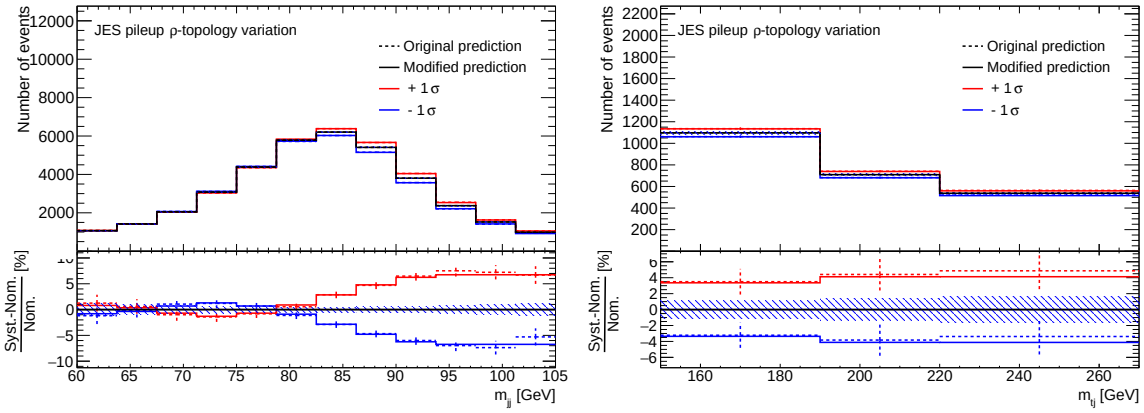


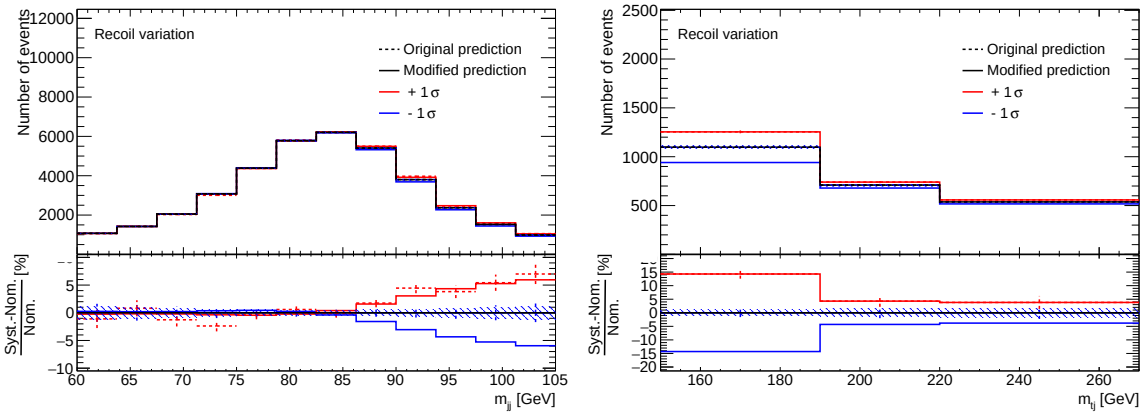
Figure 9.7: Post-fit values and uncertainties for the NPs included in the (a) 2D and (b) 3D Asimov fits. NPs associated with MC statistical uncertainties are not shown. The best fit and pre-fit values for the NP are given by $\hat{\theta}$ and θ_0, respectively. The pre-fit NP uncertainty is given by $\Delta\theta$.

m_{tj} means the majority of the large constraint observed for the 3D Asimov fit originates from the m_{tj} distribution.



(a) **(b)**

Figure 9.8: The effect of the NP relating to the component of the JES for the ρ-topology mismodelling for the pileup calibration on (a) m_{jj} and (b) m_{tj} for the (black) nominal prediction and where the associated NP is varied by (red) $+1\sigma$ and (blue) -1σ, where σ represents the pre-fit error for that parameter. Smoothing is applied to the original prediction for m_{jj}. The effect of the NP on both distributions is symmetrised. The original predictions are shown by the dashed lines, whereas the solid lines show the modified variations after smoothing and/or symmetrisation.



(a) **(b)**

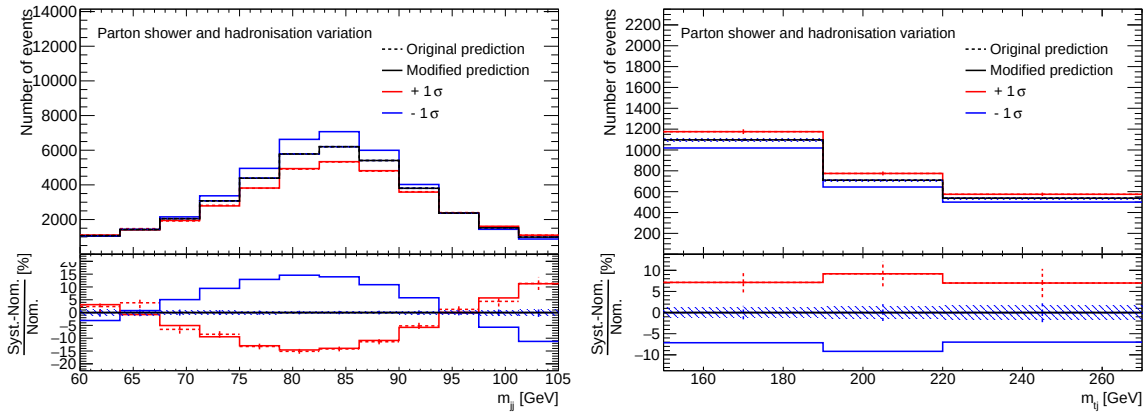
Figure 9.9: The effect of the NP relating to the choice of recoil scheme on (a) m_{jj} and (b) m_{tj} for the (black) nominal prediction and where the associated NP is varied by (red) $+1\sigma$ and (blue) -1σ, where σ represents the pre-fit error for that parameter. Smoothing is applied to the original prediction for m_{jj}. The effect of the NP on both distributions is symmetrised. The original predictions are shown by the dashed lines, whereas the solid lines show the modified variations after smoothing and/or symmetrisation.

The increase in uncertainty contribution from the choice of recoil scheme when moving from the 1D to the 2D Asimov fits can be explained by the expected physics effects. In the alternative `recoil-to-top` recoil scheme, more wide-angle radiation is

emitted from the b-quark in the top-quark decay. As a result, there are more events in which this radiation escapes the top-jet, leading to a lower reconstructed value of $\overline{m_J}$. However, some of this wide-angle radiation is captured by the small-R jets used in the reconstruction of m_{jj}. This is clearly seen in Figure 9.9, where the alternative recoil scheme ($\theta_{\text{recoil}} = +1\sigma$) causes an increase in m_{jj}. Consequently, the same physics effect that lowers $\overline{m_J}$ also increases m_{jj}. This is in contrast to the expected behaviour when varying the JES NPs, where increasing the jet energy would increase both $\overline{m_J}$ and m_{jj}. As a result, the uncertainty from the recoil scheme is not reduced in the same manner as the overall JES effect. Instead, the model becomes more sensitive to variations in the recoil scheme, making the value of m_t more correlated with the recoil NP than before. This increased sensitivity leads to a larger uncertainty contribution from the associated NP.

Other NPs that are significantly constrained include those related to the uncertainty from the parton shower and hadronisation model. Figure 9.10 shows the variations in m_{jj} and m_{tj} due to the associated NPs. As discussed in Section 9.3, the parton shower and hadronisation uncertainty is decorrelated between the observables in the fit to prevent any constraint from one distribution being propagated to another. The reasoning behind this approach can be understood by examining the related distributions in Figure 9.10 and the resulting NP constraints shown in Figure 9.7. A significant constraint arises from the difference in m_{jj} between the nominal prediction using POWHEG+PYTHIA8 and the alternative prediction using POWHEG+HERWIG7. This large difference of approximately 15% allows the profile likelihood fit model to impose a strong constraint on the associated NP, resulting in a post-fit error of 0.29. Figure 9.10 also shows a large variation in m_{tj}, though the associated NP is less constrained, with a post-fit error of 0.72. Since switching between PYTHIA8 and HERWIG7 alters multiple aspects of the simulation, propagating the associated constraint across different observable distributions is not reliable. As a result, this uncertainty was treated as uncorrelated in the fit model. Consequently, three NPs are considered, with the one related to the m_J distribution remaining unconstrained by design.

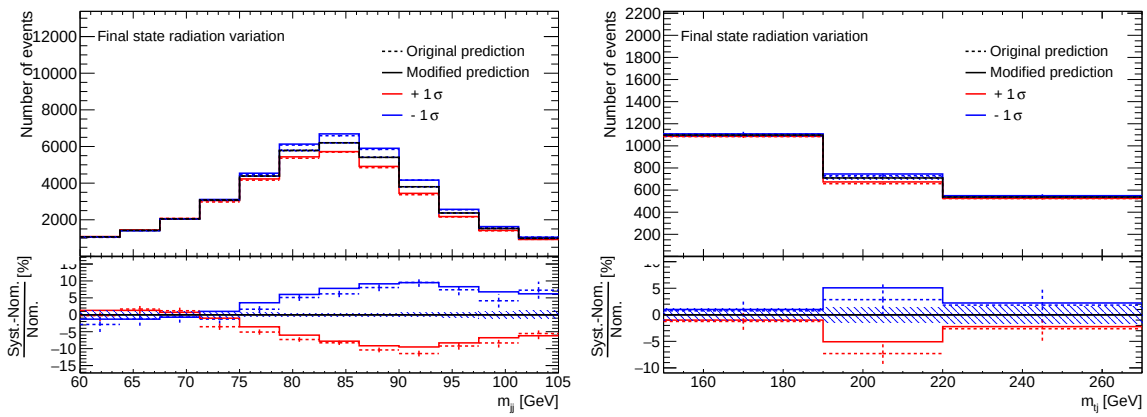
Another NP that is significantly constrained is that related to the FSR uncertainty. Figure 9.11 displays the impact of this NP on m_{jj} and m_{tj}, showing that there is a large variation in the high-m_{jj} region. Differences in this distribution are expected when varying the level of FSR, as it affects the formation of jets and the radiation captured by the large-R jet reclustering to form the top-jet. The large constraint primarily



(a) **(b)**

Figure 9.10: The effect of the decorrelated NPs relating to the parton shower and hadronisation model on (a) m_{jj} and (b) m_{tj} for the (black) nominal prediction and where the associated NP is varied by (red) $+1\sigma$ and (blue) -1σ, where σ represents the pre-fit error for that parameter. Smoothing is applied to the original prediction for m_{jj}. The effect of the NP on both distributions is symmetrised. The original predictions are shown by the dashed lines, whereas the solid lines show the modified variations after smoothing and/or symmetrisation.

arises from the impact on m_{jj}. This is reflected in how the post-fit uncertainty for the associated NP does not significantly change between the 2D and 3D Asimov fits, as seen in Figure 9.7. The variation in m_{tj} due to the modelling of FSR is also anticipated, as m_{tj} is sensitive to radiation from the top and b-quarks, which can be influenced by changes in the modelling of FSR. However, this difference is much smaller than the variation seen in m_{jj}.



(a) **(b)**

Figure 9.11: The effect of the NPs relating to the modelling of FSR on (a) m_{jj} and (b) m_{tj} for the (black) nominal prediction and where the associated NP is varied by (red) $+1\sigma$ and (blue) -1σ, where σ represents the pre-fit error for that parameter. Smoothing is applied to the original prediction for m_{jj}. The effect of the NP on both distributions is symmetrised. The original predictions are shown by the dashed lines, whereas the solid lines show the modified variations after smoothing and/or symmetrisation.

CHAPTER 10

Results

This chapter begins by discussing the fitted values and uncertainties of the NPs, assessed using real ATLAS data, without examining the values of $\overline{m_J}$ and m_t in the data. This approach is made possible by the measurement strategy of using $\overline{m_J}$ in the model, rather than the full distribution of m_J. The model, which is evaluated using a profile likelihood fit to m_{jj} and m_{tj} with data, is then validated by applying the fit result to other kinematic observables independent of m_t, checking the agreement between the data and the predictions from simulation. Subsequently, the measured value of m_t using the full Run 2 ATLAS dataset is presented. A cross-check is performed to confirm that the data are well modelled by the simulation, by splitting the data used to construct the $\overline{m_J}$ observable in different ways: by lepton flavour, by the data-taking period, or by the number of constituent jets inside the top-jet. This m_t result is compared with previous direct measurements including those performed by both ATLAS and CMS at the LHC, as well as with the combined result from the Tevatron. Finally, given the planned future HL-LHC, the projected precision of this m_t result is calculated for four different scenarios, taking into account anticipated improvements in both statistical and systematic uncertainties.

10.1 The blinded fit

The model used for the profile likelihood fit is based on the likelihood functions for $\overline{m_J}$, m_{jj}, and m_{tj}. As discussed in Section 9.3, the fitted values and uncertainties of the NPs, as well as the measured μ, are entirely determined using m_{jj} and m_{tj}. As a result, these quantities can be evaluated before investigating $\overline{m_J}$ or m_t values in the

data. This is achieved by performing the profile likelihood fit using the model defined by Equation 9.8, but with the $L_{\overline{m_J}}$ term removed. This fit is referred to as the *blinded fit*.

After performing the blinded fit, the result is $\mu = 0.87 \pm 0.09$, which is in good agreement with the ATLAS $t\bar{t}$ differential cross-section measurement using boosted top quarks [151]. A goodness-of-fit test is conducted using the saturated model (see Section 9.2.2) to assess how well the model describes the data. This test returns a fit probability of 36%, suggesting that the model provides a good fit to the data. Figure 10.1 shows the pre-fit and post-fit distributions of m_{jj} and m_{tj} for both the data and the prediction.

Figure 10.1 suggests that there may be some shape differences between the data and the prediction for m_{jj} before performing the fit, but these are acceptable within the uncertainties. The post-fit distributions show that the fit adjusts the model to better match the data for both observables, which is consistent with the result of the goodness-of-fit test. This improvement in the agreement is a direct consequence of pulling and/or constraining various NPs in the model. Figure 10.2 displays the plot of fitted values and uncertainties for every NP included in this model setup, with all NPs having pulls less than one standard deviation.

The NP related to the uncertainty associated with the choice of recoil scheme is pulled slightly in the negative direction. Negative values for this NP do not correspond to one of the two recoil models, but rather indicate that there is less wide-angle radiation in data than in the recoil scheme used in the nominal simulation, `recoilToColored`. This is possible as neither recoil scheme is expected to perfectly describe the true nature of the radiation from the b-quark in the top-quark decay. Furthermore, the central value is still consistent with zero. This central value is primarily driven by the m_{tj} observable. Figure 10.3 shows the impact of changing the recoil scheme on the m_{jj} and m_{tj} distributions for the total prediction, compared to the data. For this figure, and for all future figures showing the impact of a NP on either the m_{jj} or m_{tj} distribution, all NPs are set to zero for the nominal prediction. The data clearly favours the recoil scheme used in the nominal $t\bar{t}$ prediction for the m_{tj} distribution, reflecting how the NP is not pulled in the model.

The NPs related to the choice of parton shower and hadronisation model associated with the m_{jj} and m_{tj} distributions are pulled slightly. The impacts of these NPs on the m_{jj} and m_{tj} distributions for the total prediction, compared to the data, are shown in Figure 10.4. These plots clearly show that the agreement between data

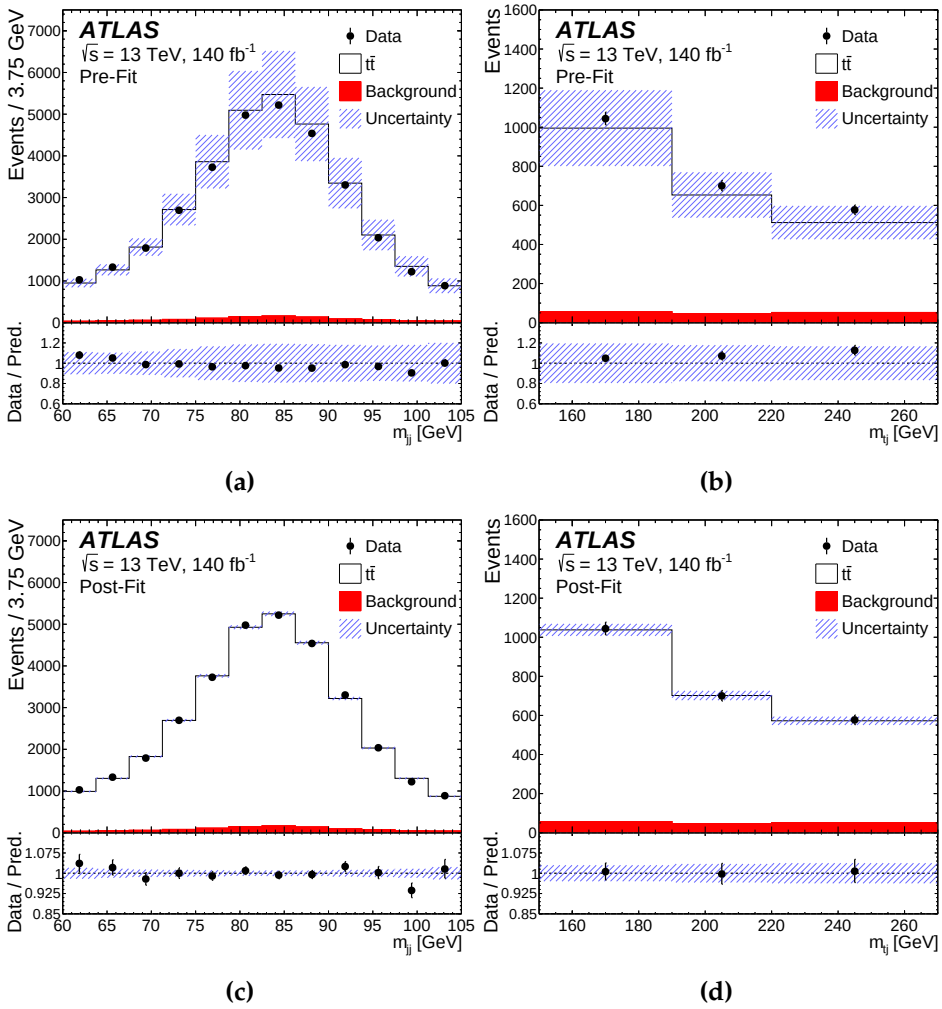


Figure 10.1: The (upper) pre-fit and (lower) post-fit expectations for the (left) m_{jj} and (right) m_{tj} distributions are compared with the data. For the pre-fit case, the $t\bar{t}$ sample is scaled to $\mu = 0.856$, which is the ratio of data to prediction found in the previous ATLAS $t\bar{t}$ differential cross-section measurement using boosted top quarks [151], and $m_t = 172.5$ GeV. There is no uncertainty on the pre-fit value of μ or m_t. The uncertainty band represents the total uncertainty in the model. The bottom panel reflects the agreement between the data and total prediction, where the error bars represent the error on the data, whereas the blue uncertainty bands denote the total uncertainty obtained from the model.

and MC is significantly better for the POWHEG+PYTHIA8 model used in the nominal simulation in m_{jj}, while the data exhibits limited sensitivity to this parameter in m_{tj}. The corresponding pulls and constraints in Figure 10.4 indicate that the data are consistent with the prediction when POWHEG+PYTHIA8 is used to model the $t\bar{t}$ events, as opposed to POWHEG+HERWIG7, which can be excluded at the approximate 2σ level based on the m_{jj} distribution.

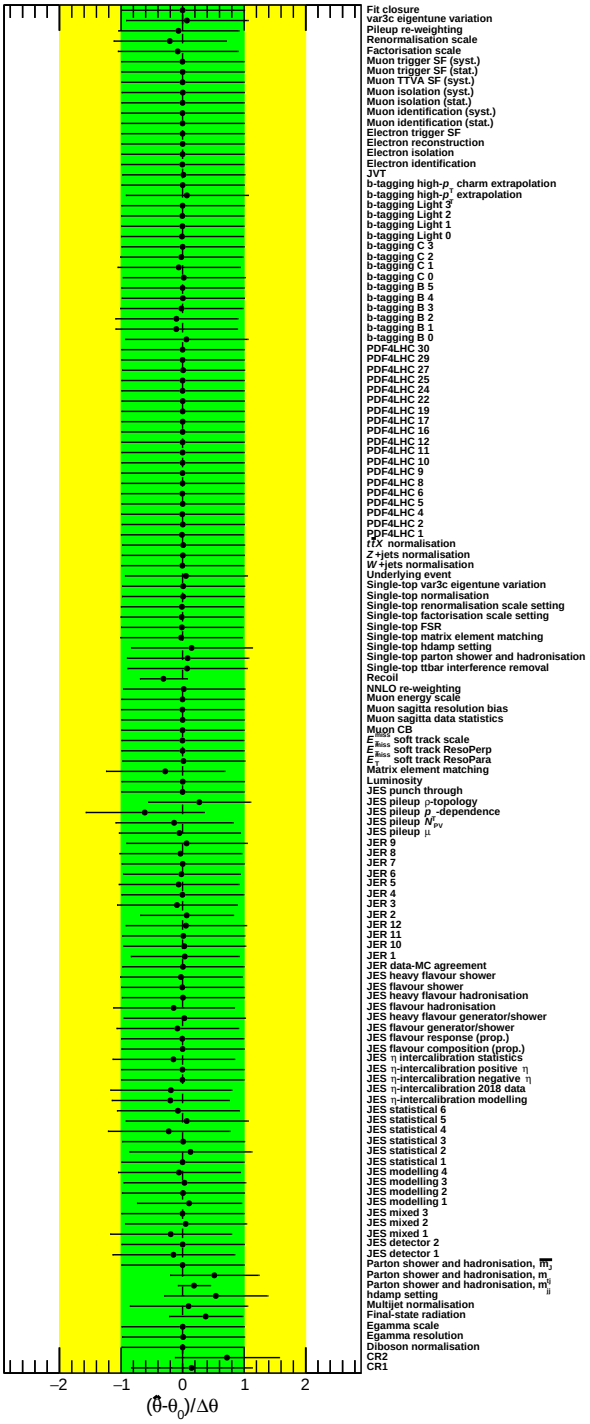
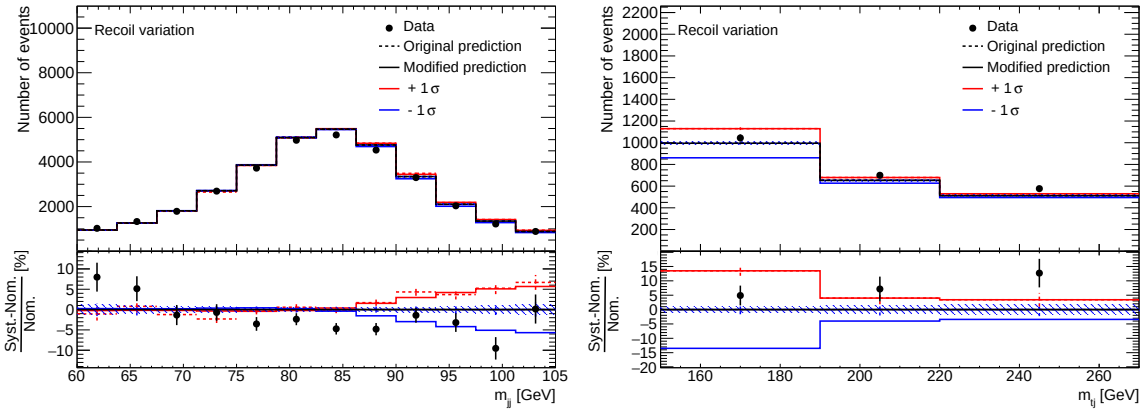
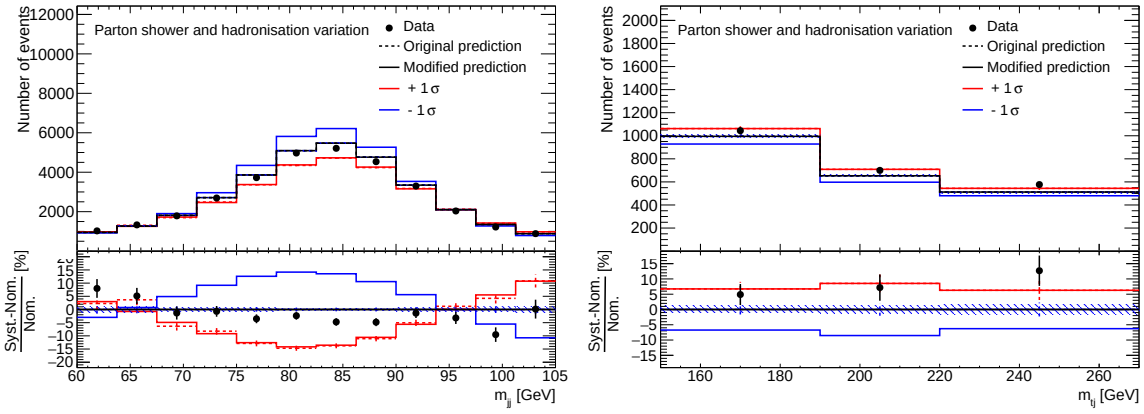


Figure 10.2: Post-fit values and uncertainties for NPs included in the profile likelihood fit to m_{jj} and m_{tj} using data. NPs associated with MC statistical uncertainties are not shown. The best fit and pre-fit values for the NP are given by $\hat{\theta}$ and θ_0, respectively. The pre-fit NP uncertainty is given by $\Delta\theta$.



(a) **(b)**

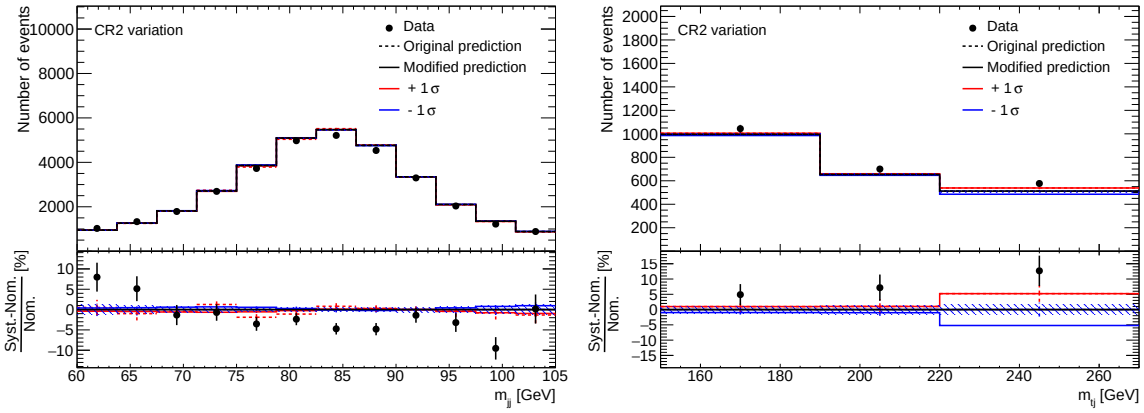
Figure 10.3: The effect of the NP relating to the uncertainty due to the choice of recoil scheme on (a) m_{jj} and (b) m_{tj} for the (black) nominal total prediction and where the associated NP is varied by (red) $+1\sigma$ and (blue) -1σ, where σ represents the pre-fit error for that parameter. Points relating to the data distributions are included. Smoothing is applied to the original prediction for m_{jj}.



(a) **(b)**

Figure 10.4: The effect of the NPs relating to the choice of parton shower and hadronisation model on (a) m_{jj} and (b) m_{tj} for the (black) nominal prediction and where the associated NP is varied by (red) $+1\sigma$ and (blue) -1σ, where σ represents the pre-fit error for that parameter. Points relating to the data distributions are included. Smoothing is applied to the original prediction for m_{jj}.

One NP with a large pull is related to the uncertainty arising from the comparison of the CR2 and CR0 CR models. Figure 10.5 illustrates the impact of this NP on the m_{jj} and m_{tj} distributions, showing the total prediction compared with the data. It is evident from the figure that the pull on this NP is primarily driven by the last bin of the m_{tj} distribution, where the positive variation induced by the NP leads to improved agreement between the data and the prediction.



(a) **(b)**

Figure 10.5: The effect of the NP relating to the comparison of the CR2 and CR0 models on (a) m_{jj} and (b) m_{tj} for the (black) nominal prediction and where the associated NP is varied by (red) $+1\sigma$ and (blue) -1σ, where σ represents the pre-fit error for a given NP. Points relating to the data distributions are included. Smoothing is applied to the original prediction for m_{jj}.

The NPs associated with the modelling of FSR, the matching of MEs to the parton shower, the variation of the h_{damp} parameter, and the p_{T} term used in the JES pileup correction also exhibit mild pulls. The impacts of these NPs on the m_{jj} and m_{tj} distributions, derived from the total prediction with data overlaid, are shown in Figure 10.6 and Figure 10.7. The FSR NP is pulled in the positive direction, reflecting better agreement between data and the up variation in the m_{jj} prediction. The ME matching NP is slightly pulled in the negative direction, though it is difficult to pinpoint the exact cause from the figure, apart from the first bin of m_{tj}, where the data prefers the down variation. The m_{tj} distribution in the data favours the up variation of the NP associated with the setting of h_{damp}, which reflects the positive pull. The NP relating to the p_{T} term used in the JES pileup correction is also pulled slightly in the negative direction, consistent with the observation that the m_{jj} distribution in data appears to favour the down variation in the high-m_{jj} tail.

This cross-check between the signs of the NP pulls in Figure 10.2 and whether these NP variations improve the agreement between data and prediction further increases confidence in the fit model. Importantly, no unexpected NP pulls are observed.

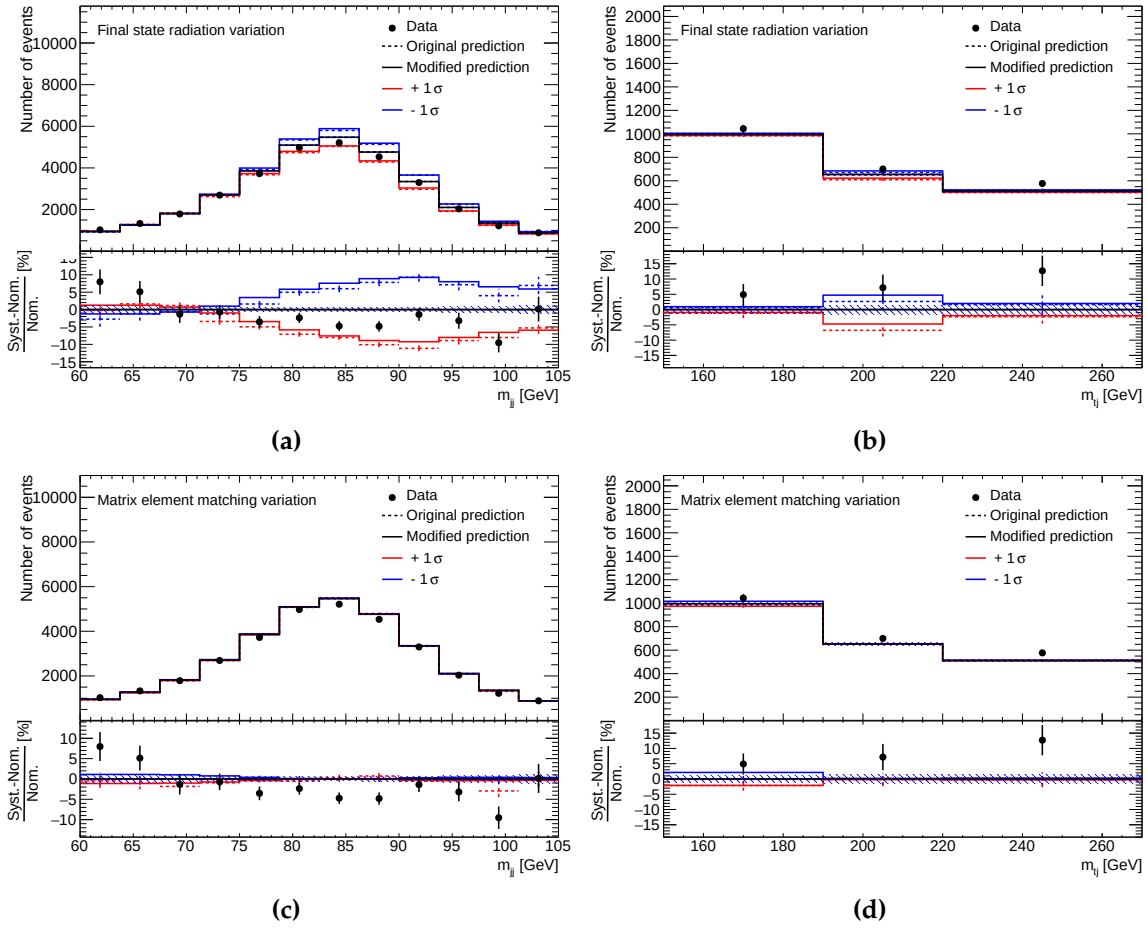


Figure 10.6: The effect of the NPs relating to (upper) the modelling of FSR and (lower) the matching of the matrix elements to the parton shower on (left) m_{jj} and (right) m_{lj} for the (black) nominal prediction and where the associated NP is varied by (red) $+1\sigma$ and (blue) -1σ. σ represents the pre-fit error for a given NP. Points relating to the data distributions are included. Smoothing is applied to the original prediction for m_{jj}.

10.2 Validating the blinded fit

To ensure confidence in the profile likelihood fit's ability to effectively model the signal events, the blinded fit result is applied to other kinematic variables, referred to as *validation observables*, that are not sensitive to m_t. For each of these observables, the level of agreement between the data and the prediction is assessed. Each bin in the resulting distributions is parametrised in terms of the NPs, with shifts in the predictions arising from the pulls on the corresponding NPs. The associated uncertainties on the predictions are evaluated by propagating the covariance matrix

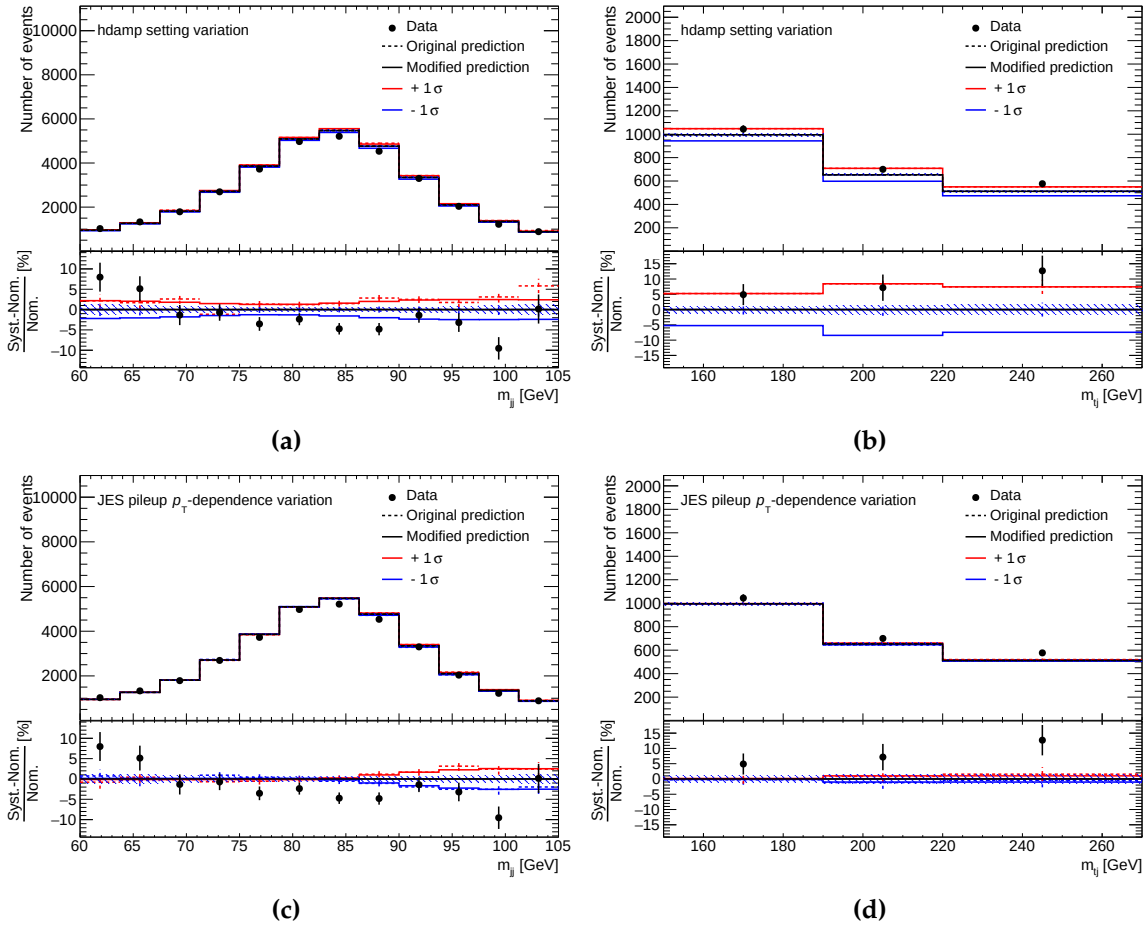


Figure 10.7: The effect of the NPs relating to (upper) the variation of the h_{damp} parameter and (lower) the modelling of the p_{T} term used in the JES calibration on (left) m_{jj} and (right) m_{tj} for the (black) nominal prediction and where the associated NP is varied by (red) $+1\sigma$ and (blue) -1σ. σ represents the pre-fit error for a given NP. Points relating to the data distributions are included. Smoothing is applied to the original prediction for m_{jj}.

obtained from the blinded fit. To be included in any of the validation observable distributions, events must be capable of reconstructing either $\overline{m_J}$ or m_{jj}.

The observables investigated include some related to the top-jet:

- p_{T} of the top-jet.
- Number of constituent jets inside the top-jet.
- p_{T} of the leading constituent jet inside the top-jet.
- p_{T} of the sub-leading constituent jet inside the top-jet.

Additionally, observables related to the other side of the event are also used to validate the blinded fit. These are:

- p_{T} of the lepton.
- ϕ of the lepton.
- η of the lepton.
- p_{T} of the b-tagged jet originating from the leptonically decaying top quark.
- Transverse W-boson mass, m_{T}^{W}.
- $E_{\mathrm{T}}^{\mathrm{miss}}$.

Three further observables are tested:

- ΔR separation between the leptonically decaying top quark and the closest additional jet, $\Delta R(t^{\mathrm{lep}}, j)$, for events with at least one additional jet.
- $\Delta\phi$ separation between the top-jet and the sub-leading additional jet, $\Delta\phi\,(j_2, \text{top-jet})$, for events with at least two additional jets.
- $\Delta\phi$ separation between the leading and sub-leading additional jets, $\Delta\phi\,(j_1, j_2)$, for events with at least two additional jets.

The decision to include $\Delta R(t^{\mathrm{lep}}, j)$ was made because both m_{tj} and $\Delta R(t^{\mathrm{lep}}, j)$ should be sensitive to the same effects of changing the recoil scheme. If the blinded fit result using m_{tj} does not accurately describe $\Delta R(t^{\mathrm{lep}}, j)$, it would indicate potential issues with the model or with the use of m_{tj}. Additionally, $\Delta\phi\,(j_2, \text{top-jet})$ and $\Delta\phi\,(j_1, j_2)$ are considered as differences in the parton shower are expected to manifest in the kinematics of the sub-leading additional jet, while the kinematics of the leading additional jet are driven by the ME calculation. Therefore, these distributions are sensitive to the uncertainty in the choice of parton shower. They are included to assess any mis-modelling due to the large constraint from m_{jj} on the NP associated with the choice of parton shower and hadronisation scheme. The decision to decorrelate the parton shower and hadronisation uncertainty between the fit observables introduces an ambiguity regarding which parton shower and hadronisation NP should be correlated with the uncertainty in the validation observables. The choice is made to correlate the NP associated with the m_{jj} distribution, as it is the most tightly constrained. This selection results in the smallest post-fit uncertainty in the validation regions, providing the most stringent test of the blinded fit result.

Pre-fit and post-fit distributions for the total prediction compared to the data for the observables relating to the top-jet are shown in Figure 10.8 and Figure 10.9. For those observables relating to the other side of the event, their pre-fit and post-fit distributions are shown in Figure 10.10 and Figure 10.11. Similar distributions for the total prediction compared to the data for $\Delta R(t^{\text{lep}}, j)$ are shown in Figure 10.12, where events entering these distributions must have at least one additional jet. Pre-fit and post-fit distributions for the observables sensitive to changes in the parton shower are generated for events that have at least two additional jets, and are presented in Figure 10.13.

From these Figures, it is clear that the agreement between data and prediction improves following the fit, with a reduction in the uncertainties. No post-fit distributions shows any sign of mis-modelling, though there may be a small shape difference between data and prediction observed post-fit for the p_{T} distributions of the top-jet and b-tagged jet on the leptonic side of the event, and the $E_{\text{T}}^{\text{miss}}$.

For each observable, the χ^2 statistic between data and expectation is calculated as described in Section 9.2.1. The goodness-of-fit test using χ^2 is considered a reasonable statistical test, provided that the tested distribution of the kinematic observable has a low correlation with the m_{jj} and m_{tj} distributions used in the profile likelihood fit. The statistical correlation factors for each tested distribution with the m_{jj} and m_{tj} distributions, based on the nominal $t\bar{t}$ prediction, are provided in Table 10.1.

From Table 10.1, the high statistical correlation between the distribution for the number of top-jet constituent jets and the m_{jj} distribution is likely due to the small number of bins used in the distribution for the number of jet constituents. Additionally, the first bin contains no events from the m_{jj} distribution because of the m_{jj} selection requirements. Some level of correlation between the distribution for the p_{T} of the top-jet and the m_{jj} distribution is expected, as a higher top-jet p_{T} value generally indicates that the top-jet is more likely to have fewer jet constituents. A large correlation is observed between m_{tj} and $\Delta R(t^{\text{lep}}, j)$, which is expected since both kinematic variables are sensitive to the same effects. The remaining distributions show little correlation with m_{jj} and m_{tj}, suggesting that the χ^2 goodness-of-fit test is reliable for these distributions. The χ^2 value and associated fit probability for each distribution that has a statistical correlation with both m_{jj} and m_{tj} below 10% are provided in Table 10.2. Table 10.2 suggests that the agreement is still acceptable for the p_{T} distributions of the top-jet and b-tagged jet on the leptonic side of the event, and the $E_{\text{T}}^{\text{miss}}$, even though they exhibit a small shape difference.

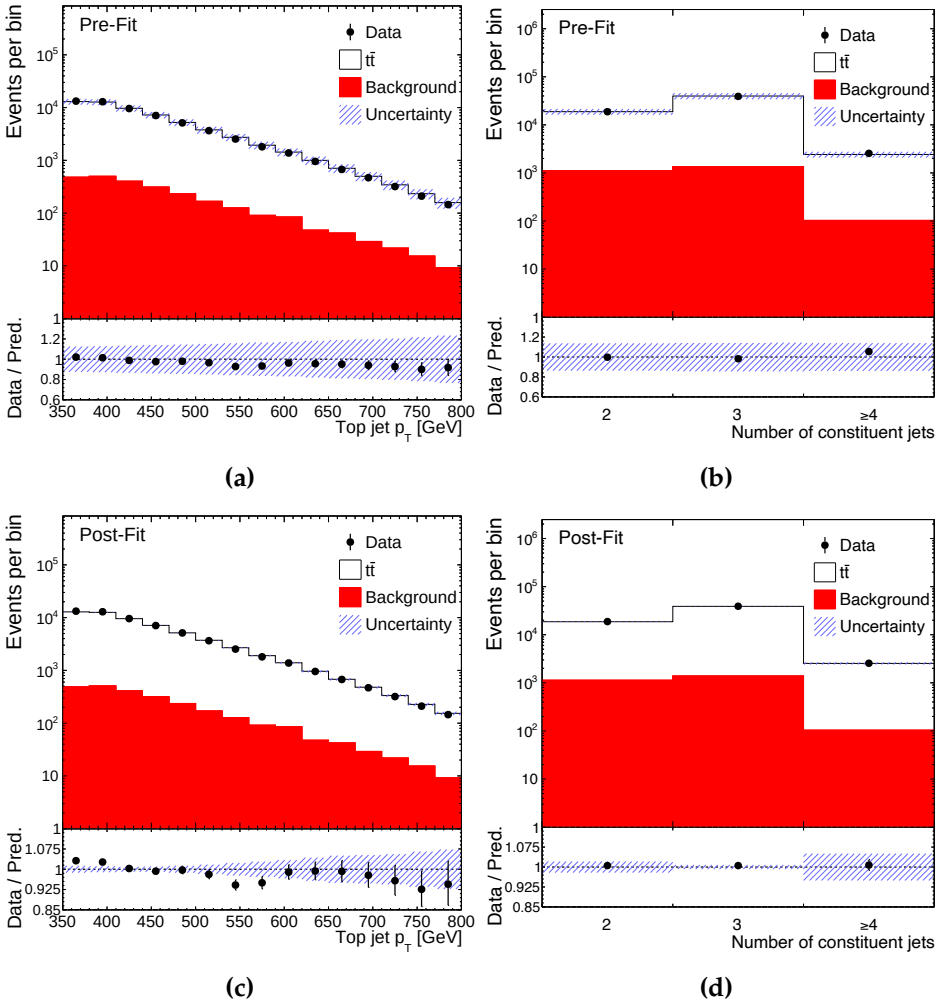


Figure 10.8: The (upper) pre-fit and (lower) post-fit expectations for (left) the top-jet p_{T} and (right) the number of constituent jets inside the top-jet are compared with the data. For the pre-fit case, the $t\bar{t}$ sample is scaled to $\mu = 0.856$, which is the ratio of data to prediction found in the previous ATLAS $t\bar{t}$ differential cross-section measurement using boosted top quarks [151], and $m_t = 172.5$ GeV. There is no uncertainty on the pre-fit value of μ or m_t. The uncertainty band represents the total uncertainty in the model. The bottom panel reflects the agreement between the data and total prediction, where the error bars represent the error on the data, whereas the blue uncertainty bands denote the total uncertainty obtained from the fit model.

This apparent agreement, despite the minor shape differences between data and prediction, can be attributed to the fact that certain NPs exhibit similar shapes in the relevant validation observables. For example, the impact of the NP associated with the NNLO re-weighting of the nominal simulation leads has a variation in the top-jet p_{T} distribution that resembles the observed shape difference in the data-prediction

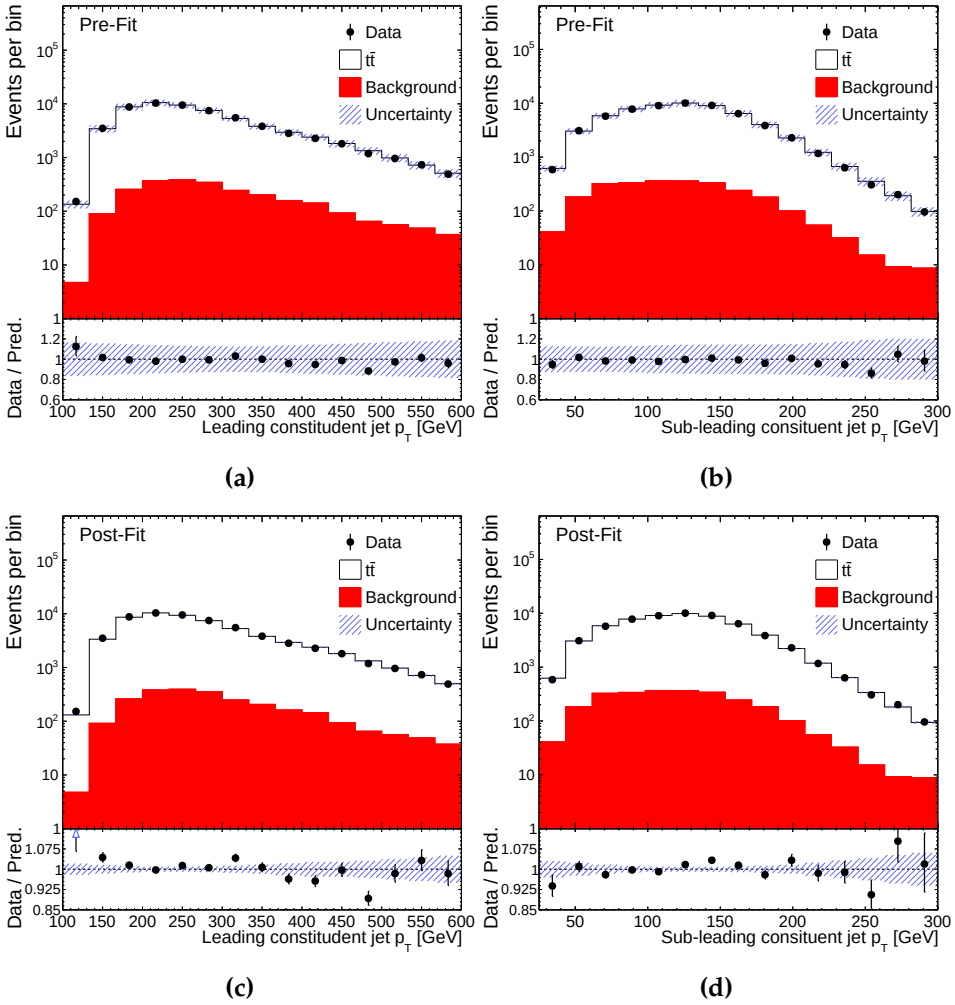


Figure 10.9: The (upper) pre-fit and (lower) post-fit p_{T} expectations for (left) the leading and (right) the sub-leading constituent jets are compared with the data. For the pre-fit case, the $t\bar{t}$ sample is scaled to $\mu = 0.856$, which is the ratio of data to prediction found in the previous ATLAS $t\bar{t}$ differential cross-section measurement using boosted top quarks [151], and $m_t = 172.5$ GeV. There is no uncertainty on the pre-fit value of μ or m_t. The uncertainty band represents the total uncertainty in the model. The bottom panel reflects the agreement between the data and total prediction, where the error bars represent the error on the data, whereas the blue uncertainty bands denote the total uncertainty obtained from the fit model.

agreement in this observable, as shown in Figure 10.14. Additionally, Figure 10.14 demonstrates that the NP related to the h_{damp} parameter introduces slopes in the p_{T} distribution of the b-tagged jet on the leptonic side, as well as in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. The shapes of these variations likewise reflect the shape of the discrepancies observed between data and prediction in these observables. Therefore, since some NP variations

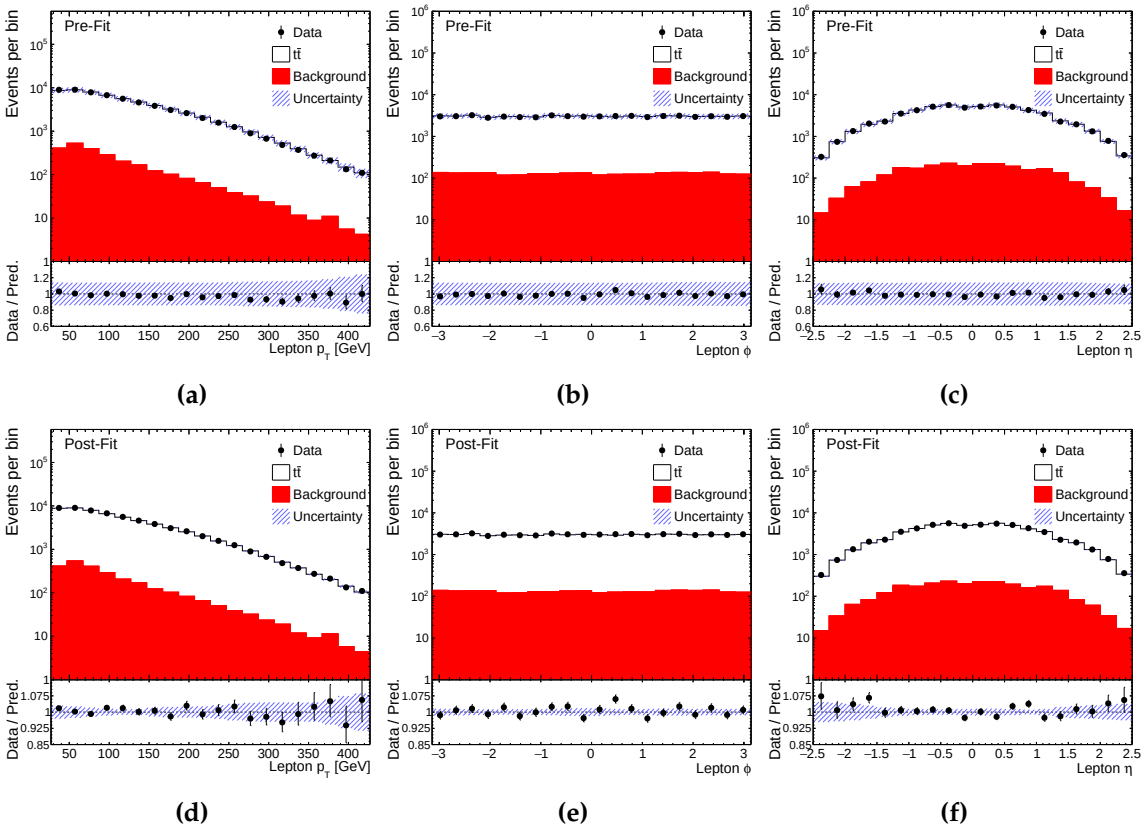


Figure 10.10: The (upper) pre-fit and (lower) post-fit expectations for (left) the selected lepton's p_{T}, (centre) ϕ, and (right) η are compared with the data. For the pre-fit case, the $t\bar{t}$ sample is scaled to $\mu = 0.856$, which is the ratio of data to prediction found in the previous ATLAS $t\bar{t}$ differential cross-section measurement using boosted top quarks [151], and $m_t = 172.5$ GeV. There is no uncertainty on the pre-fit value of μ or m_t. The uncertainty band represents the total uncertainty in the model. The bottom panel reflects the agreement between the data and total prediction, where the error bars represent the error on the data, whereas the blue uncertainty bands denote the total uncertainty obtained from the fit model.

exhibit shape differences similar to those observed between data and prediction in the validation observables, these discrepancies are not considered to be significant.

By inspecting Figures 10.8-10.13 and Table 10.2, it is evident that the fit result successfully models the signal events, with acceptable χ^2 values and associated fit probabilities for the agreement between the prediction and data seen in each distribution. Based on these observations, the fit model was concluded to be valid and reliable.

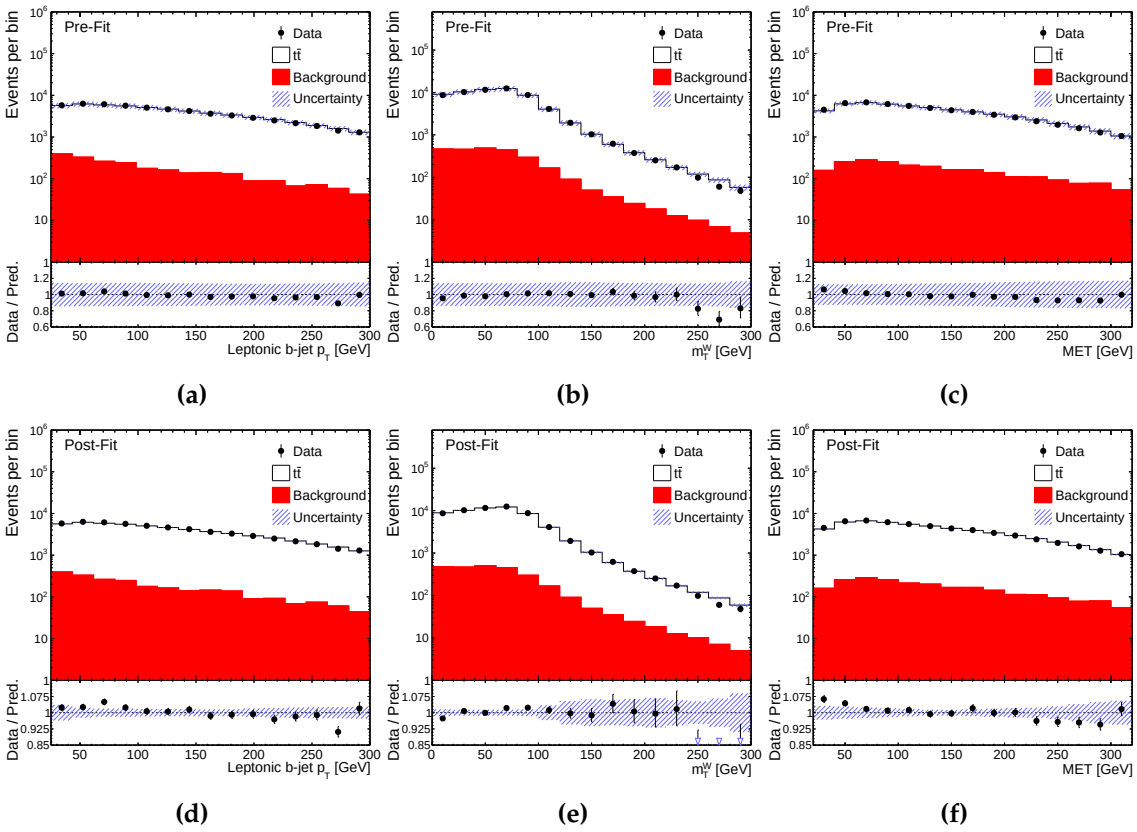
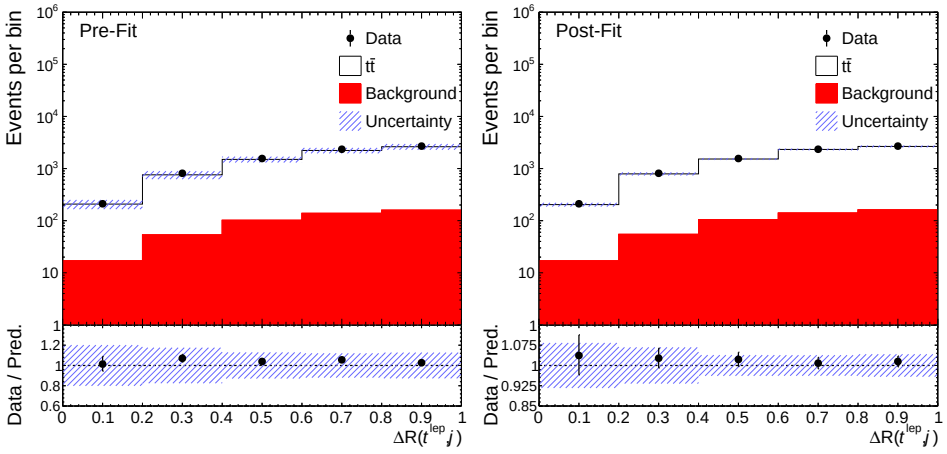


Figure 10.11: The (upper) pre-fit and (lower) post-fit expectations for (left) the p_{T} of the leptonic b-jet, as well as (centre) the m_{T}^{W} and (right) $E_{\mathrm{T}}^{\mathrm{miss}}$ are compared with the data. For the pre-fit case, the $t\bar{t}$ sample is scaled to $\mu = 0.856$, which is the ratio of data to prediction found in the previous ATLAS $t\bar{t}$ differential cross-section measurement using boosted top quarks [151], and $m_t = 172.5$ GeV. There is no uncertainty on the pre-fit value of μ or m_t. The uncertainty band represents the total uncertainty in the model. The bottom panel reflects the agreement between the data and total prediction, where the error bars represent the error on the data, whereas the blue uncertainty bands denote the total uncertainty obtained from the fit model.



(a) **(b)**

Figure 10.12: The (a) pre-fit and (b) post-fit expectations for $\Delta R(t^{\text{lep}}, j)$ are compared with the data. For the pre-fit case, the $t\bar{t}$ sample is scaled to $\mu = 0.856$, which is the ratio of data to prediction found in the previous ATLAS $t\bar{t}$ differential cross-section measurement using boosted top quarks [151], and $m_t = 172.5$ GeV. There is no uncertainty on the pre-fit value of μ or m_t. The uncertainty band represents the total uncertainty in the model. The bottom panel reflects the agreement between the data and total prediction, where the error bars represent the error on the data, whereas the blue uncertainty bands denote the total uncertainty obtained from the fit model.

Table 10.1: Statistical correlation factors for the validation kinematic distributions with the m_{jj} and m_{tj} distributions using the nominal $t\bar{t}$ MC sample.

Kinematic distribution	Correlation with m_{jj}	Correlation with m_{tj}
Top-jet p_{T}	-0.23	+0.02
Leading constituent jet p_{T}	-0.41	+0.03
Sub-leading constituent jet p_{T}	+0.05	-0.01
Number of constituent jets	+0.59	-0.04
Lepton p_{T}	-0.04	-0.05
Lepton ϕ	-0.01	+0.01
Lepton η	+0.01	+0.01
Leptonic b-jet p_{T}	-0.01	-0.05
m_{T}^{W}	-0.02	+0.01
$E_{\text{T}}^{\text{miss}}$	-0.01	-0.01
$\Delta\phi(\text{j}_1, \text{j}_2)$	+0.04	-0.06
$\Delta\phi(\text{j}_2, \text{top-jet})$	+0.04	-0.03
$\Delta R(t^{\text{lep}}, j)$	+0.06	-0.60

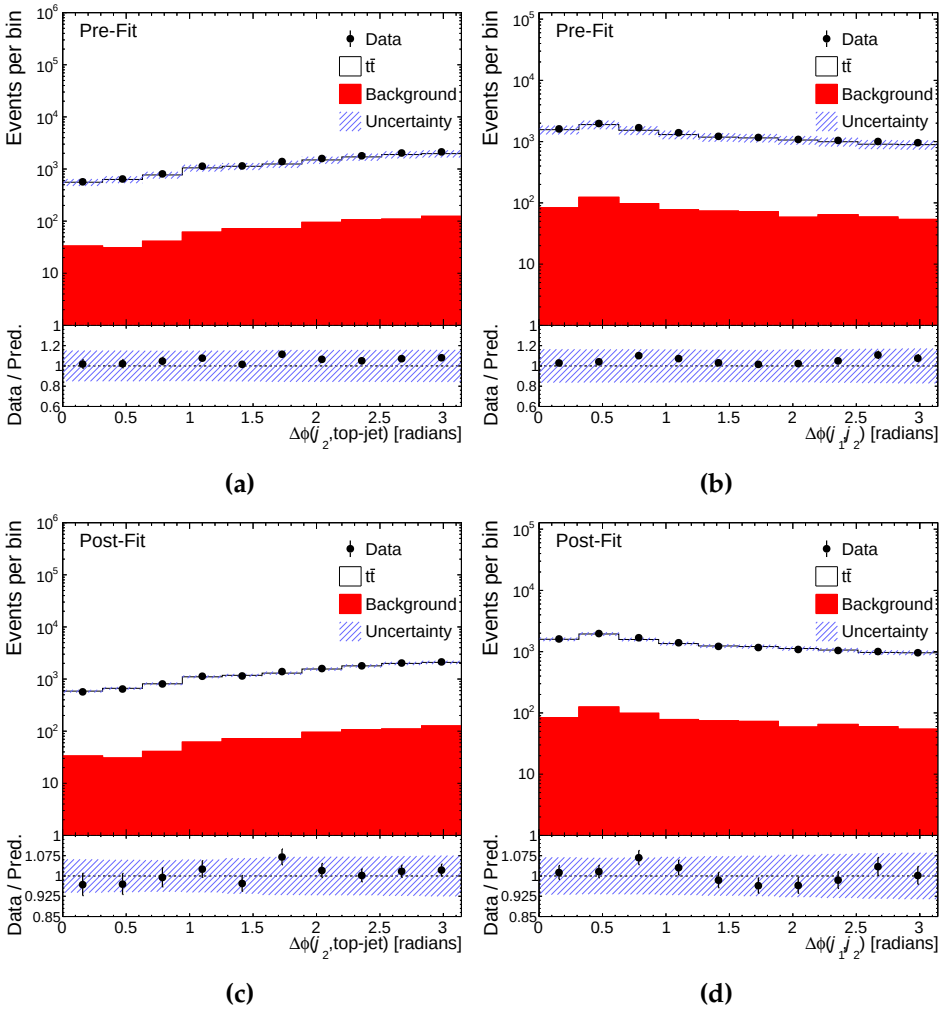


Figure 10.13: The (upper) pre-fit and (lower) post-fit expectations for (left) $\Delta\phi(\mathrm{j}_2, \text{top-jet})$ and (right) $\Delta\phi(\mathrm{j}_1, \mathrm{j}_2)$ are compared with the data. For the pre-fit case, the $t\bar{t}$ sample is scaled to $\mu = 0.856$, which is the ratio of data to prediction found in the previous ATLAS $t\bar{t}$ differential cross-section measurement using boosted top quarks [151], and $m_t = 172.5$ GeV. There is no uncertainty on the pre-fit value of μ or m_t. The uncertainty band represents the total uncertainty in the model. The bottom panel reflects the agreement between the data and total prediction, where the error bars represent the error on the data, whereas the blue uncertainty bands denote the total uncertainty obtained from the fit model.

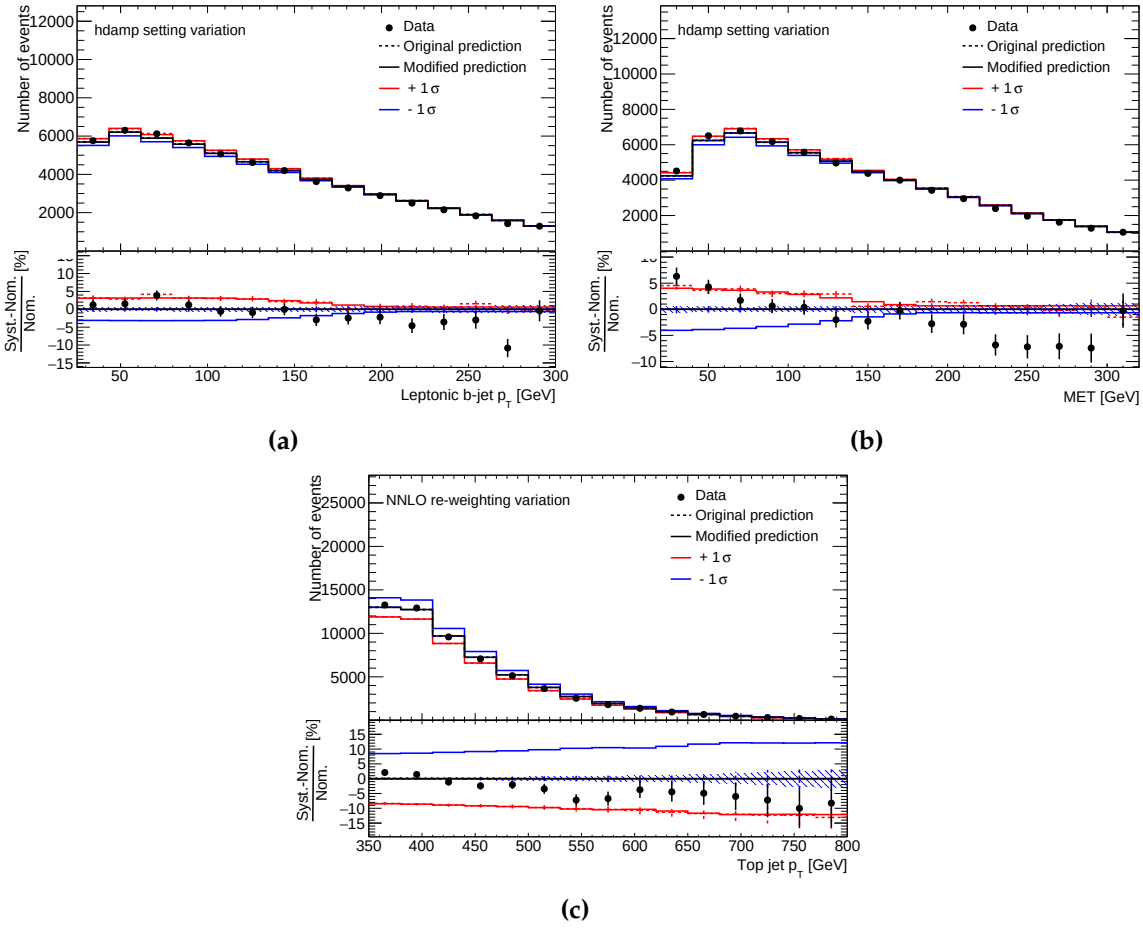


Figure 10.14: The effect of the NP relating to the variation of the h_{damp} parameter on (a) the p_{T} distribution for the b-tagged jet associated with the leptonically decaying top quark and (b) the distribution of $E_{\text{T}}^{\text{miss}}$. Additionally, (c) the effect of varying the NP related to the uncertainty in the NNLO re-weighting procedure on the p_{T} distribution for the top-jet is included. For each plot the (black) nominal prediction is varied by (red) $+1\sigma$ and (blue) -1σ. σ represents the pre-fit error for a given NP. Points relating to the data distributions are included.

Table 10.2: The agreement between data and prediction for the tested observables that have less than 10% statistical correlation with the m_{jj} and m_{tj} when using the constrained model.

Distribution	χ^2	N_{dof}	Prob (%).
Sub-leading constituent jet p_{T}	19.5	15	19
Lepton p_{T}	11.3	20	94
Lepton ϕ	23.6	20	26
Lepton η	22.9	20	29
Leptonic b-jet p_{T}	14.9	15	46
m_{T}^{W}	17.1	15	31
$E_{\text{T}}^{\text{miss}}$	13.0	15	60
$\Delta\phi(\text{j}_1, \text{j}_2)$	9.1	10	52
$\Delta\phi(\text{j}_2, \text{top-jet})$	6.1	10	81

10.3 Top-quark mass in ATLAS data

The $\overline{m_J}$ in data is measured to be 172.46 GeV. The profile likelihood fit to the three observables gives $m_t = (172.95 \pm 0.53)$ GeV, where the total uncertainty is larger than the expected uncertainty of 0.51 GeV. This discrepancy is likely due to the observed number of events being lower than expected. The contributions to the total systematic uncertainty from experimental and theoretical sources are comparable, at 0.36 and 0.28 GeV, respectively. The largest contribution to the total uncertainty comes from the JES, as shown in the breakdown of uncertainty contributions in Table 10.3. The uncertainties associated to the heavy flavour JES components contribute 0.14 GeV. This has been reduced by approximately 50% as a direct consequence of using the new precision flavour JES uncertainties discussed in Section 5.5.3. The total uncertainty also includes a sizeable statistical component of 0.27 GeV, suggesting that significant improvements to the precision are possible given larger datasets (see Section 10.4).

The fitted values of the NPs that contribute most significantly to the total uncertainty on m_t are shown in Figure 10.15. This figure illustrates that several uncertainty sources related to the modelling of the $t\bar{t}$ events have substantial contributions to the total precision on m_t. These include uncertainties arising from the modelling of additional radiation, CR, and the choice of parton shower and hadronisation model.

Figure 10.16 visualises the pre-fit and post-fit agreement between the prediction and data for the m_J distribution, where the uncertainty band shown includes the uncertainties on the fitted μ and m_t values. There is good agreement observed between the data and prediction after performing the fit.

Figure 10.17 shows the result of a profile likelihood scan, illustrating how $-2\ln\lambda$ varies as m_t is scanned over the range $171 < m_t < 174$ GeV. For each value of m_t, the conditional maximum likelihood estimates of the model parameters are determined. As expected, $-2\ln\lambda$ follows a χ^2 distribution, with the minimum indicating the best-fit m_t value. The $\pm 1\sigma$ confidence level intervals correspond to the points where the distribution rises by one unit from the minimum. Although the $\pm 1\sigma$ confidence level intervals extracted from Figure 10.17 appear slightly asymmetric compared to those quoted in Table 10.3, this is not unexpected. The covariance matrix method described in Section 9.4 yields a symmetric uncertainty, whereas the scan-based approach can accommodate asymmetries in the likelihood. Symmetrising the confidence level

Table 10.3: Categorised breakdown of the uncertainty sources contributing to the total uncertainty in m_t. Each category's contribution is equal to the sum in quadrature of the individual uncertainty contributions. 'Radiation' refers to the sources relating to the modelling of ISR and FSR for $t\bar{t}$ events. Those related to ISR are the setting of h_{damp} and the use of the Var3c eigentune of the PYTHIA8 A14 tune. 'Higher-order corrections' refers to sources relating to the μ_f and μ_r scales and the NNLO reweighting. Systematic uncertainties associated with the modelling of background processes are contained within the 'Background modelling' group.

Source	Uncertainty (GeV)
JES	$\pm$ 0.29
Radiation (ISR and FSR)	$\pm$ 0.17
Colour reconnection (CR1 and CR2)	$\pm$ 0.15
JES heavy flavour	$\pm$ 0.14
Parton shower and hadronisation model	$\pm$ 0.14
JER	$\pm$ 0.10
MC statistics	$\pm$ 0.08
Underlying event	$\pm$ 0.08
Recoil	$\pm$ 0.07
Fit closure	$\pm$ 0.07
Background modelling	$\pm$ 0.05
Matrix element matching ($p_{\text{T}}^{\text{hard}} = 1$)	$\pm$ 0.04
b-tagging	$\pm$ 0.04
Higher-order corrections	$\pm$ 0.02
$E_{\text{T}}^{\text{miss}}$	$\pm$ 0.02
Pileup	$\pm$ 0.01
JVT	$\pm$ 0.01
PDF	$\pm$ 0.01
Leptons	$\pm$ 0.01
Luminosity	$<$ 0.01
Total statistical	$\pm$ 0.27
Total systematic	$\pm$ 0.46
Total	$\pm$ 0.53

intervals obtained from the likelihood scan results in the same total uncertainty as that reported in Table 10.3.

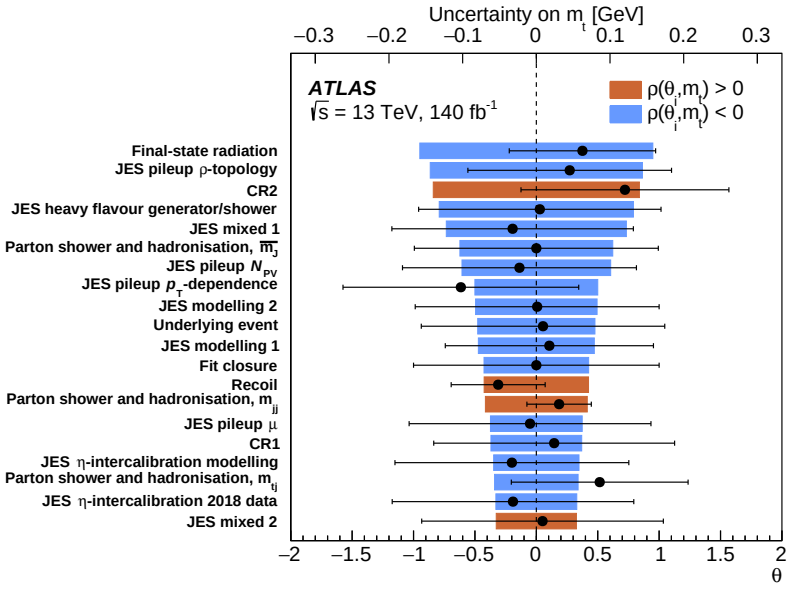
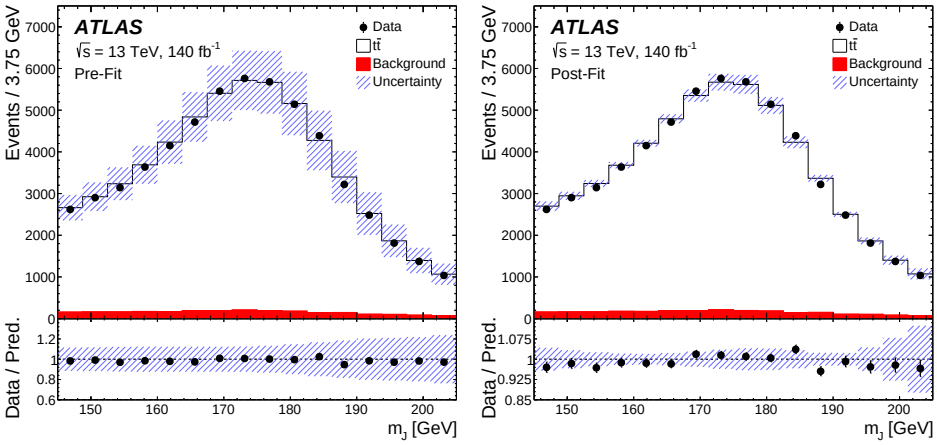


Figure 10.15: The fitted values for NPs that have a contribution to the uncertainty on m_t greater than 0.05 GeV are shown in the black points and are represented in units of σ along the lower x-axis. The error bars represent the post-fit uncertainty in the value of the NP. The corresponding contribution of each NP to the total uncertainty of m_t is shown by the coloured bars and can be interpreted using the upper x-axis. The red (blue) coloured bars represent the NPs that have a positive (negative) correlation with m_t. All modelling uncertainties displayed in the figure correspond to uncertainties on the signal $t\bar{t}$ process.

10.3.1 Cross-checking the result

It is crucial to cross-check the m_t result. Three tests are performed, each splitting the data used to construct the $\overline{m_J}$ observable in different ways: by lepton flavour, by the data-taking period (2015-2017, 2018), or by the number of constituent jets inside the top-jet. In each case, a fit to the data is performed with two m_t parameters. Figure 10.18 shows that for each test the difference between the fitted m_t values is found to be consistent with zero within one standard deviation, accounting for the correlation between the fitted parameters. The correlations between the m_t parameters for the tests split by lepton flavour, data-taking periods, and number of constituent jets inside the top-jet are 0.73, 0.64, and 0.30, respectively.



(a) **(b)**

Figure 10.16: The pre and post-fit expectations for the m_J distribution are compared with the data. For the pre-fit case, the $t\bar{t}$ sample is scaled to $\mu = 0.856$, which is the ratio of data to prediction found in the previous ATLAS $t\bar{t}$ differential cross-section measurement using boosted top quarks [151], and $m_t = 172.5$ GeV. There is no uncertainty on the pre-fit value of μ or m_t. The uncertainty band represents the total uncertainty in the model, including the uncertainty on m_t and μ. The bottom panel reflects the agreement between the data and total prediction, where the error bars represent the error on the data, whereas the blue uncertainty bands denote the total uncertainty obtained from the fit model.

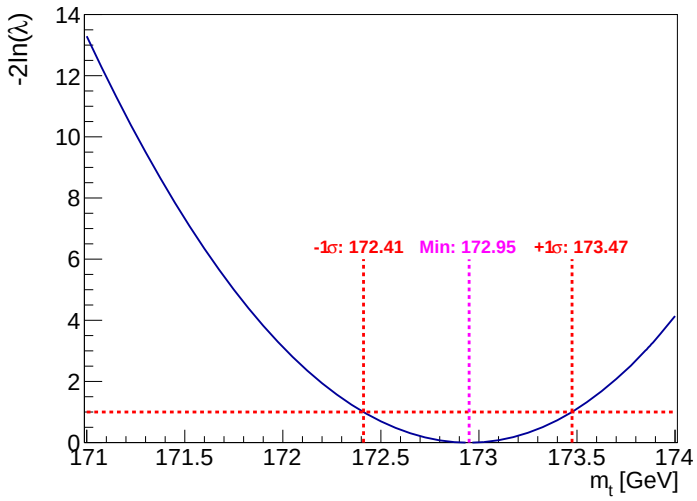


Figure 10.17: Distribution of $-2\ln\lambda$, where λ represents the profile likelihood ratio. This distribution follows a χ^2 distribution according to Wilks' theorem [240], where the minimum corresponds to the best fit value of m_t, and the $\pm 1\sigma$ confidence levels are shown where the distribution intersects with the line $-2\ln\lambda = 1$.

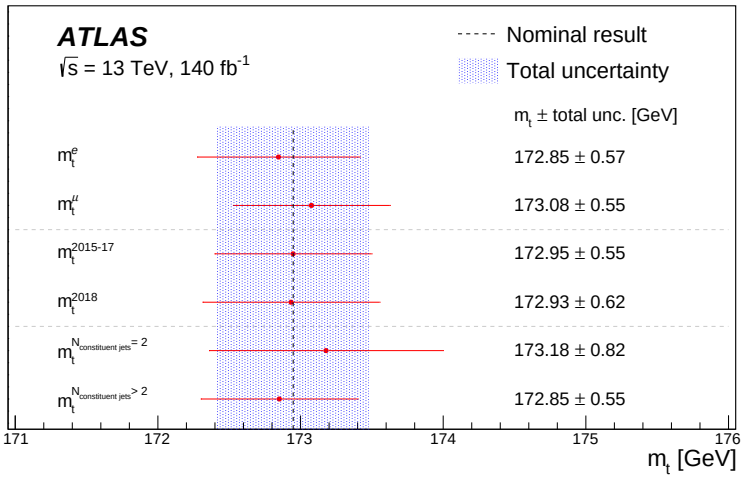


Figure 10.18: Measurements of m_t when using events split by lepton flavour, different data-taking periods, and by the number of constituent jets inside the top-jet. In each of the fits, there is one free parameter for m_t for each section of the dataset. The dashed line and shaded band represent the central value and uncertainty of the measurement with no split of the events and one m_t parameter.

10.3.2 Comparison with previous measurements

A comparison of this result with previous ATLAS m_t measurements is shown in Figure 10.19. It is clear from this comparison that the current result represents the most precise m_t measurement in a single channel from ATLAS to date. Moreover, the result is in very good agreement with the ATLAS Run 1 combination, which has an uncertainty that is only lower by 0.05 GeV [32].

Figure 10.20 compares the result presented in this thesis with other precise determinations of m_t. This measurement shows excellent agreement with the LHC Run 1 combination [32], in contrast to the most precise determination from a single channel, which shows signs of tension with the combination. Furthermore, the single measurement presented here is more precise than the Tevatron combination, which has a precision of 0.65 GeV [31]. It also surpasses the latest boosted m_t measurement from CMS, which has a very similar central value but a total uncertainty larger by 0.31 GeV [34].

To statistically test the consistency between these different measurements and the result presented in this thesis, a dedicated study would need to be performed to estimate the correlation between the measurements. This is not so simple given

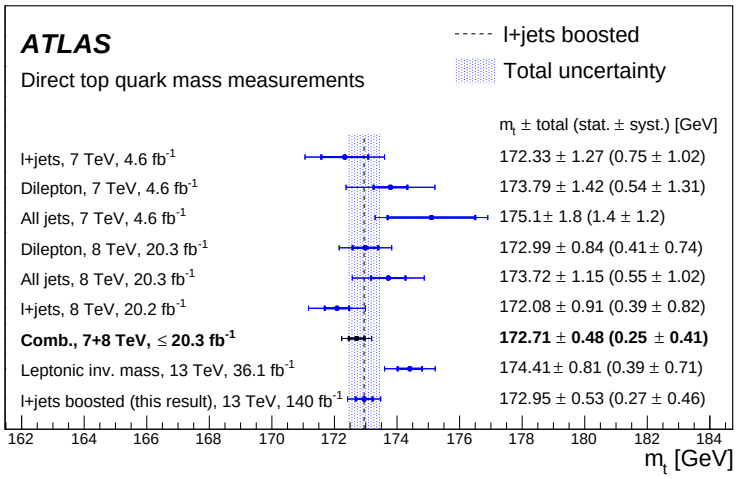


Figure 10.19: The m_t result is compared with previous ATLAS m_t measurements. The dashed line and shaded band represent the central value and uncertainty of this measurement (which is also shown as the final measurement point in the plot).

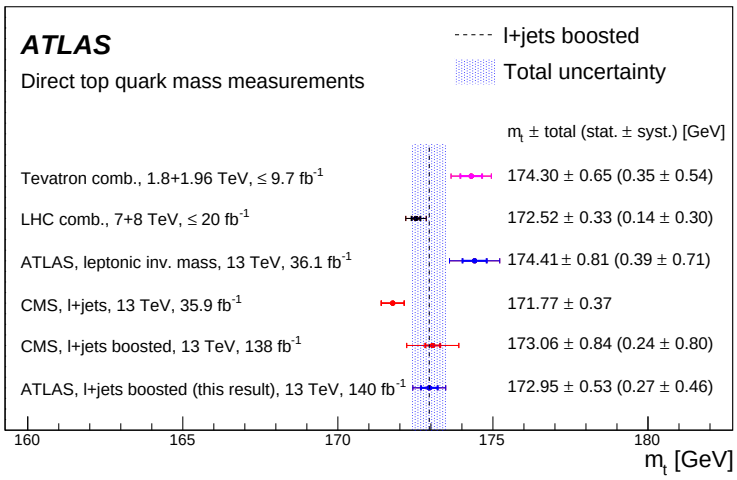


Figure 10.20: The m_t result is compared with other precise determinations of m_t. The dashed line and shaded band represent the central value and uncertainty of this measurement (which is also shown as the final measurement point in the plot).

different measurements will depend on different models and systematic uncertainty definitions.

10.4 Projection of the expected precision at the HL-LHC

The HL-LHC is expected to provide a wealth of data for particle physics analyses, with data-taking scheduled to begin in 2030. Although the precision of the m_t result

presented in Section 10.3 is currently dominated by sources of systematic uncertainty, it includes a 0.27 GeV component arising from the available number of events in the data. As a result, the precision of this measurement will improve with a dataset of higher integrated luminosity. To predict this potential improvement, four hypothetical scenarios are considered, each representing possible advances in theoretical predictions and experimental techniques that would reduce various uncertainty contributions [2]:

1. Scenario 1: The only improvement comes from the increase in data and simulation statistics with luminosity. This is considered a pessimistic view of the future.

2. Scenario 2: Starting from Scenario 1, the JES uncertainty component related to the pileup correction based on ρ is reduced by a factor of seven, following the results of dedicated studies presented in Ref. [248]. In addition, the JES component associated with the non-closure uncertainty introduced for 2018 data, which accounts for differences in the hadronic tile calorimeter calibration during that year of data-taking [133], is removed. All remaining JES uncertainty components, except those related to the flavour dependence of the initiating jets, are reduced by 50%, based on studies of new single-particle measurements that demonstrate improvements in the JES uncertainty, as described in Ref. [249].

3. Scenario 3: Starting from Scenario 2, all $t\bar{t}$ signal modelling uncertainties are reduced by 50%. This reduction is considered reasonable given ongoing theoretical improvements and the future tuning of theoretical predictions to data.

4. Scenario 4: Starting from Scenario 3, JES flavour dependence uncertainties are reduced by 50%. While this is possible, the current flavour uncertainties are already very precise.

When assessing the expected uncertainty on the extracted m_t for each scenario, the fit setup is kept identical to that used to measure m_t in this thesis. However, the expected number of events is scaled according to the increased luminosity, and the sizes of the systematic uncertainties are adjusted as specified for each scenario. This approach neglects any changes in the kinematics of the signal events, which are expected to be small, and the increase in the $t\bar{t}$ cross-section. Additionally, with more events, the binning strategy and choice of observables could be revisited, potentially leading to further improvements. Therefore, the results presented should be viewed as rough but conservative estimates.

Figure 10.21 shows how the total uncertainty in the m_t result from Section 10.3, Δm_t, would evolve as the size of the available datasets increases with the HL-LHC operation, for each of the four scenarios. Table 10.4 summarises these predictions. The best-case estimate suggests the precision of the m_t result could reach 0.23 GeV, making it more precise than the LHC Run 1 combination, which remains the most precise measurement of m_t to date. Even in the pessimistic Scenario 1, the total uncertainty is expected to decrease to 0.39 GeV. This study shows that the channel studied in this thesis should continue to be pursued at the HL-LHC with the prospect of a highly precise m_t determination. This study also underscores the growing importance of resolving the theoretical interpretation of m_t measured by direct methods.

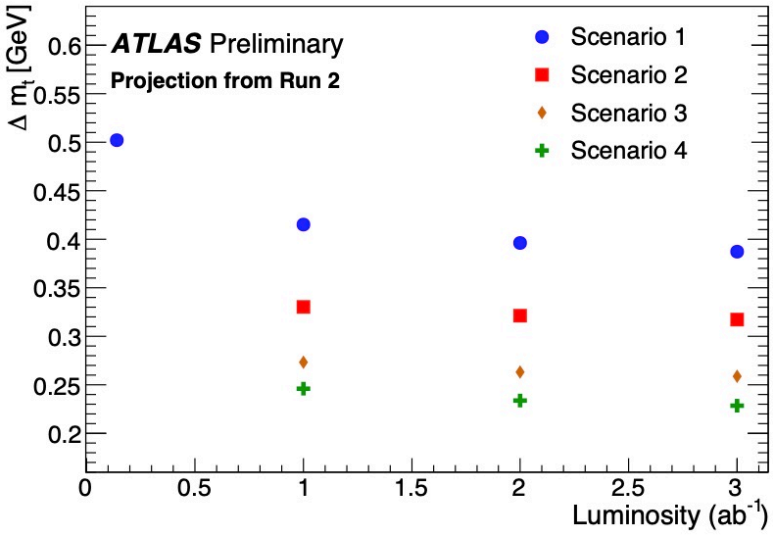


Figure 10.21: Predicted total uncertainty on the m_t result presented in Section 10.3 for the four hypothetical scenarios.

Table 10.4: Expected total uncertainty in m_t, Δm_t, for the four hypothetical scenarios for the improvements to the systematic uncertainties given different datasets, measured in integrated luminosity.

Luminosity	Δm_t [GeV]			
	Scenario 1	Scenario 2	Scenario 3	Scenario 4
1 ab^{-1}	0.42	0.33	0.27	0.25
2 ab^{-1}	0.40	0.32	0.26	0.23
3 ab^{-1}	0.39	0.32	0.26	0.23

CHAPTER 11

Conclusion

As the heaviest fundamental particle in the SM, the top quark and its mass play a pivotal role in testing the internal consistency of the SM. Precise measurements of m_t serve as key inputs to global EW fits, making it a cornerstone of particle physics research. This thesis presented a direct measurement of m_t using the full Run 2 dataset collected by the ATLAS detector during pp collisions at $\sqrt{s} = 13$ TeV delivered by the LHC, corresponding to an integrated luminosity of 140 fb^{-1}. The analysis focused on semi-leptonic $t\bar{t}$ events where the hadronically decaying top quark has high-p_{T}. In such boosted cases, the top quark is reconstructed as a single RC large-R jet, with the invariant mass of the jet providing direct sensitivity to m_t. A profile likelihood fit was used, incorporating two additional observables to mitigate the impact of systematic uncertainties. By including these observables, the profile likelihood fit was able to constrain the related specific NPs and reduce their contributions by approximately 80%. As a result, the analysis achieved the most precise ATLAS measurement of m_t in a single decay channel to date, yielding 172.95 ± 0.53 GeV. This value is in excellent agreement with the combined LHC Run 1 result, which remains the most precise m_t determination to date. The reported uncertainty includes a 0.27 GeV statistical component, indicating room for further improvement with larger datasets.

Run 3 is now well underway, with its final year of data-taking approaching and the next long shutdown scheduled to begin in late 2026. Upon completion and full calibration of this data, the largest pp collision dataset ever recorded will be available for analysis. Looking further ahead, ATLAS detector aims to collect up to 3000 fb^{-1} of data with the upcoming HL-LHC program, set to begin in 2030. This major upgrade of the LHC corresponds to an increase by up to a factor of seven in the peak delivered luminosity expected compared to the current LHC. This remarkable increase will

enable measurements with unprecedented statistical precision, including even more precise determinations of m_t and the exploration of previously inaccessible processes and rare phenomena. To fully exploit the potential of the HL-LHC, major upgrades to both the ATLAS detector and the TDAQ system are essential. Central to this upgrade is the replacement of the current ID with the all-silicon ITk, and the deployment of a fully software-based EF trigger system to succeed the current HLT. While the final architecture of the EF has yet to be chosen, it is expected to be a flexible CPU-based system, potentially enhanced by commercial accelerator technologies. As part of this broader upgrade effort, this thesis also presented the integration of the FTF tracking algorithm into the Athena framework to support simulation of Run 4 data using the ITk geometry. This implementation included a novel, pre-developed GNN-based seeding approach that significantly improves the purity of the track seeds generated by the FTF, thereby reducing unnecessary CPU usage. Additional refinements were made to ensure that the EF trigger will meet the high-performance reconstruction demands of the HL-LHC era.

Assuming the successful implementation of these upgrades to both the detector and TDAQ system, projections suggest that the m_t measurement described in this thesis could ultimately reach a precision of 0.23 GeV. As the precision of experimental results continues to increase, so too must our efforts to confront the limitations of the SM, refine the theoretical accuracy and precision of MC generators, and resolve the long-standing challenge of the top-quark mass interpretation problem. Collectively, these efforts would bring us closer to unravelling the true nature of the Universe, using one of the most advanced machines ever built to deepen our understanding.

Bibliography

[1] ATLAS Collaboration, *Measurement of the top quark mass with the ATLAS detector using $t\bar{t}$ events with a high transverse momentum top quark*, 2502.18216. Cited on pages iv, 3, 13, and 104.

[2] ATLAS Collaboration, *Projected uncertainty for a high-p_{T} top quark mass measurement with the ATLAS detector at the HL-LHC*, ATL-PHYS-PUB-2025-009, URL: https://cds.cern.ch/record/2927707, 2025. Cited on pages iv and 175.

[3] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508. Cited on page 6.

[4] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321. Cited on page 6.

[5] M. Thomson, *Modern particle physics*. Cambridge University Press, New York, 2013, 10.1017/CBO9781139525367. Cited on pages 6, 7, and 94.

[6] Particle Data Group Collaboration, *Review of particle physics*, *Phys. Rev. D* **110** (2024) 030001. Cited on pages 7, 9, 11, 12, 115, and 202.

[7] Super-Kamiokande Collaboration, *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562. Cited on page 8.

[8] SNO Collaboration, *Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **87** (2001) 071301. Cited on page 8.

[9] SNO Collaboration, *Direct evidence for neutrino flavor transformation from neutral-current interactions in SNO*, *AIP Conf. Proc.* **646** (2002) 43. Cited on page 8.

[10] ATLAS Collaboration, *The quest to discover supersymmetry at the ATLAS*

experiment, *Phys. Rept.* **1116** (2025) 261. Cited on pages 8 and 10.

[11] W. Buchmuller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, *Nucl. Phys. B* **268** (1986) 621. Cited on page 9.

[12] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652. Cited on pages 9 and 11.

[13] M. L. Perl et al., *Evidence for Anomalous Lepton Production in e+ - e- Annihilation*, *Phys. Rev. Lett.* **35** (1975) 1489. Cited on page 9.

[14] E288 Collaboration, *Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions*, *Phys. Rev. Lett.* **39** (1977) 252. Cited on page 9.

[15] C. Campagnari and M. Franklin, *The Discovery of the top quark*, *Rev. Mod. Phys.* **69** (1997) 137 [hep-ex/9608003]. Cited on page 9.

[16] D0 Collaboration, *Observation of the top quark*, *Phys. Rev. Lett.* **74** (1995) 2632 [hep-ex/9503003]. Cited on page 9.

[17] CDF Collaboration, *Observation of top quark production in $\bar{p}p$ collisions*, *Phys. Rev. Lett.* **74** (1995) 2626 [hep-ex/9503002]. Cited on page 9.

[18] ALEPH, CDF, D0, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, Tevatron Electroweak Working Group, SLD Electroweak, Heavy Flavour Groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, CERN-PH-EP-2010-095, URL: https://cds.cern.ch/record/1222431, 2010. Cited on page 9.

[19] Gfitter Group, *The global electroweak fit at NNLO and prospects for the LHC and ILC*, *Eur. Phys. J. C* **74** (2014) 3046 [1407.3792]. Cited on page 9.

[20] J. Haller, A. Hoecker, R. Kogler, K. Mönig, T. Peiffer and J. Stelzer, *Update of the global electroweak fit and constraints on two-Higgs-doublet models*, *Eur. Phys. J. C* **78** (2018) 675 [1803.01853]. Cited on pages 9, 10, and 202.

[21] CMS Collaboration, *Measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. Lett.* **116** (2016) 052002 [1510.05302]. Cited on page 10.

[22] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531. Cited on page 11.

[23] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, *Phys. Rev. D* **2** (1970) 1285. Cited on page 11.

[24] K. Zoch, *Cross-section measurements of top-quark pair production in association with a hard photon at 13 TeV with the ATLAS detector*, CERN-THESIS-2020-081, URL: https://cds.cern.ch/record/2725289, 2020. Cited on pages 12 and 202.

[25] A. H. Hoang, *What is the Top Quark Mass?*, *Ann. Rev. Nucl. Part. Sci.* **70** (2020) 225 [2004.12915]. Cited on pages 12 and 13.

[26] ATLAS collaboration, *A precise interpretation for the top quark mass parameter in ATLAS Monte Carlo simulation*, ATL-PHYS-PUB-2021-034, URL: https://cds.cern.ch/record/2777332, 2021. Cited on pages 12 and 13.

[27] ATLAS Collaboration, *Measurement of the top-quark mass in $t\bar{t}$ + 1-jet events collected with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **11** (2019) 150 [1905.02302]. Cited on page 12.

[28] ATLAS Collaboration, *Measurement of the top-quark mass using a leptonic invariant mass in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, *JHEP* **06** (2023) 019 [2209.00583]. Cited on page 13.

[29] B. Dehnadi, A. H. Hoang, O. L. Jin and V. Mateu, *Top quark mass calibration for Monte Carlo event generators — an update*, *JHEP* **12** (2023) 065 [2309.00547]. Cited on page 13.

[30] A. H. Hoang, S. Plätzer and D. Samitz, *On the Cutoff Dependence of the Quark Mass Parameter in Angular Ordered Parton Showers*, *JHEP* **10** (2018) 200 [1807.06617]. Cited on page 13.

[31] CDF, D0 Collaboration, *Combination of CDF and D0 results on the mass of the top quark using up 9.7 fb^{-1} at the Tevatron*, *arXiv* (2016) [1608.01881]. Cited on pages 13 and 173.

[32] ATLAS, CMS Collaboration, *Combination of Measurements of the Top Quark Mass from Data Collected by the ATLAS and CMS Experiments at s=7 and 8 TeV*, *Phys. Rev. Lett.* **132** (2024) 261902 [2402.08713]. Cited on pages 13 and 173.

[33] CMS Collaboration, *Measurement of the top quark mass using a profile likelihood approach with the lepton + jets final states in proton-proton collisions at $\sqrt{s}$=13 TeV*, *Eur. Phys. J. C* **83** (2023) 963 [2302.01967]. Cited on page 13.

[34] CMS Collaboration, *Measurement of the differential $t\bar{t}$ production cross section as a function of the jet mass and extraction of the top quark mass in hadronic decays of boosted top quarks*, *Eur. Phys. J. C* **83** (2023) 560 [2211.01456]. Cited on pages 13, 104, and 173.

[35] C. D. Anderson, *The Apparent Existence of Easily Deflectable Positives*, *Science* **76** (1932) 238. Cited on page 15.

[36] S. H. Neddermeyer and C. D. Anderson, *Note on the Nature of Cosmic Ray Particles*, *Phys. Rev.* **51** (1937) 884. Cited on page 15.

[37] G. D. Rochester and C. C. Butler, *Evidence for the Existence of New Unstable Elementary Particles*, *Nature* **160** (1947) 855. Cited on page 15.

[38] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001. Cited on pages 15, 17, and 18.

[39] GARGAMELLE NEUTRINO Collaboration, *Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment*, *Phys. Lett. B* **46** (1973) 138. Cited on page 15.

[40] 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. Cited on page 15.

[41] UA2 Collaboration, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, *Phys. Lett. B* **122** (1983) 476. Cited on page 15.

[42] UA1 Collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c^2 at the CERN SPS Collider*, *Phys. Lett. B* **126** (1983) 398. Cited on page 15.

[43] UA2 Collaboration, *Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider*, *Phys. Lett. B* **129** (1983) 130. Cited on page 15.

[44] NA48 Collaboration, *A New measurement of direct CP violation in two pion decays of the neutral kaon*, *Phys. Lett. B* **465** (1999) 335 [hep-ex/9909022]. Cited on page 15.

[45] 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

[1207.7214]. Cited on page 15.

[46] 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 [1207.7235]. Cited on page 15.

[47] E. Mobs, *The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018*, OPEN-PHO-ACCEL-2018-005 , URL: https://cds.cern.ch/record/2636343, 2018. Cited on pages 16 and 202.

[48] D. J. Warner, *Project study for a new 50 MeV linear accelerator for the C. P. S*, CERN-MPS-LINP-73-1 , URL: https://cds.cern.ch/record/414071, 1973. Cited on page 16.

[49] E. Boltezar, H. Haseroth, W. Pirkl, G. Plass, T. R. Sherwood, U. Tallgren et al., *PERFORMANCE OF THE NEW CERN 50-MeV LINAC.*, *IEEE Trans. Nucl. Sci.* **26** (1979) 3674. Cited on page 16.

[50] L. Arnaudon et al., *Linac4 technical design report*, CERN-AB-2006-084 , URL: https://cds.cern.ch/record/1004186, 2006. Cited on page 16.

[51] Study Group for CPS Improvements, *The second stage CPS improvement study: 800 MeV booster synchrotron*, CERN-MPS-INT-DL-B-67-19 , URL: https://cds.cern.ch/record/109292, 1967. Cited on page 16.

[52] CERN, *CERN Rapport annuel 1988*, URL: https://cds.cern.ch/record/1516883, 1989. Cited on page 16.

[53] CERN, *CERN's 25 GeV proton synchroton*, URL: https://cds.cern.ch/record/445034, Geneva, 1960. Cited on page 16.

[54] CERN, *The 300 GeV programme*, URL: https://cds.cern.ch/record/104068, Geneva, 1972. Cited on page 17.

[55] S. Myers, *The LEP collider, from design to approval and commissioning*, CERN-91-08, CERN-YELLOW-91-08, URL: https://cds.cern.ch/record/226776, 1991. Cited on page 17.

[56] 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 [2212.09379]. Cited on pages 17, 72, 108, and 121.

[57] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*,

JINST **3** (2008) S08003. Cited on pages 17, 20, 21, 22, 23, 27, 29, 202, and 203.

[58] CMS Collaboration, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004. Cited on page 17.

[59] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002. Cited on page 17.

[60] LHCb collaboration, *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005. Cited on page 17.

[61] ATLAS collaboration, *Public ATLAS Luminosity Results for Run-3 of the LHC*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun3, 2013, Accessed 09/01/2025. Cited on pages 18, 19, 72, and 202.

[62] ATLAS collaboration, *ATLAS inner detector: Technical design report. Vol. 1*, CERN-LHCC-97-16, ATLAS-TDR-4, URL: https://cds.cern.ch/record/331063, 1997. Cited on page 22.

[63] ATLAS collaboration, *ATLAS inner detector: Technical design report. Vol. 2*, CERN-LHCC-97-17, URL: https://cds.cern.ch/record/331064, 1997. Cited on page 22.

[64] ATLAS collaboration, *ATLAS pixel detector: Technical design report*, CERN-LHCC-98-13, URL: https://cds.cern.ch/record/381263, 1998. Cited on page 23.

[65] M. Capeans, G. Darbo, K. Einsweiler, M. Elsing, T. Flick, M. Garcia-Sciveres et al., *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013, ATLAS-TDR-19, URL: https://cds.cern.ch/record/1291633, 2010. Cited on page 23.

[66] ATLAS IBL Collaboration, *Production and Integration of the ATLAS Insertable B-Layer*, *JINST* **13** (2018) T05008 [1803.00844]. Cited on page 23.

[67] ATLAS collaboration, *Experiment Briefing: Keeping the ATLAS Inner Detector in perfect alignment*, ATLAS-PHOTO-2020-018, URL: https://cds.cern.ch/record/2723878, 2020. Cited on pages 24 and 203.

[68] ATLAS Collaboration, *Operation and performance of the ATLAS semiconductor tracker in LHC Run 2*, *JINST* **17** (2022) P01013 [2109.02591]. Cited on page 24.

[69] ATLAS TRT Collaboration, *The ATLAS transition radiation tracker*, ATL-CONF-2003-012, URL: https://cds.cern.ch/record/686973, 11, 2003. Cited on page 25.

[70] ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance*, *JINST* **3** (2008) P02013. Cited on page 25.

[71] ATLAS Collaboration, *ATLAS calorimeter performance Technical Design Report*, CERN-LHCC-96-40, URL: https://cds.cern.ch/record/331059, 1996. Cited on page 25.

[72] ATLAS collaboration, *ATLAS liquid argon calorimeter: Technical design report*, CERN-LHCC-96-41, URL: https://cds.cern.ch/record/331061, 1996. Cited on pages 26 and 28.

[73] ATLAS collaboration, *ATLAS tile calorimeter: Technical design report*, CERN-LHCC-96-42, URL: https://cds.cern.ch/record/331062, 1996. Cited on page 27.

[74] ATLAS collaboration, *ATLAS muon spectrometer: Technical design report*, CERN-LHCC-97-22, ATLAS-TDR-10, URL: https://cds.cern.ch/record/331068, 1997. Cited on page 28.

[75] M. Livan, *Monitored drift tubes in ATLAS*, *Nucl. Instrum. Meth. A* **384** (1996) 214. Cited on page 28.

[76] T. Argyropoulos et al., *Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance*, *IEEE Trans. Nucl. Sci.* **56** (2009) 1568. Cited on page 28.

[77] D. Boscherini, *Performance and operation of the ATLAS Resistive Plate Chamber system in LHC Run-1*, *JINST* **9** (2014) C12039. Cited on page 28.

[78] K. Nagai, *Thin gap chambers in ATLAS*, *Nucl. Instrum. Meth. A* **384** (1996) 219. Cited on page 28.

[79] ATLAS Collaboration, *Performance of the ATLAS muon trigger in pp collisions at* $\sqrt{s} = 8$ *TeV*, *Eur. Phys. J. C* **75** (2015) 120 [1408.3179]. Cited on pages 29 and 203.

[80] ATLAS TDAQ Collaboration, *The ATLAS Data Acquisition and High Level Trigger system*, *JINST* **11** (2016) P06008. Cited on page 31.

[81] ATLAS collaboration, *ApprovedPlotsDAQ*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsDAQ, accessed 06/02/2025. Cited on pages 32 and 203.

[82] J. Jamieson, *Aiming for the top corner : measuring top-quark pair production in association with additional jets in highly boosted events using the ATLAS detector*, CERN-THESIS-2022-025, URL: https://cds.cern.ch/record/2805713, 2022. Cited on pages 35 and 203.

[83] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and collider physics*, vol. 8. Cambridge University Press, 2, 2011, 10.1017/CBO9780511628788. Cited on page 36.

[84] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438. Cited on page 36.

[85] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641. Cited on page 36.

[86] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) 298. Cited on page 36.

[87] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin and R. S. Thorne, *Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs*, *Eur. Phys. J. C* **81** (2021) 341 [2012.04684]. Cited on page 36.

[88] P. Bärnreuther, M. Czakon and A. Mitov, *Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$*, *Phys. Rev. Lett.* **109** (2012) 132001 [1204.5201]. Cited on pages 37 and 109.

[89] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, *JHEP* **12** (2012) 054 [1207.0236]. Cited on pages 37 and 109.

[90] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, *JHEP* **01** (2013) 080 [1210.6832]. Cited on pages 37 and 109.

[91] M. Czakon, P. Fiedler and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$*, *Phys. Rev. Lett.* **110** (2013) 252004 [1303.6254].

Cited on pages 37 and 109.

[92] M. Czakon, D. Heymes, A. Mitov, D. Pagani, I. Tsinikos and M. Zaro, *Top-pair production at the LHC through NNLO QCD and NLO EW*, *JHEP* **10** (2017) 186 [1705.04105]. Cited on pages 37 and 111.

[93] M. Beneke, P. Falgari, S. Klein and C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, *Nucl. Phys. B* **855** (2012) 695 [1109.1536]. Cited on pages 37 and 109.

[94] M. Cacciari, M. Czakon, M. Mangano, A. Mitov and P. Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, *Phys. Lett. B* **710** (2012) 612 [1111.5869]. Cited on pages 37 and 109.

[95] W. T. Giele and E. W. N. Glover, *Higher order corrections to jet cross-sections in e+ e- annihilation*, *Phys. Rev. D* **46** (1992) 1980. Cited on page 37.

[96] W. T. Giele, E. W. N. Glover and D. A. Kosower, *Higher order corrections to jet cross-sections in hadron colliders*, *Nucl. Phys. B* **403** (1993) 633 [hep-ph/9302225]. Cited on page 37.

[97] S. Catani and M. H. Seymour, *A General algorithm for calculating jet cross-sections in NLO QCD*, *Nucl. Phys. B* **485** (1997) 291 [hep-ph/9605323]. Cited on page 37.

[98] S. Catani, S. Dittmaier, M. H. Seymour and Z. Trocsanyi, *The Dipole formalism for next-to-leading order QCD calculations with massive partons*, *Nucl. Phys. B* **627** (2002) 189 [hep-ph/0201036]. Cited on page 37.

[99] S. Frixione, Z. Kunszt and A. Signer, *Three jet cross-sections to next-to-leading order*, *Nucl. Phys. B* **467** (1996) 399 [hep-ph/9512328]. Cited on page 37.

[100] S. Frixione, *A General approach to jet cross-sections in QCD*, *Nucl. Phys. B* **507** (1997) 295 [hep-ph/9706545]. Cited on page 37.

[101] A. Buckley, C. White and M. White, *Practical Collider Physics*. IOP, 12, 2021, 10.1088/978-0-7503-2444-1. Cited on pages 37, 38, and 41.

[102] A. Buckley et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145 [1101.2599]. Cited on page 39.

[103] J. R. Forshaw, J. Holguin and S. Plätzer, *Building a consistent parton shower*, *JHEP* **09** (2020) 014 [2003.06400]. Cited on page 39.

[104] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196 [1512.01178]. Cited on pages 39, 110, and 112.

[105] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [1410.3012]. Cited on pages 39, 58, 65, 108, 109, and 112.

[106] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, *JHEP* **03** (2008) 038 [0709.1027]. Cited on pages 39 and 111.

[107] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [0709.2092]. Cited on pages 39, 109, and 112.

[108] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *JHEP* **06** (2002) 029 [hep-ph/0204244]. Cited on page 39.

[109] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146]. Cited on pages 39, 109, and 112.

[110] F. Krauss, *Matrix elements and parton showers in hadronic interactions*, *JHEP* **08** (2002) 015 [hep-ph/0205283]. Cited on page 40.

[111] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027 [1207.5030]. Cited on pages 40 and 111.

[112] T. Sjöstrand, *Jet fragmentation of multiparton configurations in a string framework*, *Nuclear Physics B* **248** (1984) 469. Cited on page 40.

[113] M. Bahr et al., *Herwig++ Physics and Manual*, *Eur. Phys. J. C* **58** (2008) 639 [0803.0883]. Cited on pages 40, 65, 110, and 112.

[114] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152. Cited on pages 41, 109, 110, and 112.

[115] ATLAS Collaboration, *Measurements of observables sensitive to colour reconnection in $t\bar{t}$ events with the ATLAS detector at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **83** (2023) 518 [2209.07874]. Cited on page 42.

[116] GEANT4 Collaboration, *GEANT4 - A Simulation Toolkit*, *Nucl. Instrum. Meth. A*

506 (2003) 250. Cited on page 42.

[117] ATLAS collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013, URL: https://cds.cern.ch/record/1300517, 2010. Cited on page 43.

[118] ATLAS Collaboration, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, *Eur. Phys. J. C* **77** (2017) 673 [1704.07983]. Cited on pages 44 and 87.

[119] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, *Nucl. Instrum. Meth. A* **262** (1987) 444. Cited on pages 45 and 87.

[120] ATLAS Collaboration, *A neural network clustering algorithm for the ATLAS silicon pixel detector*, *JINST* **9** (2014) P09009 [1406.7690]. Cited on page 46.

[121] ATLAS Collaboration, *The ATLAS inner detector trigger performance in pp collisions at 13 TeV during LHC Run 2*, *Eur. Phys. J. C* **82** (2022) 206 [2107.02485]. Cited on pages 46, 86, and 87.

[122] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV*, *Eur. Phys. J. C* **76** (2016) 292 [1603.05598]. Cited on pages 46 and 47.

[123] J. Illingworth and J. Kittler, *A survey of the hough transform*, *Computer Vision, Graphics, and Image Processing* **44** (1988) 87. Cited on page 46.

[124] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J. C* **77** (2017) 490 [1603.02934]. Cited on page 48.

[125] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, *JINST* **14** (2019) P12006 [1908.00005]. Cited on pages 49 and 50.

[126] TMVA Collaboration, *TMVA - Toolkit for Multivariate Data Analysis*, physics/0703039. Cited on page 50.

[127] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 2 data*, *JINST* **19** (2024) P02009 [2309.05471]. Cited on pages 50 and 51.

[128] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, *Eur. Phys. J. C* **80** (2020) 47 [1909.00761]. Cited on pages 50, 108, and 113.

[129] 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 [2012.00578]. Cited on pages 51, 52, and 53.

[130] 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 [2212.07338]. Cited on pages 52 and 53.

[131] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, *JINST* **15** (2020) P09015 [2004.13447]. Cited on pages 53, 108, and 113.

[132] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *Eur. Phys. J. C* **77** (2017) 466 [1703.10485]. Cited on pages 53, 54, 56, 64, 68, 85, and 204.

[133] 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 [2007.02645]. Cited on pages 54, 57, 58, 63, 64, 67, 68, 175, and 204.

[134] G. P. Salam, *Towards Jetography*, *Eur. Phys. J. C* **67** (2010) 637 [0906.1833]. Cited on pages 56 and 57.

[135] M. Cacciari, G. P. Salam and G. Soyez, *The anti-*k_t *jet clustering algorithm*, *JHEP* **04** (2008) 063 [0802.1189]. Cited on pages 56, 57, and 69.

[136] 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 [1510.03823]. Cited on page 57.

[137] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, *Phys. Rev. D* **48** (1993) 3160 [hep-ph/9305266]. Cited on page 58.

[138] M. Cacciari and G. P. Salam, *Pileup subtraction using jet areas*, *Phys. Lett. B* **659** (2008) 119 [0707.1378]. Cited on page 58.

[139] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, *Eur. Phys. J. C* **73** (2013) 2304 [1112.6426]. Cited on page 60.

[140] ATLAS Collaboration, *A measurement of the calorimeter response to single hadrons and determination of the jet energy scale uncertainty using LHC Run-1 pp-collision data with the ATLAS detector*, *Eur. Phys. J. C* **77** (2017) 26 [1607.08842]. Cited on pages 62 and 64.

[141] ATLAS Collaboration, *Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC*, *Eur. Phys. J. C* **73** (2013) 2305 [1203.1302]. Cited on pages 62 and 64.

[142] ATLAS Collaboration, *MC-to-MC calibrations*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ApplyJetCalibrationR21#MC_to_MC_calibrations, Accessed 11/03/2025. Cited on pages 62 and 66.

[143] ATLAS Collaboration, *Measurements of jet cross-section ratios in 13 TeV proton-proton collisions with ATLAS*, *Phys. Rev. D* **110** (2024) 072019 [2405.20206]. Cited on pages 63, 64, 66, 117, and 204.

[144] ATLAS Collaboration, *Dependence of the Jet Energy Scale on the Particle Content of Hadronic Jets in the ATLAS Detector Simulation*, ATL-PHYS-PUB-2022-021, URL: https://cds.cern.ch/record/2808016, 2022. Cited on page 65.

[145] SHERPA Collaboration, *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034 [1905.09127]. Cited on pages 65, 111, and 113.

[146] ATLAS Collaboration, *JES Precision Reduced Flavour Uncertainties*, URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetUncertaintiesRel21Summer2018SmallR#JES_Precision_Reduced_Flavour_Un, Accessed 11/03/2025. Cited on page 65.

[147] ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, *Eur. Phys. J. C* **83** (2023) 681 [2211.16345]. Cited on pages 68 and 114.

[148] 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 [1907.05120]. Cited on page 69.

[149] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, *JHEP* **02** (2010) 084 [0912.1342]. Cited on page 69.

[150] B. Nachman, P. Nef, A. Schwartzman, M. Swiatlowski and C. Wanotayaroj, *Jets from Jets: Re-clustering as a tool for large radius jet reconstruction and grooming at the LHC*, *JHEP* **02** (2015) 075 [1407.2922]. Cited on page 69.

[151] ATLAS Collaboration, *Measurements of differential cross-sections in top-quark pair events with a high transverse momentum top quark and limits on beyond the Standard Model contributions to top-quark pair production with the ATLAS detector at $\sqrt{s}$ = 13 TeV*, *JHEP* **06** (2022) 063 [2202.12134]. Cited on pages 69, 104, 107, 116, 121, 125, 140, 152, 153, 161, 162, 163, 164, 165, 166, 172, 210, 212, 213, and 214.

[152] ATLAS Collaboration, *Jet reclustering and close-by effects in ATLAS run II*, ATLAS-CONF-2017-062, URL: https://cds.cern.ch/record/2275649, 2017. Cited on pages 69 and 118.

[153] ATLAS Collaboration, *Measurement of the ATLAS Detector Jet Mass Response using Forward Folding with 80 fb^{-1} of $\sqrt{s} = 13$ TeV pp data*, ATLAS-CONF-2020-022, URL: https://cds.cern.ch/record/2724442, 2020. Cited on pages 69 and 118.

[154] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV*, *Eur. Phys. J. C* **78** (2018) 903 [1802.08168]. Cited on page 70.

[155] CMS Collaboration, *Evidence for Higgs boson decay to a pair of muons*, *JHEP* **01** (2021) 148 [2009.04363]. Cited on page 72.

[156] ATLAS Collaboration, *Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector*, *Phys. Lett. B* **800** (2020) 135103 [1906.02025]. Cited on page 72.

[157] I. Zurbano Fernandez et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*, . Cited on page 72.

[158] ATLAS ITK Collaboration, *ATLAS ITk pixel detector overview*, *Nucl. Instrum. Meth. A* **1070** (2025) 169978. Cited on page 72.

[159] ATLAS Collaboration, *Modelling radiation damage to pixel sensors in the ATLAS detector*, *JINST* **14** (2019) P06012 [1905.03739]. Cited on page 72.

[160] ATLAS Collaboration, *Measurements of sensor radiation damage in the ATLAS inner detector using leakage currents*, *JINST* **16** (2021) P08025 [2106.09287]. Cited

on page 72.

[161] ATLAS ITK Collaboration, *The ATLAS ITk Strip Detector System for the Phase-II LHC Upgrade*, *PoS* **DIS2024** (2025) 270. Cited on page 72.

[162] ATLAS ITK Collaboration, *The ATLAS ITk detector system for the Phase-II LHC upgrade*, *Nucl. Instrum. Meth. A* **1045** (2023) 167597. Cited on pages 73 and 76.

[163] ATLAS Collaboration, *A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade: Technical Design Report*, CERN-LHCC-2020-007, ATLAS-TDR-031, URL: https://cds.cern.ch/record/2719855, 2020. Cited on page 73.

[164] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*, CERN-LHCC-2017-017, ATLAS-TDR-026, URL: https://cds.cern.ch/record/2285580, 2017. Cited on page 73.

[165] ATLAS collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System*, CERN-LHCC-2017-020, ATLAS-TDR-029, URL: https://cds.cern.ch/record/2285584, 2017. Cited on pages 73, 78, 79, 80, 81, and 94.

[166] ATLAS collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Trigger and Data Acquisition System - Event Filter Tracking Amendment*, CERN-LHCC-2022-004, ATLAS-TDR-029-ADD-1, URL: https://cds.cern.ch/record/2802799, 2022. Cited on pages 73, 78, 79, 81, 83, 85, and 204.

[167] ATLAS collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, CERN-LHCC-2017-021, ATLAS-TDR-030, URL: https://cds.cern.ch/record/2285585, 2017. Cited on pages 74, 75, and 76.

[168] ATLAS collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, CERN-LHCC-2017-005, ATLAS-TDR-025, URL: https://cds.cern.ch/record/2257755, 2017. Cited on pages 74 and 76.

[169] ATLAS Collaboration, *Expected tracking performance of the ATLAS Inner Tracker at the High-Luminosity LHC*, *JINST* **20** (2025) P02018 [2412.15090]. Cited on pages 74, 75, 77, and 204.

[170] C. Krause, *Test beam performance of pixel detectors for the Inner Tracker Upgrade of the ATLAS experiment*, ATL-ITK-PROC-2024-002, URL:

https://cds.cern.ch/record/2888827, 2024. Cited on page 76.

[171] ATLAS collaboration, *New Small Wheel Technical Design Report*, CERN-LHCC-2013-006, ATLAS-TDR-020, CERN-LHCC-2013-006, ATLAS-TDR-020, URL: https://cds.cern.ch/record/1552862, 2013. Cited on page 80.

[172] ATLAS TDAQ Collaboration, *FELIX: the Detector Interface for the ATLAS Experiment at CERN*, *EPJ Web Conf.* **251** (2021) 04006. Cited on page 80.

[173] ATLAS collaboration, *Athena*, Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena, Accessed 17/04/2025. Cited on page 83.

[174] ATLAS collaboration, *ATLAS Software Documentation: Athena*, URL: https://atlas-software.docs.cern.ch/athena, Accessed 17/04/2025. Cited on pages 83 and 91.

[175] ATLAS Collaboration, *The ATLAS trigger system for LHC Run 3 and trigger performance in 2022*, *JINST* **19** (2024) P06029 [2401.06630]. Cited on pages 86, 87, and 88.

[176] P. Yepes, *A fast track pattern recognition*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **380** (1996) 582. Cited on page 87.

[177] ATLAS Collaboration, *Machine learning predictor for 'measurement-to-track' association for the ATLAS Inner Detector trigger*, *J. Phys. Conf. Ser.* **2438** (2023) 012101. Cited on pages 87 and 89.

[178] S. Amrouche et al., *The Tracking Machine Learning Challenge: Throughput Phase*, *Comput. Softw. Big Sci.* **7** (2023) 1 [2105.01160]. Cited on page 88.

[179] I. Abt, I. Kisel, S. Masciocchi and D. Emelyanov, *CATS: A cellular automaton for tracking in silicon for the HERA-B vertex detector*, *Nucl. Instrum. Meth. A* **489** (2002) 389. Cited on page 89.

[180] E. Watton, *ITk seeding for TrigFastTrackFinder*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/56436, 2022, Accessed 08/04/2025. Cited on pages 92 and 98.

[181] E. Watton, *New CA based trigger test for ITk*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/57318, 2022, Accessed 08/04/2025. Cited on page 92.

[182] W. Lampl, *A new approach for ATLAS Athena job configuration*, *EPJ Web Conf.* **214** (2019) 05015. Cited on page 92.

[183] ATLAS collaboration, *Fast Track Reconstruction for HL-LHC*, ATL-PHYS-PUB-2019-041, URL: https://cds.cern.ch/record/2693670, 2019. Cited on page 93.

[184] D. Emeliyanov, *EFTRACK-133, optimized settings, adaptive phi sliding windows, removed un-used variables*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/60095, 2023, Accessed 08/04/2025. Cited on page 95.

[185] D. Emeliyanov, *EFTRACK-133, additional CPU time optimizations, a new control flag to switch track refit*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/60362, 2023, Accessed 08/04/2025. Cited on page 95.

[186] E. Watton, *Use eta dependent pT cut for trigger implementation for FTF ITk*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/72449, 2024, Accessed 08/04/2025. Cited on page 97.

[187] E. Watton, *Optimisations for FastTrackFinder running as an offline algorithm for ITk*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/63391, 2023, Accessed 08/04/2025. Cited on page 97.

[188] E. Watton, *Use new tracking tools for default offline FTF and trigger FTF for ITk*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/72450, 2024, Accessed 08/04/2025. Cited on page 98.

[189] E. Watton, *Adding new nightly trigger test to run FTF with a Run 4 RDO*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/57984, 2022,

Accessed 08/04/2025. Cited on page 98.

[190] E. Watton, *New ART test for FTF as an offline algorithm motivated by work in EFTRACK-7*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/60492, 2023, Accessed 08/04/2025. Cited on page 98.

[191] E. Watton, *New ART test for FTF with ITk using ttbar 200 pileup dataset, modified single muon FTF with ITk ART test*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/60681, 2023, Accessed 08/04/2025. Cited on page 98.

[192] E. Watton, *New ttbar + PU200 nightly trigger test for FTF with ITk*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/66969, 2023, Accessed 08/04/2025. Cited on page 98.

[193] E. Watton, *New nightly trigger tests for ITk using TIDA*, Merge request for Athena Gitlab repository, URL: https://gitlab.cern.ch/atlas/athena/-/merge_requests/68090, 2024, Accessed 08/04/2025. Cited on page 98.

[194] CMS Collaboration, *Measurement of the jet mass in highly boosted* $\mathrm{t\bar{t}}$ *events from pp collisions at* $\sqrt{s} = 8$ *TeV*, *Eur. Phys. J. C* **77** (2017) 467 [1703.06330]. Cited on page 104.

[195] CMS Collaboration, *Measurement of the Jet Mass Distribution and Top Quark Mass in Hadronic Decays of Boosted Top Quarks in pp Collisions at* $\sqrt{s} =$ *TeV*, *Phys. Rev. Lett.* **124** (2020) 202001 [1911.03800]. Cited on page 104.

[196] ATLAS Collaboration, *Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, *JINST* **18** (2023) T11004 [2211.16178]. Cited on page 107.

[197] ATLAS collaboration, *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, ATLAS-CONF-2014-058, URL: https://cds.cern.ch/record/1951336, 2014. Cited on page 107.

[198] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, *JINST* **13** (2018) P07017. Cited on page 108.

[199] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, *JINST* **15** (2020) P04003 [1911.04632]. Cited on page 108.

[200] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823 [1005.4568]. Cited on page 108.

[201] 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, URL: https://cds.cern.ch/record/2206965, 2016. Cited on page 108.

[202] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244 [1207.1303]. Cited on pages 108, 109, and 112.

[203] ATLAS Collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, *Phys. Rev. Lett.* **117** (2016) 182002 [1606.02625]. Cited on page 108.

[204] 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 [1112.5675]. Cited on page 109.

[205] S. Frixione, P. Nason and G. Ridolfi, *A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, *JHEP* **09** (2007) 126 [0707.3088]. Cited on page 109.

[206] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [1002.2581]. Cited on pages 109 and 112.

[207] NNPDF Collaboration, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040 [1410.8849]. Cited on pages 109, 111, and 112.

[208] ATLAS collaboration, *Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMCNLO*, ATL-PHYS-PUB-2017-007, URL: https://cds.cern.ch/record/2261938, 2017. Cited on pages 109, 110, 112, 119, and 120.

[209] ATLAS collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, URL: https://cds.cern.ch/record/1966419, 2014. Cited on pages 109, 111, and 112.

[210] ATLAS collaboration, *Studies on the improvement of the matching uncertainty definition in top-quark processes simulated with Powheg+Pythia 8*, ATL-PHYS-PUB-2023-029, URL: https://cds.cern.ch/record/2872787, 2023. Cited on page 110.

[211] J. Bellm et al., *Herwig 7.1 Release Note*, 1705.06919. Cited on page 110.

[212] J. Bellm et al., *Herwig 7.2 release note*, *Eur. Phys. J. C* **80** (2020) 452 [1912.06509]. Cited on page 110.

[213] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *Eur. Phys. J. C* **75** (2015) 204 [1412.3989]. Cited on pages 110 and 112.

[214] ATLAS collaboration, *A study of different colour reconnection settings for Pythia8 generator using underlying event observables*, ATL-PHYS-PUB-2017-008, URL: https://cds.cern.ch/record/2262253, 2017. Cited on page 110.

[215] T. Gleisberg and S. Hoeche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039 [0808.3674]. Cited on page 111.

[216] F. Buccioni, J.-N. Lang, J. M. Lindert, P. Maierhöfer, S. Pozzorini, H. Zhang et al., *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) 866 [1907.13071]. Cited on page 111.

[217] F. Cascioli, P. Maierhofer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601 [1111.5206]. Cited on page 111.

[218] A. Denner, S. Dittmaier and L. Hofer, *Collier: a fortran-based Complex One-Loop LIbrary in Extended Regularizations*, *Comput. Phys. Commun.* **212** (2017) 220 [1604.06792]. Cited on page 111.

[219] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *JHEP* **09** (2012) 049 [1111.1220]. Cited on page 111.

[220] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063 [hep-ph/0109231]. Cited on page 111.

[221] S. Hoeche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, *JHEP* **05** (2009) 053 [0903.1219]. Cited on page 111.

[222] M. Pinamonti and L. Serkin, *TTbarNNLOReweighter package*, URL: https://gitlab.cern.ch/pinamont/TTbarNNLOReweighter, 2021. Cited on

page 111.

[223] NNPDF Collaboration, *Parton distributions with QED corrections*, *Nucl. Phys. B* **877** (2013) 290 [1308.0598]. Cited on page 111.

[224] S. Catani, S. Devoto, M. Grazzini, S. Kallweit and J. Mazzitelli, *Top-quark pair production at the LHC: Fully differential QCD predictions at NNLO*, *JHEP* **07** (2019) 100 [1906.06535]. Cited on page 111.

[225] M. Grazzini, S. Kallweit and M. Wiesemann, *Fully differential NNLO computations with MATRIX*, *Eur. Phys. J. C* **78** (2018) 537 [1711.06631]. Cited on page 111.

[226] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, *Eur. Phys. J. C* **71** (2011) 1547 [1009.2450]. Cited on page 112.

[227] S. Alioli, P. Nason, C. Oleari and E. Re, *NLO single-top production matched with shower in POWHEG: s- and t-channel contributions*, *JHEP* **09** (2009) 111 [0907.4076]. Cited on page 112.

[228] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber and C. D. White, *Single-top hadroproduction in association with a W boson*, *JHEP* **07** (2008) 029 [0805.3067]. Cited on page 112.

[229] ATLAS collaboration, *Studies of $t\bar{t}/tW$ interference effects in $b\bar{b}\ell^+\ell^{-\prime}\nu\bar{\nu}'$ final states with Powheg and MG5_aMC@NLO setups*, ATL-PHYS-PUB-2021-042, URL: https://cds.cern.ch/record/2792254, 2021. Cited on page 112.

[230] J. Campbell, T. Neumann and Z. Sullivan, *Single-top-quark production in the t-channel at NNLO*, *JHEP* **02** (2021) 040 [2012.01574]. Cited on page 112.

[231] N. Kidonakis and N. Yamanaka, *Higher-order corrections for tW production at high-energy hadron colliders*, *JHEP* **05** (2021) 278 [2102.11300]. Cited on page 112.

[232] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer 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 [1405.0301]. Cited on page 112.

[233] C. Anastasiou, L. J. Dixon, K. Melnikov and F. Petriello, *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, *Phys. Rev. D* **69** (2004) 094008 [hep-ph/0312266]. Cited on page 113.

[234] H. Brooks and P. Skands, *Coherent showers in decays of colored resonances*, *Phys. Rev. D* **100** (2019) 076006 [1907.08980]. Cited on page 120.

[235] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys. G* **43** (2016) 023001 [1510.03865]. Cited on page 120.

[236] LHC Higgs Cross Section Working Group Collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, 1610.07922. Cited on page 121.

[237] Buttinger, William, *Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data*, ATL-COM-SOFT-2015-119, URL: https://cds.cern.ch/record/2014726, 2015. Cited on page 121.

[238] ATLAS Collaboration, *Measurements of top-quark pair differential and double-differential cross-sections in the ℓ+jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *Eur. Phys. J. C* **79** (2019) 1028 [1908.07305]. Cited on page 125.

[239] CMS Collaboration, *Measurement of differential $t\bar{t}$ production cross sections in the full kinematic range using lepton+jets events from proton-proton collisions at $\sqrt{s} =$ 13 TeV*, *Phys. Rev. D* **104** (2021) 092013 [2108.02803]. Cited on page 125.

[240] G. Cowan, K. Cranmer, E. Gross and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554 [1007.1727]. Cited on pages 128, 129, 142, 172, and 214.

[241] R. D. Cousins, *Generalization of Chisquare Goodness-of-Fit Test for Binned Data Using Saturated Models, with Application to Histograms*, URL: https://www.physics.ucla.edu/~cousins/stats/cousins_saturated.pdf, 2013. Cited on page 131.

[242] L. Lista, *Statistical Methods for Data Analysis in Particle Physics*, vol. 909. Springer, 2016, 10.1007/978-3-319-20176-4. Cited on pages 131 and 137.

[243] T. Dado, *Description of the different smoothing algorithms*, URL: https://gitlab.cern.ch/atlas-phys/exot/CommonSystSmoothingTool/-/blob/master/doc/algorithm_description.md, Accessed: 04-09-2024. Cited on page 134.

[244] Tomas Dado and others, *Documentation for TRExFitter*, URL: https://trexfitter-docs.web.cern.ch/trexfitter-docs/, Accessed: 04-09-2024. Cited on page 136.

[245] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, *eConf* **C0303241** (2003) MOLT007 [physics/0306116]. Cited on page 136.

[246] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, *Comput. Phys. Commun.* **10** (1975) 343. Cited on pages 136 and 137.

[247] A. Pinto et al., *Uncertainty components in profile likelihood fits*, *Eur. Phys. J. C* **84** (2024) 593 [2307.04007]. Cited on page 137.

[248] ATLAS Collaboration, *New techniques for jet calibration with the ATLAS detector*, *Eur. Phys. J. C* **83** (2023) 761 [2303.17312]. Cited on page 175.

[249] ATLAS Collaboration, *A precise measurement of the jet energy scale derived from single-particle measurements and in situ techniques in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, 2407.15627. Cited on page 175.

List of figures

List of tables